

ARTES MULTIMEDIALES - IUNA

LENGUAJE SONORO II

LOCALIZACIÓN ESPACIAL DEL SONIDO

Dr. Pablo Cetta

Factores psicofísicos que determinan la localización de los eventos sonoros

La psicofísica, al tratar el problema de la localización espacial del sonido, se ha concentrado principalmente en las modificaciones que ocurren en las señales que llegan a nuestros oídos en relación a la dirección y a la distancia del evento sonoro. Los pabellones auditivos, la cabeza, el cuello y el torso actúan de forma combinada, recibiendo los estímulos del campo acústico y transformándolos en su recorrido al oído medio a través de los tímpanos. Cualquier modificación de las características de la señal acústica, relacionada con la posición de la fuente, se convierte en un indicio potencial para la localización.

Suele dividirse a los indicios en dos grupos: *temporales* (cuando su representación en el dominio del tiempo resulta más pertinente para su análisis) y *espectrales* (su representación es más significativa en el dominio de la frecuencia). También es posible clasificarlos como *binaurales* (basados en la comparación de las diferencias percibidas por cada oído) y *monoaurales* (aquellos que percibidos por un único oído, resultan relevantes para la localización).

A partir de los estudios de Lord Rayleigh (1907), se sostuvo por mucho tiempo que derivamos las pistas necesarias para la ubicación de una fuente de las diferencias interaurales de presión sonora y de los tiempos de arribo de las señales acústicas. La diferencia interaural de tiempo parecía aplicable a sonidos cuya frecuencia fuera menor a 1500 Hz, mientras que la diferencia interaural de intensidad resultaba útil por encima de ese valor. Hoy sabemos que esto es cierto, pero sólo si nos referimos a sonidos puros.

La diferencia interaural de nivel de presión sonora (ILD, por *Interaural Level Difference*) se debe principalmente a la acción separadora de la cabeza, que actúa como una pantalla. Para frecuencias cuya longitud de onda es menor que el diámetro de la cabeza no se produce difracción, sino sombra acústica, registrándose una diferencia de nivel que está en función del ángulo de incidencia de la onda.

La diferencia interaural de tiempo (ITD, por *Interaural Time Difference*) también depende del ángulo de posicionamiento de la fuente, y el pequeño retardo entre las señales que arriban a cada uno de los oídos está asociado a la diferencia entre los caminos recorridos.

Investigaciones posteriores aportaron un indicio monoaural de importancia, expresado a través de las modificaciones espectrales de un evento sonoro en relación a su posición. Estas transformaciones se deben principalmente a la acción del pabellón auditivo, que actúa como un filtro.

Durante el análisis de las causas que determinan la formación de una imagen espacial del sonido vamos a considerar, en primer término, los fenómenos relacionados con la audición al aire libre. En estas condiciones, las únicas señales que percibimos son las que parten directamente de la fuente y se propagan a través de ondas progresivas, sin encontrar obstáculos a su paso. Es la audición que experimentamos en una cámara anecoica, y en ciertas situaciones naturales donde la reflexión sobre el suelo se encuentra atenuada. Posteriormente, dedicaremos nuestra atención a la percepción espacial en recintos, considerando sus aspectos particulares.

Así como la frecuencia encuentra su correlato sensitivo en la altura, y la intensidad en la sonoridad, denominaremos –siguiendo a Blauert (1997)- “evento sonoro” a una manifestación perteneciente al terreno de los estímulos físicos, y “evento auditivo” a la sensación vinculada a ese estímulo.

Para ubicar a un evento sonoro en el espacio utilizaremos coordenadas esféricas. Definimos la posición de la fuente a través una distancia (r) y dos ángulos, uno medido sobre el plano horizontal - ángulo de azimut (θ), para el que consideramos 0° al frente- y otro que se eleva sobre este plano y mide la pendiente del vector con origen en el centro de la cabeza del oyente, y extremo en la fuente - ángulo de elevación (φ).

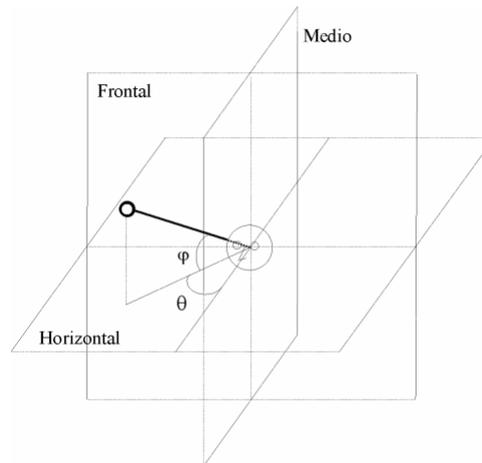


Figura 1 Planos relativos al oyente y ángulos de posicionamiento de la fuente

Funciones de transferencia del oído externo

Según mencionamos antes, el oído externo actúa como un filtro, y como tal puede ser definido a partir de su función de transferencia $H(z)$. Esta describe la relación entre la entrada y la salida.

$$X(z) \longrightarrow \boxed{H(z)} \longrightarrow Y(z) = H(z)X(z)$$

filtro

Figura 2 Función de transferencia

A fin de simplificar, podemos definir a la función de transferencia como la relación entre el espectro de Fourier de la señal de salida y el de la señal de entrada. Ella nos informa en qué medida varía la amplitud y la fase de cada una de las componentes de la señal que ingresa al filtro.

Dado que la multiplicación en el dominio de la frecuencia equivale a la convolución en el dominio del tiempo, podemos también averiguar la señal resultante de un filtro del siguiente modo:

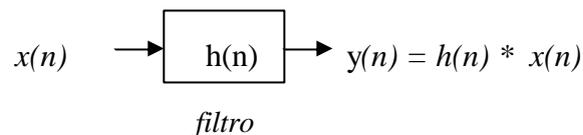


Figura 3 Filtro caracterizado por su respuesta a impulso

$h(n)$ es la respuesta a impulso, o sea la señal que egresa del filtro al ingresar un impulso, y el símbolo $*$ representa a la operación de convolución.

En la figura 4 se observa un gráfico de dos respuestas a impulso, una por cada oído, grabadas con una cabeza artificial, con la fuente ubicada a 30° de azimut y 0° de elevación. El eje horizontal corresponde al tiempo y el vertical a la amplitud. Se aprecian claramente las diferencias entre ambas.

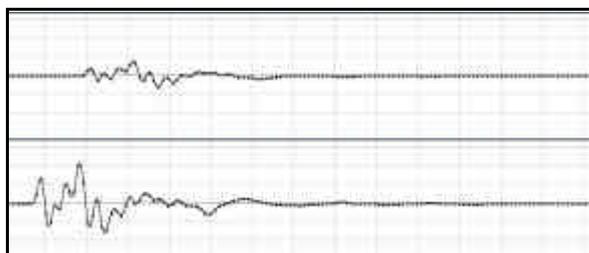


Figura 4 Respuestas a impulso registradas con una cabeza artificial para $\theta = 30^\circ$ y $\varphi = 0^\circ$.

Para clarificar estos conceptos, podemos definir al impulso como una señal discreta cuya primera muestra es unitaria y todo el resto nulo. El impulso posee una respuesta plana en frecuencia, y su utilización nos permite evaluar el comportamiento de un sistema, tanto en los aspectos temporales como espectrales por su brevedad, y porque posee energía en todas las frecuencias del rango audible.

Si registramos un sonido impulsivo en un determinado ambiente –producido por un disparo, por ejemplo- el recinto responderá a ese estímulo con múltiples reflexiones y una coloración espectral particular. Efectuando la convolución entre la respuesta a impulso de la sala y otra señal, grabada de un instrumento musical al aire libre por ejemplo, obtendremos un resultado similar al del instrumento ejecutando en ese ambiente. En este caso, hemos evaluado a la sala como si se tratara de un filtro.

Las funciones de transferencia del oído, reconocidas universalmente a través de la sigla HRTF (por *Head Related Transfer Functions*), dependen de la frecuencia, de la distancia de la fuente, del ángulo de azimut y del ángulo de elevación.

Desde un punto de vista práctico, obtenemos las HRTF de un sujeto o una cabeza artificial ubicando micrófonos en los conductos auditivos y midiendo las respuestas a impulso para diferentes direcciones de la fuente. Estas respuestas, conocidas como HRIR (por *Head Related Impulse Response*) son posteriormente transformadas al dominio de la frecuencia con la Transformada Discreta de Fourier (DFT), y conforman las HRTF medidas para posiciones particulares de la fuente. La representación gráfica de la magnitud de una HRTF nos permite apreciar cómo varía la amplitud en función de la frecuencia para una posición dada de la fuente, una vez que la señal ingresó en el conducto auditivo.

Por otra parte, la comparación de varias HRTF nos da una idea de las transformaciones espectrales que tienen lugar al cambiar la posición del estímulo.

A pesar de la complejidad demostrada por la envolvente espectral de las HRTF, y por la profunda diversidad expuesta en las funciones medidas sobre diferentes sujetos, es posible distinguir ciertas características comunes a todas, según se observa en la figura 5 (Kendall y Martens, 1984). Las curvas corresponden a las HRTF medidas sobre dos sujetos, variando el ángulo de elevación entre 0° y 30° en pasos de 10° , con la fuente ubicada a 10 metros.

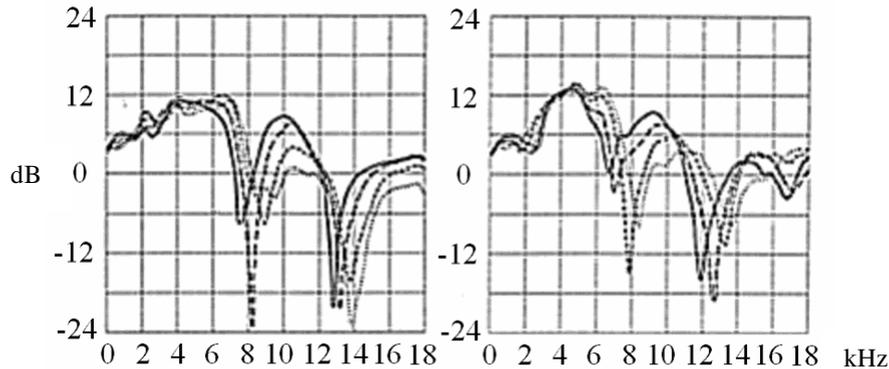


Figura 5 HRTF medidas sobre dos sujetos entre 0° y 30° de elevación, en pasos de 10° (Kendall y Martens, 1984)

La línea sólida corresponde a 0° , los guiones largos a 10° , los cortos a 20° y la línea punteada a 30° . Se observan claramente dos importantes depresiones, que aportarían un indicio válido en función del ángulo de elevación de la fuente.

Indicios binaurales

Los indicios binaurales no dependen de manera crítica de las características de la fuente, o del conocimiento previo que el oyente posea de ella. Resulta razonable creer que guardamos en nuestra memoria las características aprendidas de nuestras propias funciones de transferencia. En la audición practicada con un solo oído, en cambio, debemos conocer de antemano el sonido que esa fuente provoca.

La distancia entre la fuente y el sujeto afecta a las diferencias interaurales. La diferencia interaural de nivel sonoro depende en gran medida de la distancia, mientras que la diferencia interaural de tiempo se ve escasamente afectada.

La diferencia interaural de tiempo es uno de los indicios más importantes en la localización espacial. Experiencias realizadas por Wightman y Kistler (1992) revelaron que la ITD podría ser la pista de localización más sólida. Al presentar señales donde la ITD guiaba a la percepción hacia una posición determinada de la fuente, mientras que los otros indicios apuntaban a posiciones diferentes, los juicios se orientaban hacia aquellas ubicaciones señaladas por la ITD, en tanto el estímulo conservara energía a bajas frecuencias. Sin embargo, cuando las frecuencias por debajo de 1500 Hz eran removidas con un filtro pasa altos, la ILD se tornaba dominante, junto con los indicios espectrales monoaurales.

Si realizamos una experiencia de audición de sonidos impulsivos con auriculares, en la cual aplicamos un mínimo retardo a la señal destinada a uno de los oídos, percibimos un desplazamiento del evento auditivo hacia el oído que recibe la señal sin retardar. Aplicando un tiempo de retardo creciente,

observamos un desplazamiento lineal cuyo valor máximo ocurre a los 630 μ s. La distancia que recorre el sonido en ese tiempo es de 21 cm, y a este valor se lo conoce históricamente como “constante de Hornbostel-Wertheimer”. Hoy sabemos que no se trata de una constante, dado que depende de diversos parámetros. Desde el punto de vista físico, estos 21 cm corresponden a la diferencia de recorrido entre ambos oídos cuando la fuente está lateralizada 90°.

Estudios posteriores demostraron que el desplazamiento máximo se registra únicamente con sonidos cuyo período, dividido dos, está en el orden de los 630 μ s, o sea, para un sonido de unos 800 Hz. Por encima de este valor de frecuencia, el máximo desplazamiento obtenible se torna cada vez menor. Vale decir, que el máximo desplazamiento no ocurre para un retardo de 630 μ s, sino para $T/2$, dado que el desplazamiento depende de la frecuencia. Siempre referidos a los sonidos puros, por encima de 1.6 kHz ya no es posible distinguir ningún desplazamiento causado por la diferencia interaural de tiempo.

El sistema perceptual recurre no sólo a las diferencias provenientes de la estructura interna de los sonidos (retardos de fase), sino también a cambios temporales relacionados con las envolventes de las señales (retardos de grupo). En experiencias comienza a observarse que los corrimientos temporales a nivel de las envolventes, manteniendo en fase a los sonidos, son responsables de los desplazamientos del evento auditivo aún cuando no existen componentes por debajo de los 1600 Hz.

Hemos visto, entonces, que existen dos clases de diferencias interaurales de tiempo. El sistema auditivo puede interpretar diferencias temporales tanto a nivel de portadoras como de envolventes. Las diferencias evaluadas sobre las portadoras tienen efecto sólo por debajo de los 1600 Hz, y las de envolvente surten menos efecto a medida que la frecuencia decrece. Las diferencias interaurales de nivel de presión sonora, en cambio, son efectivas en todo el registro audible, siempre dependiendo del tipo de señal.

Tanto la ITD como la ILD contribuyen conjuntamente a la lateralización del evento auditivo. Las evaluaciones interaurales, no obstante, poseen en sí mismas algunas limitaciones. Cuando la fuente se ubica en el plano medio las diferencias son nulas, y en otros casos la diferencia interaural es la misma para varias posiciones. La figura 6 muestra el denominado “cono de confusión”, en el cual la resta de las distancias entre cualquier punto ubicado en su superficie, y los puntos que representan a los oídos, es constante.

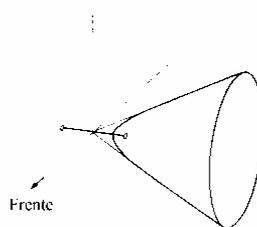


Figura 6 Cono de confusión

Estas ambigüedades generan divergencias en la localización, que es preciso resolver a partir de otros mecanismos.

Indicios espectrales monoaurales

Aún las más pequeñas alteraciones de las señales que llegan a los canales auditivos pueden producir notables alteraciones en la imagen espacial. Acústicamente hablando, el pabellón auditivo se comporta como un filtro lineal que afecta fundamentalmente a las altas frecuencias. Produce una distorsión en la señal, en relación al ángulo de incidencia y a la distancia, codificándola con atributos temporales y espectrales. El efecto acústico del pabellón se basa en los fenómenos físicos de reflexión, difracción, dispersión, sombra acústica, interferencia y resonancia, y contribuye tanto a la correcta discriminación entre frente y atrás, como a la detección del grado de elevación de la fuente.

Por mucho tiempo se consideró que los pabellones auditivos no tenían un rol importante en la audición, considerándolos meros protectores del sistema auditivo. En la actualidad sabemos que cumplen una función determinante en la audición espacial, además de servir a la eliminación del ruido del viento. El canal auditivo, ligeramente curvo y con una longitud promedio de 25 mm, parte de la concavidad central del pabellón y se extiende más allá del tímpano, siempre con un diámetro variado cuyo promedio es de 7-8 mm. El oído externo finaliza en el tímpano, que es ligeramente elíptico y dispuesto a 40-50° respecto al eje del canal. La mayor parte de la información que ingresa al oído interno pasa por el tímpano y el oído medio, mientras que otra se transmite a partir de la conducción del hueso temporal, y se le atribuye una importancia secundaria en el estudio de la audición espacial.

Batteau (1967) propuso que las reflexiones producidas en las convoluciones del oído externo eran capaces de brindar indicios adicionales para la localización. Sugirió que los retardos temporales entre el sonido directo y esas reflexiones variaban en función del ángulo de incidencia, transformando así a las señales que tomaban contacto con el tímpano. El proceso era llevado a cabo mediante la evaluación de la delicada estructura temporal resultante, que se observa al apreciar los picos de una respuesta a impulso del oído externo (figura. 1.4). La figura 7 muestra un esquema del modelo de Batteau, y las funciones que atribuyó a las componentes del pabellón auditivo.

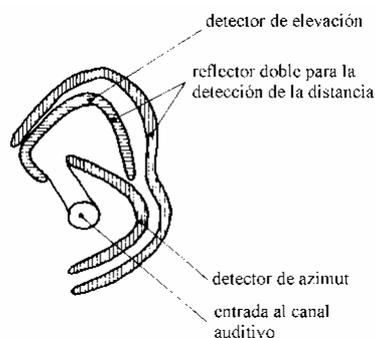


Figura 7 Funciones del pabellón auditivo, según Batteau (1967)

Estudios posteriores determinaron que las diferencias temporales percibidas monoauralmente no eran relevantes para la detección de la posición de la fuente. Las HRTF son sumamente breves, alrededor de 2 ms, y la resolución temporal del sistema auditivo se halla en ese orden. La estructura temporal de una respuesta a impulso resultaría indetectable.

Partiendo de las observaciones de Batteau, los investigadores se han concentrado más en las modificaciones espectrales derivadas de este fenómeno, que en las transformaciones en el dominio del tiempo. Las convoluciones del oído externo funcionan como resonadores, transformando el espectro del sonido de la fuente en función del ángulo de incidencia.

Mientras más ancha es la banda de frecuencias del espectro, mayor es la precisión en la localización. Pero los indicios espectrales percibidos monoauralmente son los menos confiables y parecen depender del conocimiento previo de las características de la fuente. Mediciones realizadas con espectros sistemáticamente alterados han dado como resultado un alto grado de confusión en la localización, tanto en la detección frente-atrás como en el ángulo de elevación, lo que los valida en tanto indicios. Si las diferencias interaurales fueran las únicas responsables de una correcta localización, la alteración del espectro de la fuente no debería influir negativamente. Cuando el sistema perceptual reconoce diferencias interaurales, determina la ubicación de la emisión del estímulo dando un valor más importante a estas diferencias que a los indicios espectrales.

Correcciones asociadas a los movimientos de la cabeza

Las teorías motoras han estudiado en detalle las relaciones entre la posición del evento auditivo y los cambios de las señales de entrada a los oídos durante los movimientos de la cabeza. Estos movimientos, producidos en el momento de la detección de la posición de la fuente, constituyen un indicio adicional de importancia.

Cuando un oyente mueve su cabeza en la dirección del evento auditivo, buscando la posición probable del evento sonoro, decrece el nivel de indeterminación dado que el mayor grado de precisión se encuentra en la zona frontal. Van Soest (1929) fue el primero en señalar que la resolución de las imprecisiones por falta de indicios concretos, o conflicto entre ellos, se realiza a través de los movimientos de la cabeza.

Vimos que cuando la fuente se ubica sobre el plano medio las diferencias interaurales son prácticamente nulas. En este caso, el movimiento permite generar una diferencia que sirve, en principio, para definir si el movimiento proviene del frente o de atrás. La figura 8 puede aclarar este punto. Cuando giramos la cabeza hacia la derecha, si la señal arriba antes al oído derecho, determinamos que la fuente se encuentra atrás.

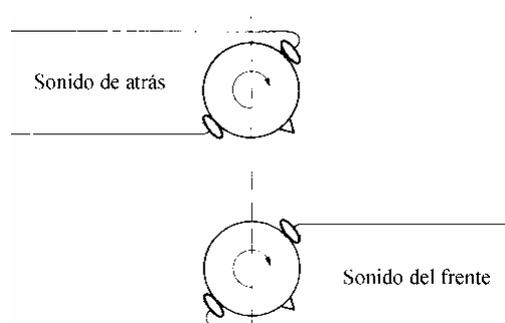


Figura 8 Indicio aportado por la rotación de la cabeza.

Van Soest atribuye al sistema perceptual la capacidad de reconocer la polaridad de la ITD como elemento aclaratorio de la posición de la fuente. Para este tipo de evaluación es necesario tener presente el sentido de la rotación realizada, que podría estar indicado visualmente, a través del órgano vestibular del equilibrio, o por el estado de los músculos del cuello.

Una anomalía en la detección de la posición de la fuente sonora, especialmente con bandas estrechas, se manifiesta en la percepción del evento en una dirección simétrica respecto al eje que cruza a ambos oídos. Una fuente ubicada en un ángulo de incidencia de 30° sobre el plano horizontal puede ser

percibida a 150° (ver figura 9). Como vimos antes, el sistema auditivo recurre a las características espectrales para resolver conflictos entre ambas direcciones, pero cuando los sonidos son de banda angosta, esta información es deficiente o nula, dando lugar a una falla en la localización. Una manera natural de resolver esta ambigüedad es mediante el movimiento de la cabeza durante la emisión del sonido. En los casos donde la localización es imprecisa el movimiento exploratorio de la cabeza ayuda a resolver el problema.

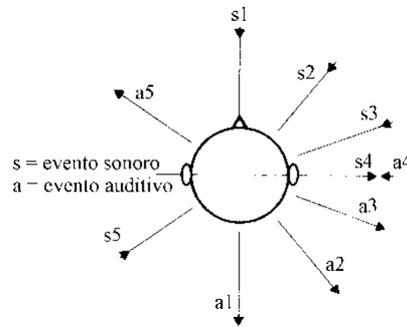


Figura 9 Percepción simétrica respecto al eje interaural

Divergencias en la posición del evento sonoro y del evento auditivo

No siempre la posición del evento sonoro y la del evento auditivo coinciden. Bajo ciertas circunstancias la localización puede ser ambigua, dado que el espacio físico y el auditivo no son necesariamente idénticos. La percepción de la localización puede variar de acuerdo a diversos motivos, entre ellos, por el tipo de señal, por la comparación con eventos anteriores, por la experiencia adquirida en la escucha repetida, por la familiaridad con el sonido de la fuente, por las expectativas del oyente.

Localización indefinida

El concepto de localización indefinida surge de considerar que el espacio auditivo es menos diferenciado que el espacio en el cual los eventos sonoros se manifiestan. El sistema auditivo posee menos resolución espacial que cualquier medición física.

La zona de mayor precisión en la ubicación de la fuente sonora ocurre en la región cercana al punto ubicado justo enfrente del oyente. Es la zona donde un mínimo desplazamiento de la fuente produce una modificación en el evento auditivo. El límite inferior corresponde aproximadamente a un ángulo de 1° , valor elevado si lo comparamos con la resolución en la visión, que es capaz de distinguir variaciones de 1 minuto de arco. Cuando nos desplazamos a derecha o a izquierda, aumenta. Hacia los lados, este valor se incrementa de tres a diez veces, y en la parte posterior decrece nuevamente hasta aproximadamente el doble del valor frontal.

El término “persistencia” suele emplearse en relación al tiempo necesario para determinar la ubicación de la fuente. La posición del evento auditivo sólo puede cambiar con una rapidez limitada. Aschhoff (1963) realizó un experimento utilizando 18 parlantes dispuestos en una circunferencia. A velocidades

bajas de rotación de la señal a través de los parlantes, los oyentes percibían la trayectoria adecuadamente. A mayor velocidad ya sólo podía percibirse el sonido a la derecha y a la izquierda, cíclicamente. Con un crecimiento aún mayor de la velocidad de rotación, el sonido parecía provenir directamente del centro de la cabeza. Como continuación de este experimento, Plath (1970) demostró que la persistencia es menor cuando la fuente alterna entre derecha e izquierda que cuando lo hace entre frente y atrás.

Los valores de precisión en la localización del ángulo de elevación difieren sustancialmente respecto a lo que sucede en el plano horizontal. Cuando la fuente se encuentra sobre el plano medio, las diferencias interaurales son prácticamente nulas, y rara vez ayudan a la interpretación de la señal. El rango de indeterminación cuando la fuente está ubicada frontalmente es de aproximadamente 17° , si la fuente es una voz hablada que no nos resulta familiar (Blauert, 1997). Si nos encontramos familiarizados con su timbre, el error disminuye a unos 9° al frente, pero es de unos 22° cuando se ubica encima de la cabeza (Damaske y Wagener, 1969). El conocimiento previo del sonido propio de la fuente se convierte en un factor determinante de la correcta localización.

Bandas direccionales

Cuando una señal de banda estrecha proviene del plano medio, o es presentada monoauralmente, y las diferencias interaurales quedan anuladas, la ubicación del evento auditivo depende más de la frecuencia central de la banda que de la ubicación física del evento sonoro. La frecuencia se constituye como un referente espacial, y el primero en estudiar sistemáticamente este fenómeno fue Pratt (1930).

Blauert, observó que modificando de forma continua la frecuencia de un sonido puro entre 200 Hz y 16 kHz se percibía un movimiento aparente que partía del frente y llegaba hasta atrás, pasando por arriba de la cabeza varias veces. Esto lo condujo a la realización de otras experiencias para determinar el porcentaje estadístico de localización de eventos auditivos al frente, arriba o atrás, con fuentes de ruido de $1/3$ de octava. El sonido era reproducido a través de cinco parlantes dispuestos de modo tal que ambos oídos recibieran señales idénticas, ya sea a través de uno o más parlantes. Demostró que los juicios sobre la posición de la fuente dependían de la frecuencia central del sonido reproducido y no de la posición del evento sonoro, ni de la configuración de parlantes usada, ni de su intensidad. Denominó a este efecto “bandas direccionales”. Si las bandas se centraban en 4 y 6 kHz aparecían al frente, las de 12 kHz por atrás, y aquellas que se centraban en 8 kHz parecían provenir de arriba de la cabeza, independientemente de la ubicación real de la fuente. Se trata de una ilusión, donde el sonido aparece en un lugar diferente al de la fuente.

Aún se continúa trabajando en esta área, bajo la hipótesis que las bandas direccionales no son sólo una mera curiosidad, sino un indicio probable de localización espacial.

Influencia del grado de coherencia de las señales

Definimos a dos señales como coherentes cuando son idénticas, o bien se diferencian en una o más de las siguientes características:

- Tienen amplitudes diferentes pero la misma forma de onda. Se produce una diferencia de nivel que es independiente de la frecuencia.
- Poseen un retardo de fase que es independiente de la frecuencia
- Se encuentran en oposición de fase

En una experiencia realizada por Chernyak y Dubrovsky (1968) se modificó el grado de coherencia de dos fuentes, presentadas simultáneamente con auriculares, con el propósito de medir la influencia de este parámetro en la representación de los eventos auditivos. Los sujetos fueron provistos con tarjetas, en las que constaba un esquema semicircular representativo del sector frontal superior de la cabeza. Allí debían marcar áreas que describieran la ubicación de los eventos auditivos percibidos. En la figura 10 pueden verse los resultados.

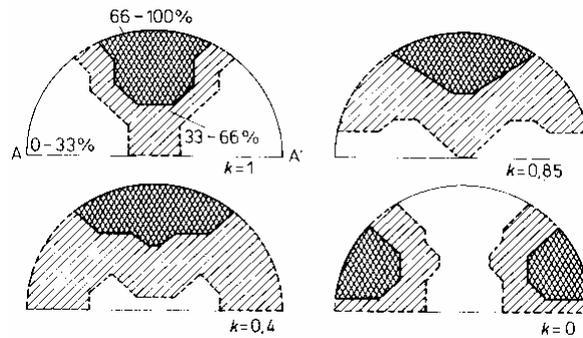


Figura 10 Influencia del nivel de coherencia en la ubicación de los eventos auditivos (Chernyak y Dubrovsky, 1968)

Cuando las señales de las fuentes son totalmente coherentes ($k = 1$) aparece un único evento auditivo, en el plano medio, que ocupa una determinada extensión. A medida que el grado de incoherencia aumenta, también crece el área ocupada por el evento. A un mayor aumento, comienzan a percibirse dos eventos diferenciados, uno en cada oído.

Percepción de la distancia

Con respecto a la determinación de la distancia de una fuente, la familiaridad con el tipo de sonido juega un importante rol en la eficacia de los resultados. La voz hablada a sonoridad normal permite una correspondencia bastante alta entre la posición del evento sonoro y el auditivo. La figura 11 muestra resultados sobre la medición de la distancia subjetiva de una voz hablada -susurrando, hablando normalmente y gritando- reproducida por un parlante que se desplaza de 0,9 a 9 metros (Gardner, 1969).

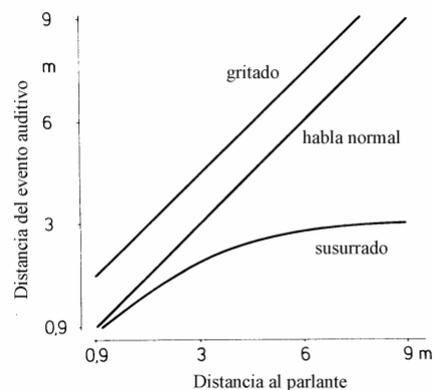


Figura 11 Evaluación de la distancia subjetiva en función de la distancia del evento sonoro, para la voz hablada (Gardner, 1969)

Según Blauert, a distancias intermedias (de 3 a 15 metros), el nivel de presión sonora es el atributo que ayuda a determinar la distancia. Al aire libre, el nivel cae 6 dB cada vez que se duplica la distancia.

Una curva típica de la dependencia de la distancia evaluada respecto al nivel de presión sonora se observa en la figura 12 (Gardner, 1969). Fue medida con un parlante ubicado a 3 metros y luego a 9 metros. Se deduce claramente que la distancia atribuida al evento auditivo no depende de la distancia del evento sonoro, sino únicamente del nivel de presión sonora que se registra en la posición del oyente.

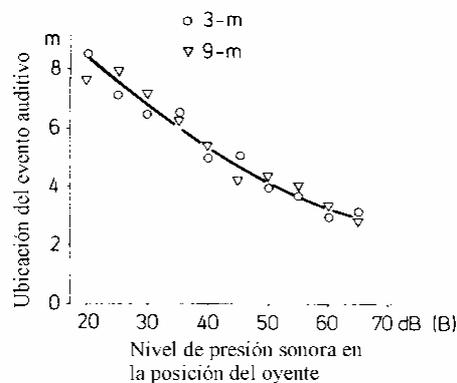


Figura 12 Relación entre la distancia estimada y el nivel de la señal, para dos posiciones fijas de 3 y 9 metros (Gardner, 1969)

Para distancias mayores, el nivel de presión sonora continúa ejerciendo su efecto, pero también la acción del aire, que atenúa las altas frecuencias y modifica el espectro de las señales.

Audición en recintos cerrados

La mayor parte de los estudios sobre los efectos ambientales en la localización se centran en la percepción de las características del recinto, en la relación entre los primeros ecos vinculada al efecto de precedencia, y en la estimación de la distancia cuando actúa la reverberación.

Cuando experimentamos la sensación sonora en un ambiente cerrado llega primero a nuestros oídos el sonido directo, que es el que proviene directamente de la fuente. En segundo término las reflexiones de primer orden, que parten de la fuente sonora, chocan con un objeto (la pared, por ejemplo) y alcanzan nuestros oídos, luego las de segundo orden (con dos reflexiones), y así hasta percibir una sensación difusa denominada reverberación. Las primeras reflexiones –en general seis, si consideramos un cuarto vacío con cuatro paredes, techo y piso- contribuyen con la determinación de la posición de la fuente, principalmente si el ataque del sonido es impulsivo. La reverberación, por su parte, nos brinda información sobre las características materiales de la sala, y sobre sus dimensiones.

Un ambiente acústicamente ideal es aquel donde la reverberación manifiesta una caída lineal, producto de una absorción uniformemente distribuida en un volumen único, y permanece estática. Esto no ocurre cuando existen áreas adyacentes con volúmenes diferenciados o niveles distintos de absorción.

El tiempo de reverberación (t_{60}) se calcula considerando el intervalo transcurrido entre el comienzo de la reverberación y el momento en que su nivel cae 60 dB. El tiempo de reverberación, cuando el volumen es medido en m^3 es, según Sabine:

$$t_{60} = 0.16 \frac{V}{A}$$

donde A representa a la sumatoria de los productos entre los coeficientes de absorción y las superficies reflectoras (el coeficiente de absorción de un piso alfombrado, por ejemplo, es de 0.05 sabins).

El gráfico que sigue (figura 13), muestra la distribución de las reflexiones en el tiempo producidas por un impulso dentro de una sala, la altura de las líneas representa su amplitud relativa.

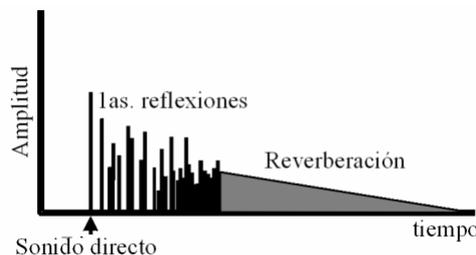


Figura 13 Respuesta típica de una sala

Si bien las reflexiones llegan desde diferentes direcciones, atribuimos una posición correcta a la fuente, pues éstas se funden con el sonido directo agregando una sensación espacial más completa y una mayor sonoridad. Este fenómeno de supremacía de la dirección del directo frente a la de los ecos es conocido como “efecto de precedencia” y ha sido objeto de amplio estudio en psicoacústica. Su acción es tan marcada que ocurre aún cuando la intensidad de las reflexiones es hasta 10 dB más pronunciada.

Cuando percibimos dos señales coherentes de forma sucesiva (producidas por dos parlantes, por ejemplo), podemos distinguir alguno de estos tres fenómenos, de acuerdo al intervalo temporal entre ambas:

- a- un único evento auditivo que depende de la posición de los dos eventos sonoros.
- b- un evento auditivo cuya posición es la misma que la de una de las fuentes.
- c- dos eventos auditivos cuyas posiciones son las posiciones de las fuentes.

El primer caso ocurre cuando los niveles y tiempos de arribo de las señales difieren ligeramente. En este caso hablamos de “localización aditiva”, dado que la posición del evento auditivo resulta del promedio de las posiciones de los eventos sonoros.

Si las señales difieren en más de 1 ms, aproximadamente, la posición del evento auditivo es determinada por la señal que llega en primer término. Cremer (1948) denominó a este fenómeno “ley del primer frente de onda”, de gran importancia en la audición en recintos cerrados. En algunos tratados la localización aditiva y la ley del primer frente de onda son consideradas parte del mismo fenómeno, y se lo denomina “efecto de precedencia”.

Si el retardo entre las señales excede cierto valor aparecen dos eventos auditivos sucesivos bien diferenciados, donde el segundo es considerado eco del primero. El límite inferior –umbral del eco- es difícil de determinar pues depende de diversos factores, como el tipo de señal, la dirección de incidencia y la intensidad. Los umbrales de eco más pequeños ocurren con los impulsos, y a medida que la intensidad crece el umbral se torna menor. Los sonidos continuantes, en cambio, poseen un umbral considerablemente mayor.

En el efecto de precedencia, las reflexiones cambian el color del sonido y refuerzan su sonoridad, y si bien no son percibidas como eventos separados del sonido directo, su anulación puede ser reconocida con claridad.

Otro tipo de vínculo se establece entre el efecto de precedencia y la percepción de la voz hablada en ambientes reverberantes, y es conocido como *efecto Haas* (Haas, 1951). Básicamente, se observa que las primeras reflexiones no alteran la comprensión de la comunicación verbal, siempre que arriben en un tiempo no demasiado largo. Es más, ayudan a la percepción al crear un refuerzo de la sonoridad. El interés puesto por Haas en la palabra hablada se debe a que la voz es muy sensible a la reverberación, en contraste con la percepción de la música.

Green (1976) describe efectivamente el efecto de integración entre el sonido directo y las primeras reflexiones. Dice que si parados a 1 metro enfrente de una pared, producimos un sonido impulsivo golpeando dos rocas entre sí, se crea una reflexión que arriba unos 6 ms más tarde, pero que no somos capaces de escuchar. Pero por otra parte, si escuchamos con auriculares dos *clicks* separados por 6 ms, ambos se perciben claramente. La supresión de la reflexión sucedida en el ambiente no ocurre en la audición con auriculares, y eso es debido al efecto de precedencia.

Un tercer vínculo lo encontramos en lo que suele denominarse “de-reverberación”. La demostración más conocida de este efecto se produce cuando escuchamos una conversación en un determinado ambiente, mientras la grabamos ubicando el micrófono cercano al punto de escucha. Al reproducirla, notamos que el sonido reverberado interfiere con el directo de la voz, dificultando la comprensión, hecho que en la audición natural no ocurría.

Las características temporales del efecto de precedencia en la localización fueron analizadas principalmente por Blauert (1997). Cuando escuchamos dos clicks provenientes de direcciones distintas, separados por menos de 1 ms, el sistema auditivo percibe al segundo clic integrado con el primero, y determina una localización que es el promedio entre ambas posiciones. Para retardos

ubicados entre 1 y 4 ms el efecto de precedencia actúa fuertemente, sin embargo, crea cierta indefinición en la localización, a la vez que aumenta la sensación de espacio. Luego, entre 5 y 10 ms, el efecto comienza a debilitarse: comenzamos a escuchar dos eventos con posiciones diferenciadas.

Las constantes de tiempo del efecto Haas son bastante distintas. Las reflexiones deterioran la comprensión de la voz hablada cuando están retardadas en más de 50 ms. El máximo poder del efecto Haas se encuentra entre los 10 y 20 ms, es decir, cuando el efecto de precedencia desaparece.

Para el efecto de de-reverberación, en cambio, las constantes de tiempo se extienden a varios segundos.

Si los tres vínculos vistos –localización, efecto Haas y de-reverberación- pertenecen al mismo principio de precedencia, sin duda se trata de un proceso sumamente complejo.

Von Békésy (1971) describió un efecto contrario a la ley del primer frente de onda, que observó frente a un retardo de 70 ms. El evento auditivo evaluado para determinar la posición de la fuente era el segundo y no el primero. Blauert (1997) llamó a este efecto “inhibición del sonido directo”. Puede suceder que la reflexión inhiba al sonido directo parcial o totalmente. Para un retardo de 20 ms, el nivel de la reflexión debe ser unos 40 dB más alto para que ocurra el enmascaramiento. Cuando ambos poseen igual nivel, la inhibición nunca ocurre.

Conclusiones

Ha quedado establecido que los indicios primarios para la determinación de la dirección de una fuente sonora están dados por las diferencias interaurales de tiempo, y de nivel de presión sonora, mientras que los indicios espectrales monoaurales también aportan información, pero fundamentalmente sobre el grado de elevación.

La diferencia interaural de tiempo es extraída por el sistema perceptual, no sólo de las diferencias de fase entre las ondas que llegan a cada oído, sino también de las envolventes dinámicas que las afectan. En el primer caso –retardos de fase- este indicio es efectivo por debajo de 1.6 kHz, mientras que en el segundo –diferencias temporales en el ataque de las envolventes- resulta apropiado para frecuencia mayores, cuyo valor está en función del ancho de banda de las señales.

La diferencia interaural de nivel de presión sonora es un indicio, en general, efectivo en todo el registro, pero el límite inferior de la zona útil varía también con el tipo de señal. Los sonidos graves de banda angosta, por ejemplo, son difíciles de localizar, debido a la difracción. Este fenómeno se produce cuando la longitud de onda es mayor al diámetro de la cabeza, y afecta a las señales –o componentes de una señal- cuyas frecuencias son inferiores a los 1800 Hz.

Los efectos de los pabellones auditivos comienzan a hacerse notorios a partir de los 3 kHz, donde la longitud de onda es comparable a su tamaño. La principal depresión en la curva de respuesta en amplitud del sistema auditivo externo, se manifiesta en la octava entre 6 y 12 kHz.

Los modelos físicos análogos contribuyen a la comprensión de los fenómenos auditivos, y facilitan la formalización de las HRTF. No obstante, es preciso considerar que las funciones de transferencia varían considerablemente de un sujeto a otro, y que cualquier intento de simulación de los indicios de localización espacial del sonido debe tener en cuenta esta particularidad, derivada de las variantes anatómicas del oído externo.

Las mayores dificultades en la localización surgen cuando la fuente se encuentra en el plano medio - donde las diferencias interaurales son prácticamente nulas- o en la superficie del denominado “cono de confusión” -donde las diferencias son constantes. La falta de indicios concretos genera, además, errores frecuentes en la determinación de posiciones al frente y atrás. Por otra parte, el nivel de definición depende de la posición; las zonas más indefinidas ocurren a los lados, y en la percepción del grado de elevación. Un sistema de simulación destinado a la ubicación precisa de fuentes aparentes, debe proveer información adicional que permita resolver estas ambigüedades.

También es importante contemplar las divergencias que ocurren entre la posición de las imágenes espaciales, y la ubicación real de los estímulos sonoros, pues inducen a errores en la localización que afectan el realismo de determinadas situaciones, propias de la simulación.

Bibliografía

- Batteau, D. W. (1967). “The role of the pinna in human localization”. *Proc. Roy. Soc. London. B-168*, 158-180.
- Blauert, J. (1997) *Spatial Hearing. The psychophysics of human sound localization*. The MIT Press. Cambridge, Massachussets.
- Cetta, P. (2003) “Modelos de localización espacial del sonido y su implementación en tiempo real”. *Altura-Timbre-Espacio*. Instituto de Investigación Musicológica “Carlos Vega”. FACM-UCA..
- Chernyak, R. y Dubrovsky, N. (1968). “Pattern of the noise images and the binaural summation of loudness for the different interaural correlation of noise”. *Proceedings, 6th International Congress on Acoustics, Tokyo*, A 3-12.
- Gilkey y Anderson (eds.) (1997). *Binaural and Spatial Hearing in Real and Virtual Environments*. Lawrence Erlbaum Associates. New Jersey.
- Green, D. M. (1976). “An introduction to hearing”. Lawrence Erlbaum Associates, Hilldale, NJ.
- Pratt, C. (1930). “The spatial character of high and low tones”. *J. Exp- Psychol.* 13, 278-285.
- Rayleigh, Lord (1907). “On our perception of sound direction”. *Philos. Mag.* 13, 214-232.
- van Soest, J. (1929). “Directional hearing of sinusoidal sound waves”. *Physica* 9, 271-282.
- von Békésy, G. (1971). “Auditory backward inhibition in concert halls”. *Science* 171, 529-536.
- Wightman, F., y Kistler, D. (1992). “The dominant role of low-frequency interaural time difference in sound localization”. *J.A.S.A.* 91, 1648-1661.