



**RIDAA**  
Repositorio Institucional  
Digital de Acceso Abierto de la  
Universidad Nacional de Quilmes



Universidad  
Nacional  
de Quilmes

Calcagno, Esteban Ramón

# Especialidad y estructura sonora en la música electroacústica



Esta obra está bajo una Licencia Creative Commons Argentina.  
Atribución - No Comercial - Sin Obra Derivada 2.5  
<https://creativecommons.org/licenses/by-nc-nd/2.5/ar/>

Documento descargado de RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes de la Universidad Nacional de Quilmes

*Cita recomendada:*

Calcagno, E. R. (2016). *Especialidad y estructura sonora en la música electroacústica. (Tesis de doctorado). Universidad Nacional de Quilmes, Bernal, Argentina Disponible en RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes <http://ridaa.unq.edu.ar/handle/20.500.11807/198>*

Puede encontrar éste y otros documentos en: <https://ridaa.unq.edu.ar>

## **Especialidad y estructura sonora en la música electroacústica**

**Esteban Ramón Calcagno**

esteban-calca@gmail.com

### **Resumen**

El espacio sonoro, o sea la posibilidad biológica, técnica y estética de percibir la ubicación de un sonido y el entorno acústico en el que este se desarrolla, fue utilizado en los últimos 60 años como un elemento que puede estructurar el discurso musical si el compositor lo desea. El presente trabajo se basó en el estudio de las diferentes áreas que se han encargado de la espacialidad del sonido: la percepción auditiva, a través de experimentos psicoacústicos, la ingeniería de sonido, desde el desarrollo de software y la estética musical, mediante el análisis estético de la espacialización de los sonidos que integran una obra. El resultado es una visión holística del espacio sonoro como elemento estructurante en obras de música electroacústica.

UNIVERSIDAD NACIONAL DE QUILMES

DEPARTAMENTO DE CIENCIAS SOCIALES

LABORATORIO DE ACÚSTICA Y PERCEPCIÓN SONORA  
((LAPSo))

## Tesis Doctoral

Presentada a los efectos de cumplimentar los requerimientos del Doctorado en Ciencias Sociales y  
Humanas

Doctorando:

LIC. ESTEBAN RAMÓN CALCAGNO

Director de Tesis:

DR. OSCAR PABLO DI LISCIA

Co-Director de Tesis:

DR. RAMIRO OSCAR VERGARA

---

ESPACIALIDAD Y ESTRUCTURA SONORA EN LA MÚSICA  
ELECTROACÚSTICA

---



---

---

## Agradecimientos

---

---

Quisiera agradecer y dedicar esta tesis a Flor, por ser el amor de mi vida y estar siempre a mi lado cuando la necesito. Sin ninguna duda esta tesis no sería posible sin su amor y su apoyo.

A mi mama Graciela, por educarme y heredarme su inteligencia. A mi papa Giuseppe, por quererme a su manera y transmitirme su pasión por el trabajo, su honradez y su inteligencia. Ambos tenían razón: tenía que dedicarme a estudiar. A mi hermana Catalina, por su arte y los buenos momentos.

A mi cuñado Juan, por ser un hermano para mí. A Alexa, mi sobrina, por iluminarnos a todos.

A mi familia de la vida: Agustina, Dani, Liliana, Celia, Mariela, Maxi, Hilario, Chiquita y Francisco. Cuando la vida se pone pesada, estar con ellos es un oasis.

A Oscar Pablo Di Liscia, por confiar en mí, por enseñarme muchas cosas, permitirme aprender otras y estar en todos los momentos en los que lo necesite.

A Ramiro Vergara, por ser mi amigo y maestro, por dejarme recorrer juntos este camino y por enseñarme a hacer ciencia. Ramiro es una de las personas más creativas que conocí en mi vida, y me enorgullece trabajar con él todos los días.

A Manuel Eguía, por los consejos, las oportunidades, la enseñanza y haberse animado a crear el Lapso, sin el cual tampoco existiría esta tesis. Siempre recuerdo con alegría el día que Manuel me invito a trabajar y me dio la oportunidad de formarme.

A Oscar Edelstein por haber ideado Teatro Acústico. A mis compañeros del Lapso: Ignacio Spiouzas, Martín Proscia, Pablo Etchemendy, Ezequiel Abregú y Pablo Riera, por las charlas, las canciones y las risas. Soy muy feliz de haber crecido junto a ellos.

A Buenos Aires Sonora: Mariano Cura, Martín Liut, Pablo Chimenti y Hernan Kerlleñevich. Por

dejar me estar, cumplirme varios sueños y hacerme ver que las ideas siempre pueden ser llevadas a cabo, aun las más descabelladas. Hemos compartido grandes momentos, y espero que haya aun más. A mis amigos de la UNQ: Martín Matus, Rafael Subía Valdez, Marcelo Martínez, Ezequiel Abregú, Pablo Bachman, Damian Anache y Emanuel Chiesa, por la honradez de sus actos y el fulgor de su trabajo, que me inspiran todos los días. A mis "ninjas", que siempre estan cuando se los necesita y con quien comparto el amor por la electrónica.

A Dani Hernandez por las enseñanzas y el soporte.

A toda la gente de la Secretaría de Servicios Informáticos de la UNQ (Ale, Big Gus, Cesar, Fernando, Mariano, Ramiro y Ariel), por elegirme, educarme, ser mis amigos e introducirme al maravilloso mundo del Software Libre. A Leo Marina por decirme en el momento justo que si no me iba de Servicios Informáticos para empezar mi carrera de Doctorado, me echaba. Fue el empujón que necesitaba.

A los que corrigieron las diferentes partes de esta tesis: Ramiro Vergara, Oscar Pablo Di Liscia, Ignacio Spiouzas, Martín Liut, Manuel Eguía y Ezequiel Abregú.

A la Universidad Nacional de Quilmes y la Licenciatura en Composición con Medios Electroacústicos por ser mi lugar en el mundo.

A la Educación Pública, sin la cual no habría llegado hasta acá.

---

---

## Índice general

---

<b>Agradecimientos</b>	<b>III</b>
<b>Introducción</b>	<b>2</b>
0.1. General . . . . .	2
0.2. Desarrollos y objetivos de la presente tesis . . . . .	6
<b>Referencias Bibliográficas</b>	<b>9</b>
<b>Primera Parte:</b>	
<b>Revisión del marco conceptual</b>	<b>10</b>
<b>1. Localización horizontal de fuentes sonoras</b>	<b>12</b>
1.1. Mínimo Angulo Audible . . . . .	12
1.2. Diferencias interaurales de tiempo e intensidad . . . . .	14
<b>Referencias Bibliográficas</b>	<b>19</b>
<b>2. Localización vertical de fuentes sonoras</b>	<b>21</b>
2.1. Mínimo Angulo Audible . . . . .	21

2.2. Información provista por la pinna . . . . . 24

2.3. Información provista por el torso . . . . . 26

**Referencias Bibliográficas 29**

**3. Localización en distancia de fuentes sonoras 32**

3.1. Introducción . . . . . 32

3.2. Indicios de percepción auditiva de distancia . . . . . 36

3.3. Intensidad . . . . . 38

3.4. Espectro . . . . . 42

3.5. Familiaridad con el estímulo auditivo . . . . . 43

3.6. Indicios acústicos binaurales . . . . . 45

3.7. Reverberación . . . . . 48

3.8. Visión . . . . . 52

**Referencias Bibliográficas 53**

**4. Enfoques analíticos de la espacialidad en la música electroacústica 57**

4.1. Introducción . . . . . 57

4.2. Gary Kendall – Atributos espaciales y juegos perceptivos . . . . . 58

4.3. Oscar Pablo Di Liscia – Análisis del tratamiento estético de los procedimientos constructivos del espacio . . . . . 65

4.4. Martín Liut – Análisis espacial de la música a partir de la praxis compositiva . . . . . 69

4.5. Dante Grela – La música como un arte temporal y espacial . . . . . 72

**Referencias Bibliográficas 75**



---

<b>5. Técnicas y sistemas de espacialización</b>	<b>77</b>
5.1. Técnicas y sistemas . . . . .	78
5.2. Panorámico de intensidad . . . . .	81
5.3. Panorámico de intensidad basado en vectores - VBAP (Vector Base Amplitude Panning)	84
5.4. Sistemas Surround . . . . .	85
5.5. Reconstrucción Binaural . . . . .	87
5.6. Síntesis del campo sonoro . . . . .	90
5.6.1. Ambisonics . . . . .	90
5.6.2. WFS – Wave Field Synthesis: síntesis de campo de onda . . . . .	93
5.7. Emulación de las características de un ambiente acústico mediante grabación y mo- delado de respuestas impulso . . . . .	94
<b>Referencias Bibliográficas</b>	<b>99</b>
<b>Segunda Parte:</b>	
<b>Trabajo de campo y resultados</b>	<b>101</b>
<b>6. El rol de la visión en la percepción auditiva de distancia</b>	<b>103</b>
6.1. Introducción . . . . .	103
6.2. Percepción de distancia auditiva y visual . . . . .	105
6.3. Efecto de la información visual en la PAD . . . . .	108
6.4. Métodos Generales . . . . .	111
6.5. Entorno experimental . . . . .	112
6.6. Participantes . . . . .	112

---

6.7. Pistas visuales . . . . .	113
6.8. Estímulos visuales . . . . .	114
6.9. Estímulos auditivos . . . . .	114
6.10. Procedimiento general . . . . .	116
6.11. Experimento 1 . . . . .	116
6.11.1. Resultados . . . . .	117
6.12. Experimento 2 . . . . .	120
6.12.1. Resultados . . . . .	121
6.13. Experimento 3 . . . . .	125
6.13.1. Resultados . . . . .	126
6.14. Discusión . . . . .	127
6.15. Relación con los resultados pasados . . . . .	128
<b>Referencias Bibliográficas</b>	<b>134</b>
<b>7. El rol del contenido espectral en la percepción auditiva de distancia</b>	<b>139</b>
7.1. Introducción . . . . .	139
7.2. Relación entre el ancho de banda de un sonido y la localización horizontal . . . . .	142
7.3. Relación entre el ancho de banda de un sonido y la localización vertical . . . . .	144
7.4. Relación entre el ancho de banda espectral y la PAD . . . . .	147
7.5. Métodos Generales . . . . .	149
7.6. Entorno experimental, set up y procedimiento . . . . .	150
7.7. Participantes . . . . .	150
7.8. Estímulos auditivos . . . . .	151

---

7.9. Experimento 1 . . . . .	152
7.9.1. Resultados . . . . .	152
7.10. Experimento 2 . . . . .	153
7.10.1. Resultados . . . . .	154
7.11. Experimento 3 . . . . .	156
7.11.1. Resultados . . . . .	157
7.12. Experimento 4 . . . . .	158
7.12.1. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 500Hz . . . . .	159
7.12.2. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 1KHz . . . . .	161
7.12.3. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 2KHz . . . . .	164
7.12.4. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 4KHz . . . . .	166
7.13. Discusión . . . . .	168
<b>Referencias Bibliográficas</b>	<b>174</b>
<b>8. Desarrollo de aplicaciones informáticas para la espacialización del sonido</b>	<b>178</b>
8.1. La espacialidad en la música electroacústica y el diseño de programas informáticos para su tratamiento . . . . .	178
8.2. Gtkonvolve: una aplicación para realizar convolución múltiple . . . . .	181
8.2.1. Convolución y procesamiento digital de señales de audio . . . . .	182
8.2.2. Convolución lineal . . . . .	183

---

8.2.3.	Convolución Rápida . . . . .	184
8.2.4.	Método Overlap/add . . . . .	186
8.2.5.	Gtkonvolve . . . . .	186
8.2.6.	Desarrollo de la interfaz . . . . .	188
8.2.7.	Lógica de funcionamiento . . . . .	188
8.2.8.	Ejemplo básico de funcionamiento . . . . .	190
8.3.	3DEV: Una herramienta para el diseño de trayectorias de múltiples sonidos direccionales en un espacio en 3 dimensiones . . . . .	192
8.3.1.	Características básicas de 3DEV . . . . .	194
8.3.2.	Descripción del programa 3DEV . . . . .	195
8.3.3.	Ventana de trayectorias espaciales . . . . .	196
8.3.4.	Ventana de archivos de audio . . . . .	198
8.3.5.	Conexión a los motores de audio y espacialización . . . . .	199
8.3.6.	Csound . . . . .	199
8.3.7.	Pure Data . . . . .	201
8.3.8.	Consideraciones finales . . . . .	202
8.4.	GUISPA: control de espacialización de fuentes sonoras en tiempo real . . . . .	202
8.4.1.	Características y desarrollo . . . . .	203
8.4.2.	Interfaz Gráfica de Usuario Principal . . . . .	205
8.4.3.	Interfaz Gráfica de Usuario Secundaria . . . . .	207
8.4.4.	Espacialización de las fuentes sonoras . . . . .	210
8.4.5.	Performance en vivo . . . . .	211
8.4.6.	Mejoras a futuro . . . . .	212

---

<b>Referencias Bibliográficas</b>	<b>214</b>
<b>9. Análisis de la espacialidad del sonido en obras de Música Electroacústica: el caso de Buenos Aires Sonora</b>	<b>217</b>
9.1. Introducción . . . . .	217
9.2. Diseño y construcción del espacio sonoro para sitios específicos . . . . .	218
9.3. Arte sonoro y Arte Para Sitios Específicos . . . . .	219
9.4. Tres obras, tres formas de pensar el espacio musical . . . . .	221
9.5. Oí(r) el ruido . . . . .	223
9.5.1. Descripción general . . . . .	223
9.5.2. Orgánico sonoro . . . . .	225
9.5.3. Sistemas de captura, reproducción y espacialización . . . . .	226
9.5.4. Análisis del espacio como estructura de la obra . . . . .	228
9.5.5. Análisis de los diferentes tipos de FUENTE . . . . .	229
9.5.6. Señales Fuente . . . . .	229
9.5.7. Fuentes conceptuales . . . . .	230
9.5.8. Imágenes Fuente . . . . .	232
9.5.9. Niveles de organización . . . . .	234
9.5.10. Juegos con los agrupamientos perceptivos . . . . .	235
9.6. Ocupación Invisible . . . . .	240
9.6.1. Descripción General . . . . .	240
9.6.2. Diseño del espacio, sistema de reproducción y espacialización . . . . .	241
9.6.3. Sistema de Reproducción . . . . .	242
9.6.4. Sistema de sensado . . . . .	245

---

9.6.5. Diseño de interacción . . . . .	248
9.6.6. Análisis . . . . .	251
9.6.7. Análisis de los diferentes tipos de FUENTE . . . . .	253
9.6.8. Señales Fuente . . . . .	253
9.6.9. Fuentes conceptuales . . . . .	254
9.6.10. Imágenes Fuente . . . . .	256
9.6.11. Niveles de organización espacial . . . . .	256
9.6.12. Juegos con los agrupamientos perceptivos . . . . .	258
9.7. 120 Máquina Lírica . . . . .	263
9.7.1. Descripción General . . . . .	263
9.7.2. Análisis “120 Máquina Lírica” – Escena Final . . . . .	264
9.7.3. Análisis de los diferentes tipos de FUENTE . . . . .	267
9.7.4. Señales Fuente . . . . .	267
9.7.5. Fuentes conceptuales . . . . .	269
9.7.6. Imágenes Fuente . . . . .	270
9.7.7. Niveles de organización . . . . .	271
9.7.8. Juegos con los agrupamientos perceptivos . . . . .	273
9.7.9. Conclusiones . . . . .	278
<b>Referencias Bibliográficas</b>	<b>280</b>
<b>Tercera Parte:</b>	
<b>Conclusiones</b>	<b>281</b>
9.8. Conclusiones sobre el estudio de la percepción auditiva de distancia . . . . .	283

---

9.9. Conclusiones sobre el desarrollo de aplicaciones informáticas dedicadas a la espacialización de fuentes sonoras . . . . .	292
9.10. Conclusiones sobre el análisis del espacio en obras electroacústicas y mixtas . . . . .	296
<b>Referencias Bibliográficas</b>	<b>299</b>
<b>Anexo</b>	<b>300</b>
<b>A. Artículos publicados, presentaciones realizadas y financiamiento obtenido en el marco esta la tesis</b>	<b>302</b>
A.1. Artículos publicados . . . . .	302
A.2. Presentaciones a congresos y conferencias . . . . .	303
A.3. Financiamiento obtenido . . . . .	304





---

---

## Introducción

---

### § 0.1 GENERAL

El sonido se desarrolla siempre en un tiempo y un espacio determinado. El espacio en la música está implícito desde la generación del sonido, pero es necesario que este espacio pueda ser ordenado, reconstruido y moldeado para influir en la forma musical, por lo que la utilización del espacio en la música obedece a razones que exceden a las cuestiones meramente acústicas (Cetta, 2007), aunque sin duda las contiene.

No es igual escuchar un piano en una sala de concierto que escucharlo al aire libre, no es igual percibir que una fuente de sonido se acerca que sentir que se aleja, no es igual un recinto sumamente reverberante que uno anecoico. Las sensaciones serán diferentes, porque de hecho los sonidos estarán contenidos en espacios distintos y generaran sensaciones diferenciadas y, por lo tanto, sus características primordiales serán diferentes, a pesar de que nos encontremos con las mismas alturas, ritmos y texturas. Todos esos parámetros seguirán estando, pero serán modelados por el espacio en el que se desarrolla la música. La influencia de la espacialidad en el arte sonoro varía desde muy leve, casi nula, hasta decisiva: todo depende del peso y el rol que el compositor le haya dado en su obra, de las características espaciales que se implanten en los

sonidos (o de las que ya posean originalmente) y de las posibilidades de los oyentes de decodificar estas características en términos de atributos espaciales (Di Liscia, 2005). El conocimiento de estos atributos es un punto de partida insoslayable para su uso como un elemento estructurante dentro de la música electroacústica.

La intención principal de esta Tesis doctoral es la de indagar en los diferentes campos en que la espacialidad del sonido se ha desarrollado (y aun lo hace), e integrar estos conocimientos como herramientas en composiciones de música electroacústica.

A partir de esto se puede decir que –en la música electroacústica y sus variantes- la espacialidad del sonido posee determinados rasgos que la caracterizan y que estos son investigados en tres áreas que se interconectan: la percepción auditiva, la ingeniería del sonido y la creación artística.

Dentro de las ramas de estudio de la percepción auditiva se puede distinguir la percepción auditiva espacial. Esta área del conocimiento estudia cuales son los rasgos de las señales acústicas que nos hacen pensar o percibir la espacialidad del sonido en un determinado ambiente (Di Liscia, 2010). A grandes rasgos existen investigaciones sobre las características espaciales de entornos acústicos y sobre la localización de fuentes sonoras.

Dentro de la localización de fuentes sonoras por parte de seres humanos, la percepción auditiva de la distancia de una fuente sonora es uno de los campos menos estudiados, quizás debido a la gran cantidad de indicios que involucra. A pesar de que es evidente la importancia de una buena percepción de distancia desde el punto de vista biológico (por ejemplo, en la supervivencia) y hasta musical (diferencias entre sonidos cercanos y lejanos), este aspecto

de la localización de sonido no ha sido tan estudiado como la localización en acimut y elevación. Por ello, un trabajo de campo importante dentro de esta tesis ha sido el estudio de la percepción de distancia auditiva y la forma en que determinados indicios la afectan, con el afán de extender las posibilidades del entendimiento de la percepción espacial. Dichas investigaciones se encuentran presentadas en las siguientes publicaciones: Vergara et al, 2010; Calcagno et al, 2012; Abregú et al, 2012.

Dentro de la ingeniería de sonido, los datos que se obtienen de las investigaciones de percepción auditiva del espacio son utilizados por los ingenieros de sonido para diseñar e implementar técnicas y tecnologías para el tratamiento espacial del sonido. Esta rama de investigación se vio potenciada en los últimos años gracias al desarrollo de la ingeniería del software y los grandes avances del procesamiento digital de señales. Dentro de esta área se pueden encontrar muchas aplicaciones destinadas al trabajo con la espacialidad del sonido, que si bien son utilizadas por músicos electroacústicos, en general no están diseñadas para tal fin. Por ello, se ha trabajado también dentro de esta tesis en el desarrollo de varias aplicaciones informáticas diseñadas especialmente para su uso en la composición de música electroacústica (Gtikonvolve, 3DEV y GUISPA - <http://estebanalcagno.com.ar/programas.html>; Di Liscia y Calcagno, 2010).

La música electroacústica estudia y aplica las características espaciales de determinados ambientes y sonidos a partir del impacto de los recursos mencionados anteriormente, logrando una integración estética de espacio y sonido. Los músicos electroacústicos se forman, en diversos grados de profundidad, en los tres campos del conocimiento antes mencionados, y de esta manera apare-

cen puntos de encuentro entre ellos. Kendall (2010) lo explica muy claramente: En un nivel alto de apreciación, la música electroacústica se vale de las tecnologías de reproducción de sonido (basadas como vimos en las investigaciones perceptivas del espacio sonoro) para lograr la generación de fenómenos irreales. En un nivel más profundo esta paleta espacial es un producto que se encuentra entre la tecnología y la percepción espacial que tiene el oyente. O sea la música electroacústica obtendría así una primera relación con los datos científicos perceptivos, aunque se deben aclarar varios puntos: la percepción de un estímulo simple en condiciones experimentales controladas, y la percepción en el contexto del trasfondo histórico, cultural y social que trae aparejada la música en sí no son equivalentes. Por lo tanto esta relación debe ser investigada, desde el punto de vista musical, con un enfoque holístico, donde también intervienen efectos psicológicos, recuerdos, sentimientos, experiencias y asociaciones. La espacialidad del sonido en la música electroacústica manifiesta un sentido artístico al reformar y reformular patrones normativos de la experiencia espacial. Consecuentemente, la música electroacústica puede resignificar la experiencia espacial creada por la experiencia del oyente en la vida diaria (Kendall, 2010). Harley (1999) aclara que la noción de espacialización dentro de la música requiere un cambio de perspectiva que no invalida ni contradice a la tradición. La espacialidad musical, aunque siempre está presente, no siempre es crucial para la comprensión de una obra de arte en la que las relaciones espaciales sean prescindibles. Cuando el rol de la proyección espacial, la localización sonora y la direccionalidad de una fuente sonora se convierten en elementos importantes de la estructura musical, se puede hablar claramente de espacialización musical (Harley, 2009).

En este marco, el tercer trabajo de campo de esta tesis doctoral presenta el análisis de obras a partir de la relación entre la percepción auditiva espacial de los sonidos utilizados, la descripción de las características acústicas de los lugares donde las obras se desarrollan, las tecnologías de espacialización que se utilizan y las funciones estructurales que determinados atributos espaciales cumplen dentro de la obra.

Para tales fines, se ha tomado como base el enfoque analítico de Gary Kendall (Kendall, 2010) y se lo ha extendido en varios aspectos que se detallarán más adelante. Tales extensiones constituyen otro aporte de la presente Tesis en el área específica del análisis musical aplicado a la música electroacústica.

## § 0.2 DESARROLLOS Y OBJETIVOS DE LA PRESENTE TESIS

En el marco de esta tesis doctoral se presentarán los aportes puntuales que se han hecho sobre cada uno de los campos que estudian la espacialidad del sonido. Estos aportes buscan como fin proporcionar al compositor de música electroacústica los medios tecnológicos y conceptuales para expresar más profundamente la poética del espacio en términos sonoros y generar un entorno apropiado que posibilite la realización de la obra (Cetta, 2007). Por lo antedicho, se ha organizado esta tesis en tres partes principales:

En la primera parte se realiza una revisión de la espacialidad del sonido en las tres áreas antes mencionadas: la percepción espacial del sonido, el análisis del espacio en la música y la emulación de la espacialidad del sonido mediante técnicas y sistemas de espacialización. Dentro del área de percepción de sonido, el Capítulo 1 estará destinado a la revisión de bibliografía sobre la

localización de fuentes sonoras en el plano horizontal (acimut), el Capítulo 2 a la localización en el plano vertical (elevación) y el Capítulo 3 abordará la revisión de la localización de fuentes sonoras en distancia. Debido a que el trabajo sobre percepción auditiva dentro de esta tesis apunta al estudio de la percepción auditiva de distancia, la revisión de este aspecto será mucho más exhaustiva que en los otros aspectos de la localización espacial. Por otra parte, el Capítulo 4 revisará algunas de las metodologías de análisis de la espacialidad como elemento estructural de la música, donde se pondrá especial énfasis en el enfoque desarrollado por Gary Kendall (2010), el que luego se utilizará como herramienta de análisis de las obras que se estudian en esta tesis. El Capítulo 5 de esta primera parte, abordará una revisión de las técnicas y sistemas de espacialización con el fin de sentar las bases de nociones que luego serán utilizadas tanto en la descripción de los sistemas utilizados en las obras aquí analizadas, como en el desarrollo de programas informáticos que se presentan en el marco de este trabajo.

La segunda parte estará destinada a mostrar los resultados puntuales obtenidos en el trabajo de campo de esta tesis. Así, el Capítulo 6 mostrará los resultados obtenidos en experimentos psicofísicos que estudian el rol de la visión en la percepción auditiva de distancia. El Capítulo 7 mostrará también resultados obtenidos a partir de experimentos psicofísicos, pero que estudian la influencia del contenido espectral de un estímulo sonoro sobre la percepción auditiva de distancia.

El Capítulo 8 abarcará el desarrollo de tres aplicaciones informáticas enfocadas en la simulación y el manejo de la espacialidad de fuentes sonoras. Por último, en el Capítulo 9, se llevará a cabo el análisis de la espacialidad del

sonido como elemento estructurante de la música electroacústica a partir de la descripción de tres obras que se destacan por un manejo diferenciado del espacio sonoro. Este último capítulo integrará nociones desarrolladas en los capítulos anteriores con el fin de obtener un análisis exhaustivo de las posibilidades del espacio dentro de la música electroacústica.

Como corolario de esta tesis, en la tercera parte se desarrollarán las conclusiones obtenidas en el marco de esta tesis doctoral, poniendo énfasis en los resultados que vinculan las tres grandes áreas aquí tratadas.

## REFERENCIAS BIBLIOGRÁFICAS

- ABREGÚ E.L., CALCAGNO E.R y R.O., VERGARA: «La distancia como factor estructural de la música». *Revista Argentina de Musicología. En prensa.*, 2012.
- CALCAGNO E.R., ABREGÚ E.L., EGUÍA M.C. y R.O., VERGARA: «The Role of Vision in Auditory Distance Perception». *Perception*, 2012, **41(2)**, pp. 175–192.
- CETTA, P.: *Un modelo para la simulación del espacio en música*. Editorial de la Universidad Católica Argentina, Buenos Aires, Argentina, 2007.
- DI LISCIA, O.P.: «Los modos de vínculo de la concepción espacial del sonido con la poiesis de la música electroacústica». *Tesis Doctoral, Universidad Nacional de Rosario, Facultad de Humanidades y Arte.*, 2005.
- : «Algunas reflexiones sobre la espacialidad del sonido en el marco de la producción discográfica comercial y la música electroacústica». *Revista Letra Imagen y Sonido: La construcción Mediática de la ciudad. FCS-UBA*, 2010, **5**, pp. 158–162.
- DI LISCIA, P.O. y CALCAGNO, E.R.: «3DEV: A tool for the control of multiple directional sound source trajectories in a 3D space». *En Proceedings de la Audio Linux Conference 2010. Utrecht, Holanda.* <http://lac.linuxaudio.org/2010/download/lac2010proceedings.pdf>, 2010, pp. 161–165.
- HARLEY, M.A.: «Spatiality of sound and stream segregation in twentieth century instrumental music». *Organised Sound*, 1999, **3(2)**, pp. 147–66.
- KENDALL, G.: «La interpretación de la espacialización electroacústica: atributos espaciales y esquemas auditivos». *En Música y espacio: ciencia, tecnología y estética. Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 241–259.
- VERGARA O.R., CALCAGNO E.R. y M.C., EGUÍA: «The role of spectral cues and minimum bandwidth in the auditory perception of distance». *J. Acoust. Soc. Am.*, 2010, **128(4)**, pp. 2455–2455.



---

PRIMERA PARTE:  
REVISIÓN DEL MARCO CONCEPTUAL



## - Capítulo 1 -

---

---

### Localización horizontal de fuentes sonoras

---

#### § 1.1 MÍNIMO ANGULO AUDIBLE

El sistema auditivo humano posee mayor resolución espacial en la localización horizontal que en los otros planos del espacio auditivo. La resolución espacial al estimar la posición angular de una fuente sonora se mide en términos del mínimo ángulo audible, conocido como MAA (Mills, 1958). El MAA es la diferencia mínima en el ángulo de incidencia de dos sonidos de idénticas características que un sujeto es capaz de diferenciar. En general, el MAA no tiene un valor absoluto en cada plano del espacio auditivo ya que su valor depende de diferentes factores externos como por ejemplo el ángulo de incidencia y el contenido espectral de los sonidos. En los seres humanos, la resolución espacial en el plano horizontal es mayor cuando la fuente sonora está ubicada al frente de la cabeza del oyente que cuando está ubicada a los costados. De este modo, el MAA es mínimo cuando la fuente se encuentra cerca del plano sagital medio del oyente y aumenta gradualmente cuando los estímulos son presentados desde ángulos con valores de acimut crecientes. Cuando se utilizan estímulos ubicados en la parte frontal de la cabeza, el MAA en el plano

horizontal tiene valores cercanos a  $1^\circ$ . Sin embargo, cuando el ángulo de incidencia del estímulo auditivo aumenta, el valor del MAA también aumenta gradualmente hasta alcanzar valores de  $10^\circ$  cuando los sonidos provienen de un ángulo de  $80^\circ$ . Cuando el ángulo de incidencia del sonido supera los  $80^\circ$ , la resolución espacial disminuye considerablemente alcanzando valores mayores a  $40^\circ$ . Esta caída tan importante en la resolución espacial se debe a que cuando el sonido proviene desde ángulos cercanos a  $90^\circ$  los sujetos no son capaces de diferenciar si éste proviene de una dirección lateral anterior o lateral posterior. Esto hace que, para sonidos provenientes de lugares lateralizados, los oyentes cometan errores mayores en la ubicación espacial de los mismos; esta zona de incerteza es conocida como el “cono de confusión” (Basso, 2006). Generalmente podemos mejorar nuestra estimación en la ubicación de sonidos que provienen con ángulos mayores a  $80^\circ$  moviendo nuestra cabeza para que la fuente quede ubicada en algún sitio fuera de la zona de incerteza donde la resolución espacial es más precisa.

Además del ángulo de incidencia, el MAA depende de las características espectrales del estímulo auditivo. En un trabajo publicado por Stevens y Newman en 1936 se describe como el MAA en el plano horizontal es dependiente del contenido espectral del estímulo. En este trabajo se utilizaron tonos puros como estímulos y se demostró que los seres humanos tenemos más precisión para ubicar el ángulo de incidencia en acimut cuando se utilizan tonos de frecuencias menores a 1 kHz o mayores a 5 kHz que cuando son utilizados tonos de frecuencias intermedias, entre 2 y 4 kHz. Cuando los sonidos provienen desde el frente de la cabeza, el MAA para un tono puro menor a 1 kHz es aproximadamente  $1^\circ$ , y  $1.5^\circ$  para tonos con frecuencias mayores a 5

kHz. Sin embargo, el MAA asciende a  $3^\circ$  para tonos con frecuencias ubicadas entre 2 y 4 kHz. El hecho de que la resolución espacial en acimut descrita por Stevens y Newman sea dependiente de la frecuencia del estímulo se debe a las características que poseen los dos indicios principales que utiliza nuestro sistema nervioso para estimar la ubicación espacial de una fuente sonora en el plano horizontal. Estos indicios son: la Diferencia Interaural de intensidad (del inglés, Interaural Level Difference, ILD) y la Diferencia Interaural de Tiempo (del inglés, Interaural Time Difference, ITD).

### § 1.2 DIFERENCIAS INTERAURALES DE TIEMPO E INTENSIDAD

La diferencia interaural de intensidad (ILD) es la diferencia de intensidad, en decibeles, que tiene la señal auditiva al llegar a cada oído cuando esta proviene de una ubicación espacial con un acimut diferente de  $0^\circ$ . Esta diferencia en la intensidad recibida se produce principalmente por la absorción y reflexión (sombra acústica) que produce la cabeza sobre el sonido en su recorrido desde un oído al otro. La disminución en la intensidad del sonido producida por la sombra acústica de la cabeza depende del ángulo de incidencia con que este llega al oyente. Cuando el sonido proviene desde el frente (acimut  $0^\circ$ ), ambos oídos reciben la señal con la misma intensidad y el ILD es, por lo tanto, cero. A medida que la señal se desplaza hacia zonas con valores de acimut crecientes el valor del ILD va aumentando gradualmente hasta alcanzar su valor máximo cuando el sonido llega con un ángulo de incidencia de  $90^\circ$ . Además del ángulo de llegada del sonido, el ILD depende de la relación entre el tamaño de la cabeza y la longitud de onda de la señal. La primera vez que

se reportó la importancia del ILD y del ITD como indicios binaurales para la ubicación angular de sonidos en el plano horizontal fue en el año 1907 en un trabajo publicado por Lord Rayleigh (1907). Lord Rayleigh presentó una teoría, todavía vigente bajo el nombre de Teoría Dúplex, en la cual se postula que el ILD es el principal indicio que utilizamos para la ubicación espacial (en el plano horizontal) de sonidos de frecuencias altas mientras que el ITD es el principal indicio en la ubicación espacial de sonidos de frecuencias bajas. Rayleigh observó que, para tonos puros de alta frecuencia, las variaciones del ILD producían en los oyentes la ilusión de que la fuente sonora cambiaba de posición dentro del plano horizontal. En estas condiciones bastaban variaciones del ILD de solo 1-2 dB para que los sujetos reportaran diferencias en la ubicación de la fuente en acimut. Rayleigh reportó que la imagen auditiva iba lateralizándose cuando los valores de ILD aumentaban. Cuando la diferencia de intensidades era de alrededor de 15 dB los sujetos reportaban la posición aparente del sonido hacia una posición de azimut cercana a  $90^\circ$ . Sin embargo, cuando se utilizaron como estímulos tonos puros de baja frecuencia (por debajo de 1.5 kHz) los sujetos no eran capaces de percibir la ubicación de la fuente percibían el origen del sonido en una ubicación cercana a  $0^\circ$ . Como vimos más arriba, la sombra acústica responsable del ILD es producida principalmente por la difracción, reflexión y refracción del sonido cuando éste choca con la cabeza del oyente (aunque también influyen la forma del tronco y la pinna). Rayleigh explicó que, el hecho de que los tonos de baja frecuencia no provoquen ILD, se debe a que cuando la longitud de onda de la señal es igual o mayor que el diámetro de la cabeza el sonido se “curva” alrededor de la misma. Este proceso es conocido como difracción, y el resultado es que la

sombra acústica disminuye (hasta, eventualmente, desaparecer) y con ella los valores de ILD. En la década del 70, estudios realizados con micrófonos colocados dentro del canal auditivo de diferentes especies, en una posición cercana a la membrana timpánica, permitieron medir minuciosamente las diferencias en la intensidad del sonido en cada oído en función de: la ubicación de la fuente en acimut, la frecuencia de la señal y el tamaño de la cabeza (Harrison y Downey, 1970). Harrison y Downey utilizaron esta técnica para medir el ILD en seres humanos, ratas, murciélagos y monos. Estos experimentos, confirmando lo propuesto por Rayleigh en la teoría dúplex, mostraron que el ILD es efectivo para las señales con longitudes de onda que son menores que el diámetro de la cabeza del oyente y además que el ILD aumenta cuanto más chica es la longitud de onda de la señal. Sus resultados demostraron que, en los seres humanos, estímulos de frecuencias bajas (por ejemplo, 500 Hz) producen un ILD muy pequeño (cercano a 0 dB) mientras que los estímulos de frecuencias elevadas (por ejemplo, 8 kHz) producen valores de ILD superiores a 20 dB. En los seres humanos, la distancia interaural es de aproximadamente 22 cm y en este caso el ILD es efectivo para frecuencias mayores a 1.5 kHz. Estos experimentos demostraron que el rango de frecuencias en donde el ILD es efectivo es dependiente del tamaño de la cabeza del oyente. Por ejemplo, en algunas especies de murciélagos, con una distancia interaural de solo 2 cm, el ILD es efectivo solamente para frecuencias mayores a 7 kHz. Esto no quiere decir que el ILD en animales con cabeza pequeña sea menor que en animales de mayor tamaño. En estos murciélagos por ejemplo, si se estimula con tonos de frecuencias mayores a 7 kHz, el ILD puede alcanzar valores de hasta 40 dB.

Sabemos entonces que el ILD es el indicio que nos permite ubicar la procedencia en acimut de sonidos con frecuencias mayores a 1.5 kHz. Sin embargo, no solamente somos capaces de ubicar la procedencia de sonidos de alta frecuencia sino que podemos detectar la ubicación, en azimut, de sonidos a lo ancho de todo el espectro de frecuencias audible. ¿Cómo hacemos entonces para detectar la procedencia de sonidos de baja frecuencia en donde el ILD deja de ser efectivo? Rayleigh pensó que, si el ILD no sirve para identificar la procedencia de tonos de baja frecuencia, el único indicio que podía ayudar a ubicar un sonido en el plano horizontal era la diferencia temporal en la llegada de una señal sonora a cada oído cuando llega con valores de acimut diferentes a  $0^\circ$ . Este indicio se conoce como diferencia interaural de tiempo o ITD (del inglés, Interaural Time Difference). De esta manera, el tiempo que tarda la señal en llegar desde la fuente a cada oído es diferente: el oído que está más cerca de la fuente recibirá el sonido antes que el oído que se encuentra más lejos de esta. Cuando el sonido proviene de un ángulo de acimut de  $90^\circ$  el ITD tiene su valor máximo ya que la señal, luego de llegar al oído más próximo, debe recorrer la distancia máxima (dada por el diámetro de la cabeza) para llegar al otro. Por el contrario si la señal proviene desde el frente de la cabeza (acimut cero) la distancia desde la fuente a cada oído es la misma y el valor de ITD es cero. Rayleigh se dio cuenta, además, que la diferencia temporal en la llegada de una señal de baja frecuencia a los oídos haría que cada uno de estos reciba la señal en diferentes momentos de fase de la onda sonora (diferencia de fase). Rayleigh confirmó la sensibilidad de los oyentes al ITD en una serie de experimentos muy ingeniosos en los cuales utilizó un complicado mecanismo que activaba un par de diapasones intercambiables lo cual



le permitía estimular cada oído con diferentes retardos temporales y con diferentes retardos en la fase de la onda sonora. Estos experimentos demostraron que, para tonos puros de frecuencias menores a 1 kHz el ITD es muy efectivo para detectar la procedencia angular de sonidos en el plano horizontal. Sin embargo, cuando la frecuencia del estímulo es aumentada por encima de 1.3 kHz el ITD se hace menos efectivo y los sujetos muestran una reducción en su capacidad de reconocer la ubicación de la fuente (la imagen auditiva tiende a moverse hacia el centro de la cabeza) para perderla totalmente cuando la señal alcanza valores de frecuencia mayores a 1.5 kHz. Esta reducción en la efectividad del ITD corresponde aproximadamente con el valor de frecuencia para el cual la diferencia de fase entre los oídos comienza a ser mayor a un ciclo de la onda sonora. Esto genera una ambigüedad en la diferencia de fase ya que resulta imposible decidir, por ejemplo, si la diferencia es de medio ciclo o de un ciclo y medio, lo que podría estar relacionado con la reducción mencionada.

## REFERENCIAS BIBLIOGRÁFICAS

- BASSO, G.: *Percepción Auditiva*. Editorial de la Universidad Nacional de Quilmes, Quilmes, Argentina, 2006.
- HARRISON, J.M. y DOWNEY, P.: «Intensity changes at the ear as a function of the azimuth of a tone source: a comparative study». *J. Acoust. Soc. Am.*, 1970, **47**, pp. 1509–1518.
- MILLS, A.W.: «On the Minimum Audible Angle». *J. Acoust. Soc. Am.*, 1958, **30(4)**, pp. 237–246.
- RAYLEIGH, L.: «On our perception of sound direction». *Philosophical Magazine Series 6*, 1907, **13(74)**, pp. 214–232.
- STEVENS, S.S. y NEWMAN, E.B.: «The localization of actual sources of sound». *The American Journal of Psychology*, 1936, **48(2)**, pp. 297–306.



## - Capítulo 2 -

---

---

### Localización vertical de fuentes sonoras

---

#### § 2.1 MÍNIMO ANGULO AUDIBLE

Además de localizar fuentes sonoras en el plano horizontal (acimut) de manera eficaz, los seres humanos también pueden percibir el ángulo de incidencia del sonido en el plano vertical con cierta precisión, y por supuesto localizar sonidos cuyas posiciones combinan ambos planos.

Se habla de elevación cuando el sonido proviene con un ángulo de incidencia diferente de cero en el eje vertical (Blauert, 1997). Al igual que en acimut, el umbral de discriminación en elevación se mide mediante el Mínimo Angulo Audible (MAA), que indica la mínima diferencia angular perceptible entre dos posiciones de una fuente sonora en el plano vertical. Algunos estudios han mostrado que los umbrales de MAA obtenidos mediante cambios en la elevación de una fuente sonora son generalmente de dos a cuatro veces más grandes que los observados en el plano horizontal (Wettschurek, 1973; Morrongiello y Rocca, 1987). Por ejemplo, Perrot y Saberi (1989) reportaron un MAA en elevación de  $3.6^\circ$ , utilizando ruido blanco como estímulo auditivo. Makous y Middlebrooks (1989) reportaron un error en la percepción de una

fuente sonora de  $3.5^\circ$  cuando la fuente se encontraba al frente y al centro del oyente, el cual aumentaba gradualmente cuando la elevación de la fuente se incrementaba (máximo  $20^\circ$ ). En este estudio también se reportó que cuando la fuente se encontraba en posiciones que combinaban acimut y elevación, los errores podían llegar como máximo a los  $15^\circ$ , mostrando una mejoría en la precisión cuando la información de ambos planos esta disponible. Como vimos, las diferencias interaurales de tiempo e intensidad que brindan las pistas utilizadas para la localización de fuentes sonoras en acimut son puramente binaurales (Blauert, 1997; Sodnik et al, 2004). En cuanto a la elevación, existen muchos estudios que indican que los humanos utilizan pistas monoaurales del dominio de las frecuencias para realizar estas localizaciones (Garas, 2000). Mediante la reflexión y difracción de los sonidos en nuestro cuerpo, cabeza y oído externo, el cual se denomina pinna, se proveen las principales pistas para la localización vertical (Algazi et al, 2001; Bloom, 1977). Estas modificaciones espectrales son dependientes de la dirección con que los sonidos llegan al cuerpo (Hebrank y Wright 1974; Shimoda et al 2007).

Según Makous y Middlebrooks (1989), cuando una fuente sonora se encuentra en una posición que combina posiciones de acimut y elevación el sistema auditivo es capaz de combinar información de diferentes clases de pistas para sintetizar una imagen espacial unitaria.

Gracias a los avances tecnológicos con respecto a los sistemas de captura, fue posible el desarrollo de micrófonos lo suficientemente pequeños como para ser introducidos sin inconvenientes dentro del canal auditivo. De esta manera fue posible obtener la Función de Transferencia Relativa a la Cabeza, mejor conocida como HRTF (del inglés, Head Related Transfer Function), la cual

permite observar las variaciones en el espectro que se producen en los sonidos que chocan contra el torso, cabeza y pinna de un oyente (Rakerd et al, 1999; Searle et al, 1975; Middlebrooks, 1999; Shaw, 1966; Hebrank y Wright, 1974; Wightman y Kistler, 1989).

Estas mediciones tienen la particularidad de capturar las características espectrales afectadas por la dirección con que llegan los sonidos, y por lo tanto poder determinar cuáles son los cambios espectrales que se producen al variar la elevación de una fuente sonora.

Según Middlebrooks (1999) se puede obtener las HRTF de forma efectiva mediante la utilización de sonidos de banda ancha, para tener información en el mayor rango de frecuencias posibles, reproducidos en campo libre (o cámara anecoica), para que el sonido no sea afectado por ningún tipo de reflexiones, y capturado mediante pequeños micrófonos con una respuesta en frecuencia plana, ubicados dentro o en la entrada del canal auditivo. En el Capítulo 5 de la presente tesis (técnicas y sistemas de espacialización), se explica de manera más extensa como se realiza esta grabación y como pueden ser utilizadas estas HRTF para generar un entorno acústico virtual.

Mediante experimentos psicofísicos y grabaciones de HRTF se ha determinado que las señales que son filtradas por el cuerpo, la cabeza y la pinna presentan picos y valles en diferentes lugares del espectro. Estos picos y valles, producidos por la interferencia constructiva y destructiva de nuestra fisonomía, contiene información relacionada con la elevación de una fuente sonora, siendo esta información la que hace posible estimar la elevación de una fuente sonora (Shimoda et al, 2007).

El hecho de que los indicios que utilizamos para percibir el ángulo de inciden-

cia del sonido en el plano vertical son indicios monoaurales fue demostrado en experimentos donde la fuente sonora fue ubicada a diferentes alturas en el centro del plano horizontal (acimut cero), zona conocida como plano medio sagital (PMS). En esta posición particular, las diferencias interaurales de tiempo y amplitud no existen ya que las señales llegan a los dos oídos de idéntica forma. A pesar de no contar con las diferencias interaurales fundamentales en la localización binaural, los oyentes con audición normal pueden localizar de manera efectiva la procedencia angular de sonidos en el plano vertical utilizando solo las pistas espectrales monoaurales (Rakerd et al, 1999). De hecho, la mayoría de los estudios realizados en elevación ubican la fuente sonora en el plano medio sagital, donde no existen diferencias de ITD e ILD, para aislar las pistas monoaurales localización vertical.

En general, se puede decir que la pinna realiza modificaciones espectrales en el sector agudo y sobreagudo del sonido y el torso y la cabeza en el sector agudo y grave del espectro. Debido a esto, para que el ángulo con que incide un sonido pueda ser percibido correctamente el estímulo debe tener un contenido espectral amplio. A continuación repasaremos los estudios que han trabajado sobre estas características de la localización vertical de fuentes sonora.

## § 2.2 INFORMACIÓN PROVISTA POR LA PINNA

La pinna ha sido objeto de estudio desde el siglo XIX, aunque a mediados del siglo XX se han desarrollado los estudios más abarcativos sobre la utilidad de la misma en la localización de fuentes sonoras en el plano medio sagital, específicamente en la localización vertical. Mach (1874), Raleigh (1876)

y Steinhauser (1879) han hecho varias consideraciones sobre la pinna. Mach conjeturó que la forma de la misma podría ser importante en la localización de las fuentes sonoras y fue uno de los primeros en detectar cambios en la calidad del sonido cuando se producían cambios en la orientación de la fuente sonora que se estaba escuchando, cambios que, según Mach, se debían a la transformación que recibían por parte de la pinna. Steinhauser, indicó que la construcción de la pinna del oído humano era similar a un cono que colecta los sonidos de las diferentes posiciones y que a la vez funcionaba como un reflector de determinados sonidos.

Ya en el siglo XX, Batteau (1967) reportó que la pinna podía funcionar como un transformador acústico esencial para la localización de las fuentes sonoras, reforzando esta teoría mediante la producción de un modelo matemático.

Rofler y Butler (1968) han indicado que si bien el oído externo había sido considerado por mucho tiempo un órgano vestigial un trabajo de Jonkees y Groen (1946) reportó que la localización en el plano vertical sería imposible sin éste. Para reforzar esta hipótesis, Fisher y Fredman (1968) y Gardner y Gardner (1973), realizaron experimentos cubriendo las circunvoluciones de la pinna, reportando que la habilidad de los oyentes para localizar sonidos en el plano medio y en elevación se reducía. Diversos estudios han reportado que la pinna influye sobre sonidos que poseen un contenido espectral por encima de los 3-4Khz, debido a que estas frecuencias poseen una longitud de onda semejante a las cavidades que presenta el oído externo humano (Chum y Kim, 2011; Hofman et al, 1998; Bronkhorst, 1995). Algazi et al (2001) reportaron que uno de los efectos más característicos de la pinna sobre el espectro es el que se denomina “muesca de la pinna” (del inglés, “pinna notch”), la cual



aparece dentro de la octava entre 6 y 12Khz. Hoffman et al (1998), se basan en estas características para apoyar la creencia de que para que una fuente sonora pueda ser bien localizada en el eje vertical, debe tener energía en la zona sobreaguda del espectro. Algazi et al (2001) han reforzado esta hipótesis al reportar que la pinna tiene claramente un mayor efecto sobre el espectro por encima de los 3Khz, pero relativamente poco efecto por debajo de esta frecuencia.

Hebrank and Wright (1974), Searle et al, (1975) y Butler y Belendiuk, (1977) han medido la respuesta en frecuencia del oído externo para diferentes ángulos de incidencia en el plano sagital. Estos han mostrado que por encima de los 5Khz aparecen sucesivos picos y valles muy selectivos en la respuesta en frecuencia. También han observado que la posición de estos picos y valles espectrales cambia cuando el ángulo vertical de la fuente sonora se incrementa. Esta revisión de los estudios más importantes sobre el efecto de la pinna en el espectro de una fuente sonora, muestra que esta tiene influencia sobre el espectro de frecuencias por encima de los 3-4Khz, por lo que para realizar una correcta localización de estos sonidos es necesario que los estímulos contengan energía en las frecuencias por encima de estos valores.

### § 2.3 INFORMACIÓN PROVISTA POR EL TORSO

La información que el torso aporta a la percepción de fuentes sonoras en el plano vertical no fue estudiada hasta mediados del siglo XX. El torso y la cabeza perturban el espectro de frecuencias por debajo de los 3Khz (Hanson, 1944). Sin embargo se ha reportado que estos cambios tienen un efecto leve

sobre la localización angular en el plano vertical y los resultados obtenidos generan información confusa. El efecto del torso en la localización vertical ha sido estudiado sistemáticamente por Gardner y Gardner en 1973, quien midió las HRTF de un maniquí con y sin pinna, concluyendo que esta no tenía influencia por debajo de los 3.5 KHz, pero que el torso introducía importantes modificaciones en el espectro de frecuencias entre los 700Hz y los 3.5KHz, las cuales considero “pistas secundarias”. Gardner advirtió que la sola presencia de estas modificaciones no asegura que el sistema auditivo las utilice en la localización de sonidos en el plano medio sagital. También observó que algunos sujetos podían localizar fuentes sonoras ubicadas en el plano medio sagital, a pesar de que los estímulos no poseían información por encima de los 4KHz (la zona de acción de la pinna) aunque la sensación de localización de los estímulos disminuyó drásticamente en comparación con estímulos que contenían información espectral en la zona aguda del espectro.

Kuhn (1987) realizó grabaciones binaurales con un maniquí, combinando condiciones con torso, sin torso, con pinna y sin pinna, reportando que observó variaciones por debajo de los 2KHz en las condiciones sin pinna. Kuhn también observó que el filtrado debido al torso presento forma de filtro peine (del inglés, “comb filter”), introduciendo muescas periódicas en frecuencias por debajo de los 3KHz. La ubicación de estas muescas produce un patrón que varía directamente con la elevación de la fuente.

Por otro lado, Searle et al (1975) identifico seis pistas de localización en su modelo estadístico de la localización auditiva humana, y utilizó los datos de Gardner y Gardner (1973) para estimar la varianza proveniente de las reflexiones del torso o “el rebote en el hombro” (del inglés, “shoulder bounce”),

concluyendo que el rebote en los hombros es la pista más débil en la percepción de la elevación.

## REFERENCIAS BIBLIOGRÁFICAS

- ALGAZI V.R., AVENDANO C. y R.O., DUDA: «Elevation localization and head-related transfer function analysis at low frequencies». *J. Acoust. Soc. Am.*, 2001, **109(3)**, pp. 1110–1122.
- BATTEAU, D.W.: «The Role of the Pinna in Human Localization». *Proc. R. Soc. Lond.*, 1967, **168**, pp. 158–180.
- BLAUERT, J.: *Spatial hearing. The psychophysics of human sound localization*. Massachusetts, USA, 1997.
- BLOOM, P.J.: «Determination of monaural sensitivity changes due to the pinna by use of minimum-audible-field measurements in the lateral vertical plane». *J. Acoust. Soc. Am.*, 1997, **61(3)**, pp. 820–828.
- BRONKHORST, A.W.: «Localization of real and virtual sound sources». *J. Acoust. Soc. Am.*, 1995, **98(5)**, pp. 2542–2553.
- BUTLER, R.A. y BELENDIUK, K.: «Spectral cues utilized in the localization of sound in the median sagittal plane». *J. Acoust. Soc. Am.*, 1977, **61(5)**, pp. 1264–1269.
- CHUM, C.J. y KIM, H.K.: «Sound Source Elevation Using Spectral Notch Filtering and Directional Band Boosting in Stereo Loudspeaker Reproduction». *Consumer Electronics, IEEE Transactions on*, 2011, **57(4)**, pp. 1915–1920.
- FISHER, H.G. y FREEDMAN, S.J.: «The role of the pinna in auditory localization». *J. Aud. Res.*, 1968, **8**, pp. 15–26.
- GARAS, J.: *Adaptive 3D sound systems*. Springer, 2000.
- GARDNER, M.B. y GARDNER, R.S.: «Problem of localization in the median plane: effects of pinnae cavity occlusion». *J. Acoust. Soc. Am.*, 1973, **53**, pp. 400–408.
- HANSON, W.W.: «The baffle effect of the human body on the response of a hearing aid». *J. Acoust. Soc. Am.*, 1944, **16**, pp. 60–62.
- HEBRANK, J. y WRIGHT, D.: «Spectral cues used in the localization of sound sources on the median plane». *J. Acoust. Soc. Am.*, 1974, **56(6)**, pp. 1829–1834.
- HOFMAN P.M., VANRISWICK J.G.A y A.J., VANOPSTAL: «Relearning sound localization with new ears». *Nature Neuroscience*, 1999, **1(5)**, pp. 417–421.
- JONGKEES, L.W.B. y GROEN, J.J.: «On Directional hearing». *J. Laryngol. Otol.*, 1946, **61**, pp. 494–504.
- KUHN, G.F.: «Physical acoustics and measurements pertaining to directional hearing». *En Directional Hearing, editado por W. A. Yost y G. Gourevitch (Springer-Verlag, New York)*, 1987, pp. 3–21.
- MACH, E.: «Archiv fur Ohrenheilkunde», 1874, **7(2)**, pp. 14–26.
- MAKOUS, J.C. y MIDDLEBROOKS, J.C.: «Two-dimensional sound localization by human listeners». *J. Acoust. Soc. Am.*, 1989, **87(5)**, pp. 2188–2200.
- MIDDLEBROOKS, J.C.: «Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency». *J. Acoust. Soc. Am.*, 1999, **106(3)**, pp. 1493–1510.
- MORRONGIELLO, B.A. y ROCCA, P.T.: «Infants' localization of sounds in the horizontal plane: effects of auditory and visual cues». *Child Development*, 1987, **58(4)**, pp. 918–927.

- PERROT, D.R. y SABERI, K.: «Minimum audible angle thresholds for sources varying in both elevation and azimuth». *J. Acoust. Soc. Am.*, 1989, **87(4)**, pp. 1728–1731.
- RAKERD B., HARTMANN A.M. y MCCASKEY, T.L.: «Identification and localization of sound sources in the median sagittal plane». *J. Acoust. Soc. Am.*, 1999, **106(5)**, pp. 2812–2820.
- RALEIGH, L.: «On Waves». *Phil. Mag.*, 1876, **1**, pp. 257–279.
- ROFLER, S.K. y BUTLER, R.A.: «Factors That Influence the Localization of Sound in the Vertical Plane». *J. Acoust. Soc. Am.*, 1968, **43(6)**, pp. 1255–1259.
- SEARLE C.L., CUDDY C.R., BRAIDA L.D. y M.F., DAVIS: «Binaural pinna disparity: another auditory localization cue». *J. Acoust. Soc. Am.*, 1975, **57(2)**, pp. 448–455.
- SHAW, E.A.G.: «Ear canal pressure generated by a free sound field». *J. Acoust. Soc. Am.*, 1966, **39**, pp. 465–470.
- SHIMODA T., KUMON M. KOHZAWA R. MIZUMOTO I., NAKASHIMA T. y Z., IWAI: «Sound Localization of Elevation using Pinnae for Auditory Robots». *En Robust Speech Recognition and Understanding, I-Tech Education and Publishing*, 2007, pp. 421–438.
- SODNIK J., BOBOJEVIC G., SUSNIK R. y S., TOMAZIC: «The directional resolution of virtual sound sources in humans». *Electrotechnical Review*, 2004, **71(3)**, pp. 121–127.
- STEINHAUSER, A.: «Phenomeno of binaural audition». *Phil. Mag.*, 1879, **7(5)**, pp. 181–261.
- WETTSCHUREK, R.G.: «The absolute difference limen of directional perception in the median plane under conditions of both, natural hearing and hearing with artificial-head-system». *Acustica*, 1973, **28(4)**, pp. 197–208.
- WIGHTMAN, F.L. y KISTLER, D.J.: «Headphone simulation of free-field listening. I: Stimulus synthesis». *J. Acoust. Soc. Am.*, 1984, **85(2)**, pp. 1829–1834.



## - Capítulo 3 -

---

---

### Localización en distancia de fuentes sonoras

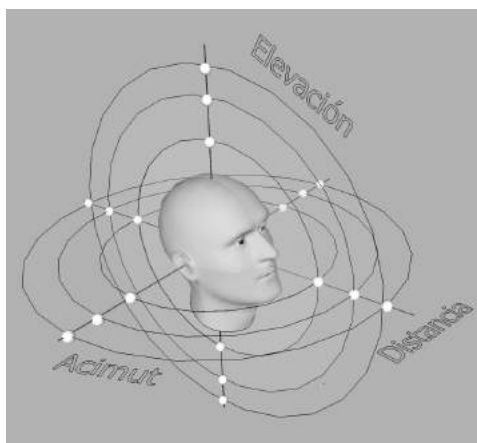
---

#### § 3.1 INTRODUCCIÓN

La evolución de los sofisticados mecanismos que conforman el sistema auditivo de los vertebrados superiores, capaz de detectar, reconocer y ubicar en el espacio diversas fuentes sonoras, representó una importante ventaja para estos en la lucha por la supervivencia. La capacidad de detectar la ubicación de fuentes sonoras de forma rápida y precisa es un atributo muy importante, principalmente en situaciones donde el sentido de la vista no puede intervenir como, por ejemplo, en situaciones de plena oscuridad o cuando los eventos sonoros ocurren fuera del campo visual.

Como vimos en los capítulos anteriores, para percibir la ubicación de una fuente sonora el cerebro debe reconocer tanto la dirección (en acimut y elevación) de la que proviene el sonido, como también la distancia a la que se encuentra la fuente que lo emite (Figura 3.1).

A diferencia de la percepción espacial en acimut y elevación, la Percepción Auditiva de Distancia (PAD) ha recibido menos atención por parte de los investigadores. Para percibir la distancia a una fuente sonora, el cerebro debe



**Figura 3.1:** Representación esquemática del acimut, elevación y distancia.

reconocer y analizar una gran cantidad de indicios, los cuales tienen diferente grado de importancia dependiendo de las características del estímulo, el medio ambiente acústico, la experiencia del oyente, etc. Por estas razones, diseñar experimentos en donde estas pistas puedan ser aisladas ha sido uno de los principales obstáculos a sortear para abordar el estudio de la PAD.

El estudio de la PAD se ha desarrollado mediante diferentes metodologías experimentales, en diferentes ambientes acústicos (tanto reales como virtuales) y utilizando estímulos de diversas características. Los experimentos realizados pueden dividirse, a grandes rasgos en dos grandes ramas: los diseñados para medir umbrales mínimos y los diseñados para medir la precisión de un oyente al estimar la distancia a una fuente. En general, para medir umbrales de PAD se ha utilizado un método de elección forzada de dos alternativas. En este tipo de experimentos se le presenta al oyente un sonido a una distancia fija (estímulo de referencia o estándar) y luego se presenta un estímulo de comparación ubicado a una distancia diferente. La tarea del sujeto es informar si el estímulo de comparación se encuentra a una distancia mayor o menor que el estímulo estándar. Este umbral se revisó en algunos casos mediante la comparación entre un parlante móvil y un parlante fijo (Strybel y Perrot,

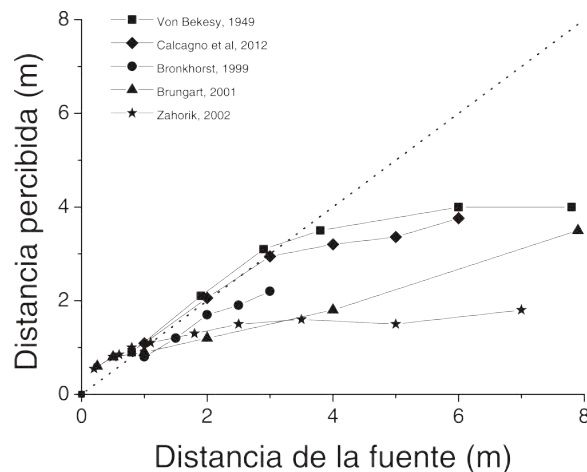


1984) y en otros casos comparando las diferencias producidas por una única fuente sonora ubicada a diferentes distancias (Simpson y Stanton, 1973; Edwards, 1955). También suelen medirse, utilizando un parlante fijo, el mínimo cambio de distancia perceptible en respuesta a variaciones de diferentes indicios de distancia como por ejemplo la intensidad del estímulo (Gamble, 1909; Edwards, 1955).

Para medir la precisión con que un oyente percibe la distancia física real (distancia absoluta) de una fuente sonora, es necesario obtener curvas psicofísicas de PAD. Sin dudas la metodología más utilizada en estos casos es la estimación mediante informes verbales. En este tipo de experimentos el oyente debe indicar, mediante una sentencia verbal, la distancia aparente de la fuente utilizando escalas de distancia explícitas (ej, metros, pies) o mediante estimaciones de magnitud (ej, la mitad de, el doble de). Otra metodología utiliza acciones motoras como método de respuesta (perceptually directed action). El ejemplo más utilizado es conocido como “walking directed action” donde el oyente debe caminar (con los ojos tapados) hasta la ubicación aparente de la fuente sonora (Loomis et al, 1998).

En general, los experimentos para medir la precisión con que un oyente percibe la distancia a una fuente se realizan reproduciendo el estímulo desde diferentes distancias, para luego obtener un promedio de las respuestas y el error con respecto a la distancia física real. Sin embargo, también existen experimentos donde solo se realiza una única presentación del estímulo a una distancia determinada. Como veremos más adelante, este método, conocido como método de presentación única, se utiliza para diferenciar entre pistas absolutas y relativas de PAD (Mershon y Bowers, 1989).

En la mayoría de los estudios, la respuesta obtenida con cualquiera de estos métodos muestra que la distancia a una fuente sonora es percibida con menor precisión y mayor variabilidad que el ángulo de incidencia del sonido (Zahorik et al, 2005). De hecho, varios estudios han reportado que, incluso en presencia de múltiples indicios, la PAD suele ser pobre (Coleman, 1962; Middlebrooks y Green, 1991; Wensel et al, 1993; Zahorik, 2002a). Estos trabajos muestran que los humanos tienden a sobrestimar la distancia a la fuente cuando ésta se encuentra a menos de 2 metros del oyente, mientras que su distancia es substancial y progresivamente subestimada cuando se encuentra a distancias mayores (Figura 3.2).



**Figura 3.2:** *Diferentes curvas de PAD obtenidas de artículos que utilizan varios sistemas de medición. Es interesante observar como los trabajos realizados en entornos acústicos virtuales presentan un nivel de subestimación más marcado.*

En la mayoría de los casos, la relación entre la distancia real a la fuente y la distancia percibida puede ajustarse mediante una función de potencia ( $Y = a \cdot x^b$ ) con un exponente considerablemente menor que uno, lo que corresponde a una no linealidad compresiva entre la distancia física y la percibida (Zahorik et al, 2005). Sin embargo, se han encontrado algunas excepciones, donde dife-

rentes tipos de información adicional, como por ejemplo la familiaridad con el estímulo auditivo o la presencia de pistas provistas por la modalidad visual, mejoran notablemente la efectividad en los juicios de distancia (Brungart y Scott, 2001; Mershon y Philbeck, 1991; Calcagno et al, 2012).

Como hemos observado, la PAD es un fenómeno complejo que involucra una gran cantidad de indicios diferentes (tanto acústicos como no acústicos). Vimos también que tanto el rol como la importancia de cada una de estas pistas pueden ser estudiadas mediante diferentes metodologías experimentales. Una exhaustiva revisión de todos estos aspectos es fundamental para un entendimiento profundo de la PAD.

### § 3.2 INDICIOS DE PERCEPCIÓN AUDITIVA DE DISTANCIA

Cuando un sonido atraviesa un espacio sufre múltiples cambios en sus propiedades físicas antes de llegar a los oídos del oyente. Muchas de estas variaciones están relacionadas con la distancia entre el oyente y la fuente sonora. El cerebro es capaz de reconocer estos cambios y utiliza esta información para generar una representación mental de la distancia a la fuente. Sin embargo, como veremos más adelante, existen pistas que no tienen relación directa con los cambios que sufre la señal en su interacción con el medio y afectan de manera importante a la PAD.

Entre el amplio espectro factores que influyen sobre la PAD podemos mencionar: la intensidad del sonido, los cambios en el espectro de frecuencias del estímulo, la relación de intensidades entre el sonido directo y el reverberante, los indicios binaurales, la influencia de pistas visuales, el conocimiento previo

del estímulo, etc.

Los indicios de PAD pueden dividirse en relativos y absolutos. Los indicios relativos (la intensidad por ejemplo) proporcionan información sobre la distancia a una fuente sonora mediante la comparación entre las distancias de dos o más fuentes. Estos indicios se estudian mediante procedimientos experimentales donde los estímulos son presentados de manera repetida. Por ejemplo, una forma de hacerlo sería repitiendo el mismo estímulo en diferentes distancias a testear. Este tipo de procedimiento experimental es conocido como método de presentaciones repetidas (o repeated-presentation method). Bajo tales condiciones, el oyente puede percibir los cambios relativos en el estímulo, e inconscientemente hacer juicios relativos de la distancia a la fuente en comparación con un estímulo presentado previamente.

Por el contrario, los indicios absolutos de distancia (como la reverberación) le brindan al oyente la información necesaria para percibir la distancia a la fuente sin necesidad de un estímulo de referencia. Los indicios absolutos de PAD se estudian mediante un método conocido como método de presentación única (Mershon y Bowers, 1989). La técnica consiste en obtener la distancia aparente de la fuente en respuesta a un único estímulo auditivo proveniente de una distancia determinada. De este modo, el oyente debe indicar la distancia a la fuente valiéndose solamente de los indicios absolutos de PAD producidos en esa única presentación, y no puede, como en el procedimiento de presentaciones repetidas, relacionar esta información con la de otros estímulos presentados previamente. Esta clasificación es importante para comprender el tipo de información que contiene cada pista para así determinar la contribución de cada una de ellas a la PAD.

A continuación se explicaran los indicios más estudiados en PAD y revisaremos los resultados más relevantes publicados hasta la fecha en la literatura.

### § 3.3 INTENSIDAD

La pista más evidente en la PAD es la relación entre la intensidad del sonido que llega al oyente y la distancia a la fuente sonora. Cualquier persona con audición normal puede notar que la intensidad de un sonido proveniente de una fuente fija disminuye a medida que nos alejamos de ella y viceversa. Lo mismo sucede cuando una fuente móvil se acerca o se aleja de nosotros. En ambientes anecoicos (sin reflexiones) la relación entre la distancia a la fuente y la intensidad con que el estímulo llega al oyente obedece a la ley del inverso del cuadrado de la distancia. Es decir, la intensidad de la señal disminuye 6dB cada vez que la distancia a la fuente se duplica y aumenta 6dB cuando la distancia a la fuente disminuye a la mitad (Gamble, 1909; Coleman, 1963). Sin embargo, esta ley no puede aplicarse en ambientes reverberantes, ya que las reflexiones del sonido en las diferentes superficies de un recinto hacen que la pérdida de intensidad general del estímulo en función de la distancia sea menor (Griffin, 1958; Lochner y de V. Keet, 1960; Begault, 1991). Por ejemplo, se ha reportado para una sala reverberante de 3.5 x 3.8 m una caída en la intensidad de 4dB cuando la distancia a la fuente aumentó al doble de su valor original (Simpson y Stanton, 1973). La ley del inverso del cuadrado de la distancia tampoco puede aplicarse en el campo cercano (distancia menores a 1m) ya que en este rango el frente de onda es aun esférico y por ende los cambios en la intensidad de la señal producidos por cambios en la distancia

dependen, también, del ángulo de incidencia con el que el sonido interactúa con la cabeza del oyente (efecto de paralaje acústico, Ashmead et al, 1990; Blauert, 1997; Brungart, 1999).

La simpleza de la relación entre la distancia a la fuente y la intensidad hacen de éste un indicio muy influyente en PAD. Sin embargo, es importante aclarar que, desde la perspectiva del oyente, esta pista puede ser ambigua, ya que variaciones en la potencia acústica de la fuente pueden ser confundidos con cambios en la distancia de la misma (Zahorik, 2002b). De hecho, experimentos psicofísicos han demostrado que, cuando se manipula la intensidad de un estímulo auditivo reproducido desde una fuente fija, los oyentes tienen la sensación de que la distancia a la fuente cambia (Békésy, 1949; Gardner, 1969).

Como mencionamos anteriormente, la intensidad es un indicio que brinda información relativa sobre la distancia a una fuente sonora. Esta pista contiene información que nos ayuda a percibir si una fuente se encuentra más cerca o más lejos que otra o, en caso de fuentes móviles, si se acercan o se alejan. Sin embargo, cuando este indicio opera aisladamente no brinda información sobre la distancia física real de la fuente. Esto fue demostrado en un artículo publicado por Mershon y King en 1975. Mershon y King utilizaron un método que consiste en obtener juicios de distancia en respuesta a un único estímulo auditivo y por ende sin ninguna referencia previa. Esta metodología les permite a los investigadores asegurar que la respuesta de los oyentes no está influenciada por indicios relativos de distancia. Mershon y King utilizaron este método en una cámara anecoica y no encontraron diferencias significativas en la distancia percibida por dos grupos de oyentes en respuesta a un mismo estímulo pero con

intensidades diferentes (20 dB de diferencia). Sin embargo, la distancia aparente de la fuente entre ambos grupos fue significativamente diferente cuando la intensidad del estímulo fue variada en una segunda y tercera presentación. En un segundo experimento, los autores presentaron a cada grupo un estímulo proveniente de dos distancias diferentes (2,74 y 5,49 m). De manera similar a lo obtenido en el primer experimento, la respuesta de ambos grupos tampoco fue significativamente diferente. Sin embargo, los participantes fueron capaces de percibir diferencias entre ambas distancias cuando se le permitió comparar el estímulo proveniente de ambas en una segunda y tercera presentación. Por último, Mershon y King encontraron que, si el mismo experimento era repetido en una sala reverberante, bastaba una sola presentación del estímulo para que la respuesta entre ambos grupos fuera significativamente diferente. Los autores concluyeron entonces que la intensidad es un indicio relativo de distancia mientras que factores relacionados con la reverberación pueden servir como indicios absolutos.

Debido al carácter relativo de la intensidad, los estudios interesados en estudiar el rol de este indicio en la PAD deben utilizar como mínimo dos presentaciones del estímulo para que el oyente pueda comparar entre ambos. Estudios pioneros en este tema han examinado la capacidad de los oyentes para discriminar cambios en la distancia aparente de la fuente en condiciones anecoicas donde la intensidad es el principal indicio disponible. Gamble (1909) y Edwards (1955) reportaron que se necesita un cambio del 20 % en la distancia real de la fuente para que un oyente sea capaz de percibirlo. Sin embargo, en trabajos posteriores se obtuvieron umbrales menores. Experimentos realizados por Simpson y Stanton (1973) y Strybel y Perrott (1984) reportaron

umbrales de un 13 y 3 % respectivamente. Ambos trabajos muestran además que los umbrales de discriminación fueron mucho mayores cuando la fuente se encontraba a menos de 3 metros del oyente (48 % y 20 % respectivamente). Sin embargo, en un trabajo reciente, Ashmead et al (1990) midieron el umbral de PAD para distancias pequeñas (entre 1 y 2 metros) y obtuvieron valores por debajo del 6 %.

Además de estudiar la relación entre la intensidad del sonido y la distancia a la fuente se han realizado estudios específicos sobre la relación entre la sonoridad y la PAD. La sonoridad (en inglés "loudness") es una medida subjetiva de la intensidad con la que un sonido es percibido por un oyente. No obstante, la sonoridad no depende sólo de la intensidad del sonido, sino también de otros factores como el ancho de banda, el contenido espectral y la duración del estímulo auditivo. Stevens y Guirao (1962) demostraron que la sonoridad y la distancia aparente de la fuente están inversamente relacionadas en condiciones donde la intensidad es el principal indicio presente. Sus datos muestran que se requiere variar la intensidad de la señal en aproximadamente 10 dB para que el oyente reporte una disminución de la sonoridad a la mitad de su valor inicial o un aumento de la distancia del doble de su valor original. Resultados similares fueron obtenidos por Warren (1958) y Begault (1991). Warren reportó que fueron necesarios cambios de  $\pm 6$ dB para que los oyentes perciban variaciones en la distancia a la fuente del doble o la mitad (consistente con la ley de la inversa del cuadrado de la distancia) mientras que Begault reportó que, en un ambiente acústico virtual, necesitó disminuir la intensidad de la señal entre 8 y 10dB para simular un aumento en la distancia de la fuente al doble de su valor original.



Una forma más directa de estudiar el rol de la intensidad en la PAD es mediante la obtención de curvas psicofísicas. Como mencionamos antes, en este tipo de experimentos el oyente debe indicar a qué distancia percibe la fuente sonora. Casi sin excepción, los estudios que miden la eficacia de los oyentes para percibir la distancia a la fuente en ambientes reverberantes reportan una importante subestimación de la distancia a la fuente cuando esta se encuentra a más de 2m (Simpson y Stanton, 1973). De este modo, para mantener una relación lineal entre la distancia real de la fuente y la distancia percibida se necesitan cambios de intensidad mayores a los que predice la ley del inverso del cuadrado (Mershon y King, 1975; Begault, 1991). Sin embargo, como veremos luego, esta subestimación puede estar relacionada, entre otros factores, con el tipo de estímulo que se utiliza. Por ejemplo, Mershon y King (1975) y Coleman (1962) indican que esta fuerte subestimación puede ser revertida significativamente si los oyentes están familiarizados con el sonido que proviene de la fuente.

### § 3.4 ESPECTRO

Las pistas espectrales que intervienen en la PAD serán revisadas en el Capítulo 7 de la presente tesis (El rol del contenido espectral en la percepción auditiva de distancia) donde se presentan además resultados experimentales obtenidos en el marco de esta tesis.

### § 3.5 FAMILIARIDAD CON EL ESTÍMULO AUDITIVO

La precisión para estimar la distancia a una fuente sonora puede verse afectada notablemente si el oyente está familiarizado con el sonido que ésta emite. Uno de los sonidos más familiares para el ser humano es sin duda la voz hablada. Es por esto que la voz humana ha sido utilizada como modelo para estudiar el efecto de la familiaridad del estímulo auditivo en la PAD. Sin embargo, los estudios en este tema son escasos y, como veremos a continuación, arrojan resultados contradictorios. En general, existe consenso en que la precisión de la PAD aumenta cuanto más familiar es el estímulo auditivo para el oyente. Como vimos en el apartado de intensidad, en ambientes anecoicos los oyentes tienden a subestimar la distancia a la fuente cuando esta se encuentra a distancias mayores a 2 metros. Sin embargo, estudios realizados en condiciones anecoicas pero utilizando como estímulo una voz hablada muestran que esta subestimación es revertida significativamente (Békésy, 1949; Gardner, 1969; Cochran et al, 1968). Es importante destacar que este aumento en la precisión no depende de la persona que habla, ya que la precisión es similar cuando se compara la respuesta a voces de personas conocidas por los oyentes con la respuesta a voces de personas desconocidas (Brungart y Scott, 2001). Sin embargo, existen reportes que contradicen lo antedicho. Por ejemplo, en experimentos realizados en una sala reverberante por Nielsen (1991), no se encontraron diferencias significativas cuando se comparó la distancia aparente de la fuente en respuesta a voz humana grabada con la distancia aparente en respuesta a otros tipos de estímulos (ruido blanco, música grabada y grabaciones de guitarra). En el mismo orden de cosas, en un experimento realizado

por Zahorik (1998) en una sala reverberante virtual, la distancia aparente de la fuente ante estímulos de voz hablada y ruido blanco no presentó diferencias significativas.

Sin embargo, Brungart y Scott (2001) sostienen que el hecho de que esos experimentos fueran realizados en ambientes reverberantes y que hayan utilizado un método de presentaciones repetidas, puede haber aumentado la precisión de los oyentes para percibir la distancia de los sonidos desconocidos disminuyendo así el efecto del conocimiento previo de la voz humana. Además, estos autores sostienen que en los experimentos de Nielsen y Zahorik no se tuvo en cuenta un importante atributo de la voz humana: el nivel de producción del discurso (la intensidad con que la persona habla). Las propiedades acústicas de la palabra cambian de forma sistemática cuando varía el nivel de producción del discurso. Cuando una persona eleva la voz en una conversación, el incremento de la presión vocal hace que aumente la frecuencia fundamental, la energía de la primera formante y la relación de energías entre las frecuencias agudas y graves del espectro vocal (Lienard y Benedetto, 1999). Varios autores han sugerido que la relación entre el nivel de presentación del estímulo con el nivel de producción puede funcionar como un indicio de PAD. Los sonidos del habla con niveles altos de presentación pero bajos niveles de producción (ej, susurros que llegan con alta intensidad) indicarían transmisores cercanos, mientras que los sonidos del habla con niveles de presentación bajos, pero con altos niveles de producción (ej, gritos que llegan con baja intensidad) indicarían que la persona se encuentra a distancias lejanas.

En 1991 Mershon y Philbeck pusieron a prueba esta hipótesis y obtuvieron resultados contundentes. Los autores midieron la distancia aparente a una

fuelle fija utilizando voces pregrabadas con el mismo nivel de presentación pero diferentes niveles de producción (susurro, voz normal y grito). Sus resultados muestran que un susurro fue percibido a menos de la mitad de distancia que una voz normal mientras que un grito se percibió a una distancia tres veces mayor. Resultados similares fueron obtenidos en otros estudios tanto en ambientes reales (Traunmüller y Eriksson, 2000) como en entornos virtuales (Brungart y Scott, 2001).

Como hemos indicado, algunos investigadores han propuesto que este efecto se debe a los cambios espectrales que se producen en respuesta a los distintos niveles de producción. Sin embargo, existe evidencia que indica que el contenido espectral de la voz hablada no puede explicar por completo este fenómeno. Esto ha sido demostrado en un estudio donde los oyentes fueron menos sensibles a los cambios en el nivel de producción del habla cuando el estímulo fue reproducido al revés a pesar de que en ambos casos el contenido espectral de cada nivel de producción era idéntico. Este resultado sugiere que, además del contenido espectral inherente a cada nivel de producción, existen factores cognitivos que afectan la manera en que los oyentes interpretan los indicios acústicos presentes en las señales de voz.

### § 3.6 INDICIOS ACÚSTICOS BINAURALES

Además de ser crucial para la percepción del ángulo de incidencia del sonido en acimut, la información que el sistema auditivo obtiene de las diferencias interaurales de tiempo (ITD) e intensidad (ILD), puede servir como indicio de PAD cuando la fuente sonora se encuentra a distancias menores a un metro

(Morimoto et al, 1975; Simpson y Stanton, 1973). En este rango, los cambios de distancia a la fuente producen cambios de ILD mayores que de ITD (Shinn-Cunningham, 2000; Brungart y Scott, 2001). En Zahorik et al (2005) se sugiere que los oyentes pueden usar esta combinación de ILD variable e ITD estático como un indicio de PAD. Esta hipótesis surge en parte en relación a resultados obtenidos por Brungart y Rabinowitz en 1999. Para aislar las pistas binaurales y observar como estas afectan a la PAD, Brungart y Rabinowitz realizaron sus experimentos en un ambiente anecoico (evitando la pista de reverberación) y controlando la intensidad de los estímulos de tal manera que desde cualquiera de las distancias utilizadas la intensidad sonora en la posición del oyente fuese la misma. Los resultados de los experimentos realizados por Brungart y Rabinowitz indicaron que, en el campo cercano, los participantes son capaces de realizar estimaciones precisas de distancia utilizando solamente pistas binaurales cuando la fuente se presenta con un ángulo de incidencia de  $90^\circ$ .

Otro efecto que ha sido propuesto como una posible pista binaural es el paralaje acústico. Este efecto se produce cuando una fuente sonora se encuentra lo suficientemente cerca de la cabeza de un oyente como para producir diferencias substanciales entre los ángulos con que el sonido incide en cada oído. Esta forma de paralaje modifica el filtrado dependiente de la direccionalidad resultante de la difracción producida por los oídos, el torso y la cabeza (Ashmead et al, 1990; Blauert, 1997). Estas variaciones espectrales pueden observarse mediante la medición de funciones de transferencia relativas a la cabeza (HRTF del inglés Head Related Transfer Function), las cuales pueden brindar la información exacta de como varía el espectro de un estímulo al

ser filtrado por el cuerpo de una persona. Un estudio realizado por Otani y Hirahara (2009) estudió la relación entre el espectro obtenido de las HRTF, calculadas mediante una simulación numérica, y la distancia a la fuente sonora para distancias de la fuente menores a 3 metros. El análisis reportó que los espectros de las HRTF simuladas variaban marcadamente en respuesta a cambios en la distancia de la fuente sonora. Para HRTF ipsilaterales, observaron un aumento en la amplitud de las frecuencias ubicadas en la zona baja del espectro a medida que la distancia a la fuente disminuía, mientras que en las HRTF contralaterales, observaron que la amplitud en todo el rango de frecuencias disminuía, cuando la distancia disminuía.

Además de los trabajos que estudian el efecto de las pistas binaurales a cortas distancias, existen algunos estudios que han reportado el efecto de estas pistas para distancias de la fuente mayores a un metro. Lokki et al (2011) sugieren que las reflexiones laterales afectan la PAD. Este estudio reportó que las reflexiones laterales fueron beneficiosas para incrementar el efecto de sonoridad binaural, en donde, debido a la forma de la cabeza humana, las reflexiones laterales son más amplificadas que las reflexiones provenientes del plano medio. A su vez Lokki y Pätynen (2011) reportaron que este efecto varía con la distancia de la fuente sonora. En este estudio se realizaron una serie de experimentos utilizando grabaciones de obras sinfónicas reproducidas en salas en donde se contaba con reflexiones tempranas laterales o medias. Cuando los oyentes disponían de reflexiones laterales la fuente sonora fue percibida a menor distancia que cuando solo disponían de reflexiones provenientes del plano medio. Resumiendo, los resultados de Lokki y Pätynen demuestran que las reflexiones laterales mejoran la calidad de la escucha en profundidad (distancia)

de una fuente sonora determinada (en este caso música) dentro de diferentes salas de concierto y además que estas reflexiones afectan la distancia aparente de la fuente sonora. Sin embargo, no reportan si esta pista mejora o empeora la efectividad de los oyentes para determinar la distancia a una fuente sonora.

### § 3.7 REVERBERACIÓN

Las ondas sonoras pueden propagarse en ambientes abiertos o cerrados. Dentro de los ambientes cerrados podemos pensar en dos grandes divisiones: los ambientes anecoicos y los reverberantes. Los primeros se asemejan mucho a los ambientes abiertos debido a que el tratamiento acústico que presentan no permite que el sonido se refleje en ninguna superficie, permitiendo que solo se escuche el sonido directo emitido por la fuente. En los ambientes reverberantes, existen múltiples superficies u objetos que hacen que la señal se refleje, produciendo copias modificadas del sonido producido por la fuente. En los ambientes con reflexiones el sonido llega al oyente a través de dos vías: la vía directa (o sonido directo) y la reverberante (o sonido reverberante). El sonido directo es el que llega al oyente desde la fuente sin reflejarse en ninguna superficie. Por el contrario, el sonido reverberante llega al oyente después de reflejarse en, al menos, una superficie del recinto (piso, techo, paredes, muebles, etc). Al hacer contacto con las superficies de un recinto las características originales del sonido (frecuencias, fase e intensidad) sufren cambios importantes. Estas variaciones dependen de las características acústicas del medio (tamaño de la sala, materiales de las superficies reflectantes, etc) por lo que el sonido reflejado contiene información que cuantifica notablemente las

cualidades acústicas de un recinto. Además de proveer información auditiva relacionada con las características acústicas del ambiente, la reverberación es un importante indicio de PAD. Como veremos a continuación, existe abundante evidencia que indica que la reverberación mejora notablemente la precisión de un oyente para percibir la distancia a una fuente sonora.

Como vimos en el apartado de intensidad, en ambientes sin reflexiones, la PAD suele ser ineficaz (la distancia a la fuente es fuertemente subestimada; Zahorik et al, 2005). Sin embargo, se ha demostrado que en ambientes reverberantes, esta subestimación disminuye significativamente (Mershon et al 1989). Numerosos trabajos han reportado que la distancia aparente a una fuente ubicada a una distancia determinada es mayor en ambientes reverberantes que la obtenida bajo las mismas condiciones en salas anecoicas (Nielsen, 1993; Békésy, 1938; Butler et al, 1980; Begault, 1992). En el mismo sentido, Bronkhorst y Houtgast (1999) demostraron, mediante experimentos psicofísicos en un entorno acústico virtual, que el número de reflexiones simulado correlaciona positivamente con la distancia percibida, es decir, un incremento en el número de las reflexiones resulta en un aumento de la distancia aparente a la fuente.

Varios autores han sugerido que el aumento en la precisión de los juicios de distancia observado en ambientes reverberantes se debe al aporte de un importante indicio: la relación de energías que existe entre el sonido directo y el reverberante (relación D/R). En ambientes reverberantes, la intensidad del sonido directo disminuye 6dB cada vez que la distancia a la fuente se duplica, mientras que la energía del sonido reverberante permanece relativamente constante. La relación D/R es un indicio interesante ya que no depende de



la intensidad de la señal y puede ser eficaz en cualquier entorno donde se produzcan reflexiones (incluso en campo abierto hay al menos una reflexión proveniente del suelo).

Múltiples trabajos indican que la relación D/R, supeditada al cambio de la distancia de la fuente sonora, conforma una pista absoluta de PAD. Mershon y Bowers (1979) han reportado que, a diferencia de lo que ocurre en ambientes anecoicos, en ambientes reverberantes basta una sola presentación del estímulo para que los oyentes reporten la distancia a la fuente con bastante precisión. Sin embargo, los juicios de distancia mejoran aun más si se realizan múltiples presentaciones del estímulo (Mershon y King, 1975). La importancia de la relación D/R fue estudiada por primera en un estudio realizado por Von Békésy (1938). Békésy demostró que un decrecimiento en la relación D/R resulta en un incremento en la distancia aparente de un fuente sonora. Mershon y King (1979) demostraron que el efecto de la relación D/R se da en un amplio rango de distancias entre la fuente sonora y los oyentes. La efectividad de este indicio fue probada también en ambientes acústicos virtuales (Butler et al, 1980; Lounsbury y Butler, 1979; Begault, 1992) y mediante la utilización de técnicas de reverberación digital (Schroeder y Logan, 1961). Desde hace años, la relación D/R ha sido considerada uno de los indicios más influyentes en PAD. No obstante, esta señal auditiva tiene varias limitaciones. Zahorik (2002a) reportó que el umbral de sensibilidad a la relación D/R resulta en valores de alrededor de 5-6 dB. Este umbral es mucho mayor que la mínima diferencia perceptible para la intensidad ( $\approx 1$ dB) (Jesteadt et al, 1977) y, siguiendo la ley del inverso del cuadrado, representaría un cambio en la distancia de más del 100%. Zahorik concluye entonces que la relación D/R no proporciona la

información suficiente para percibir cambios finos en la distancia de la fuente sonora. Sin embargo, en un estudio reciente realizado por Larsen et al (2008), similar al publicado por Zahorik, se reportaron umbrales mucho menores ( $\approx$ ) para la relación D/R. Otra limitación de la relación D/R es que su valor depende fuertemente de las características acústicas del medio ambiente. Por ejemplo, una fuente cercana en una sala reverberante puede tener la misma relación D/R que una fuente distante en una habitación con menor reverberación.

En un artículo publicado recientemente por Calcagno et al (2012) se discute la posibilidad de que el aumento en la precisión de los juicios de distancia observados en ambientes reverberantes puede ser explicado, además de la relación D/R, por la asociación entre el tiempo de reverberación y la percepción del tamaño de un recinto. Esta hipótesis parte de una idea simple: si la fuente sonora y el oyente se encuentran en el mismo recinto, distancias a la fuente grandes solo serán posibles dentro de recintos también grandes. Aunque muchos factores contribuyen significativamente al valor exacto del tiempo de reverberación (tamaño de la sala, materiales de las paredes, muebles, etc), en ausencia de información adicional, el sistema nervioso interpreta el tiempo de reverberación en relación directa con el tamaño de la sala: un gran tiempo de reverberación es consistentemente asociado con salas de gran tamaño (Cabrera et al 2006). Tal vez este hecho explique por qué la distancia aparente de una fuente en ambientes reverberantes es significativamente mayor que la percibida en ambientes anecoicos.

### § 3.8 VISIÓN

La influencia de la visión en la PAD será revisada en el Capítulo 6 de este trabajo (El rol de la visión en la percepción auditiva de distancia), donde se presentan además resultados experimentales obtenidos en el marco de esta tesis.

## REFERENCIAS BIBLIOGRÁFICAS

- ASHMEAD D.H., LEROY D. y R.D., ODOM: «Perception of the relative distances of nearby sound sources». *Perception and Psychophysics*, 1990, **47**, pp. 326–331.
- BEGAULT, D.R.: «Preferred sound intensity increase for sensation of half distance». *Perceptual and Motor Skill*, 1991, **71**, pp. 1019–1029.
- : «Perceptual effects of synthetic reverberation on three-dimensional audio systems». *Journal of the Audio Engineering Society*, 1992, **40**, pp. 895–904.
- BÉKÉSY, G.V.: «Über die Entstehung der Entfernungsempfindung beim Hören (On the origin of the sensation of distance in hearing)». *Akustische Zeitschrift*, 1938, **3**, pp. 21–31.
- : «The moon illusion and similar auditory phenomena». *American Journal of Psychology*, 1949, **62**, pp. 540–552.
- BLAUERT, J.: *Spatial hearing. The psychophysics of human sound localization*. Massachusetts, USA, 1997.
- BRONKHORST, A.W. y HOUTGAST, T.: «Auditory distance perception in rooms». *Nature*, 1999, **397**, pp. 517–520.
- BRUNGART, D.S.: «Auditory parallax effects in the hrtf for nearby sources». *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1999, pp. 171–174.
- BRUNGART, D.S y RABINOWITZ, W.M.: «Auditory localization of nearby sources I: Head-related transfer functions». *J. Acoust. Soc. Am.*, 1999, **106**, pp. 1465–1479.
- BRUNGART, D.S y SCOTT, K.R.: «The effects of production and presentation level on the auditory distance perception of speech». *J. Acoust. Soc. Am.*, 2001, **110**, pp. 425–440.
- BUTLER R.A., LEVY E.T. y W.D., NEFF: «Apparent distance of sounds recorded in echoic and anechoic chambers». *The Journal of Experimental Psychology*, 1980, **6**, pp. 745–750.
- CABRERA D., CAPRA A. FARINA A., AZZALI A. y P., MARTIGNON: «Perceived room size and source distance in five simulated concert auditoria». *12th International Congress on Sound and Vibration. Lisbon*, 2006.
- CALCAGNO E.R., ABREGÚ E.L., EGUÍA M.C y R., VERGARA: «The role of vision in auditory distance perception». *Perception*, 2012, **41(2)**, pp. 175–192.
- COCHRAN P., THROOP J. y W.E., SIMPSON: «Estimation of distance of a source of sound». *American Journal of Psychology*, 1968, **81**, pp. 198–206.
- COLEMAN, P.D.: «Failure to localize the source distance of an unfamiliar sound». *J. Acoust. Soc. Am.*, 1962, **34**, pp. 345–346.
- : «An analysis of cues to auditory depth perception in free space». *Psychological Bulletin*, 1963, **60**, pp. 302–315.
- EDWARDS, A.S.: «Accuracy of auditory depth perception». *The Journal of General Psychology*, 1955, **52**, pp. 327–329.
- GAMBLE, E.A: «Intensity as a criterion in estimating the distance of sounds». *Psychological Review*, 1909, **16**, pp. 416–426.
- GARDNER, M.B.: «Distance estimation of 0 degrees or apparent 0 degree-oriented speech signals in anechoic space». *J. Acoust. Soc. Am.*, 1969, **45**, pp. 47–53.

- GRIFFIN, D.R.: «Listening in the dark: The acoustic orientation of bats and men», 1958.
- JESTEADT W., LUCE R.D. y D.M., GREEN: «Sequential effects in judgments of loudness». *Journal of Experimental Psychology: Human Perception and Performance*, 1977, **3**, pp. 92–104.
- LARSEN E., LANSING C.R., IYER N. y A.S., FENG: «On the minimum audible difference in direct-to-reverberant energy ratio». *J. Acoustic. Soc. Am.*, 2008, **124**, pp. 450–461.
- LIENARD, J.S. y BENEDETTO, M.G.D: «Effect of vocal effort on spectral properties of vowels». *J. Acoustic. Soc. Am.*, 1999, **106**, pp. 411–422.
- LOCHNER, J.P.A y DEVONKEET, W.: «Stereophonic and quasi-stereophonic reproduction». *J. Acoustic. Soc. Am.*, 1960, **32**, pp. 393–401.
- LOKKI, T. y PATYNEN, J.: «Lateral reflections are favorable in concert halls due to binaural loudness». *J. Acoustic. Soc. Am.*, 2011, **130(5)**, pp. 345–351.
- LOKKI T., TERVO S. SILTANEN S., PATYNEN J. y SAVIOJA, L.: «Engaging concert hall acoustics is made up of temporal envelope preserving reflections». *J. Acoustic. Soc. Am.*, 2011, **129**, pp. 223–228.
- LOOMIS J.M., PHILBECK J.W., KLATZKY R.L. y R.G., GOLLEDGE: «Assessing auditory distance perception using perceptually directed action». *Perception and Psychophysics*, 1998, **60(6)**, pp. 966–980.
- LOUNSBURY, B.F. y BUTLER, R.A.: «Estimation of distances of recorded sounds presented through headphones». *Scandinavian Audiology*, 1979, **8**, pp. 145–149.
- MERSHON, D.H. y BOWERS, J.N.: «Absolute and relative cues for the auditory perception of egocentric distance». *Perception*, 1979, **8**, pp. 311–322.
- MERSHON, D.H. y KING, L.: «Intensity and reverberation as factors in the auditory perception of egocentric distance». *Perception and Psychophysics*, 1975, **18**, pp. 409–415.
- MERSHON, D.H. y PHILBECK, J.W.: «Auditory perceived distance of familiar speech sounds». *Proceedings of the Psychonomic Society 32nd Annual Meeting*, 1991.
- MERSHON D.H., LITTLE A.D. MCMURTRY P.L., BALLENGER W.L. y J.L., BUCHANAN: «Effects of room reflectance and background noise on perceived auditory distance». *Perception*, 1989, **18**, pp. 403–416.
- MIDDLEBROOKS, J.C. y GREEN, D.M.: «Sound Localization by Human Listeners». *Annual Review of Psychology*, 1991, **42**, pp. 135–159.
- MORIMOTO M., ANDO Y. y Z., MAEKAWA: «On head-related transfer function in distance perception». *Proceedings of the Congress of the Acoustical Society of Japan*, 1975, pp. 137–138.
- NIELSEN, S.H.: *Distance Perception in Hearing*. Aalborg University Press, Aalborg, 1991.
- : «Auditory distance perception in different rooms». *Journal of the Audio Engineering Society*, 1993, **41**, pp. 755–770.
- OTANI, M. y HIRAHARA, T.: «Numerical study on source-distance dependency of head-related transfer functions». *J. Acoustic. Soc. Am.*, 2009, **125(5)**, pp. 3253–3261.
- SCHROEDER, M.R. y LOGAN, B.F.: «Colorless artificial reverberation». *Journal of the Audio Engineering Society*, 1961, **9**, pp. 192–197.
- SIMPSON, W.E. y STANTON, L.D.: «Head movement does not facilitate perception of the distance of a source of sound». *American Journal of Psychology*, 1973, **86**, pp. 151–159.

- STEVENS, S.S. y GUIRAO, M.: «Loudness, reciprocity and partition scales». *J. Acoustic. Soc. Am.*, 1962, **34**, p. 1466–1471.
- STRYBEL, T.Z. y PERROTT, D.R.: «Discrimination of relative distance in the auditory modality: The success and failure of the loudness discrimination hypothesis». *J. Acoustic. Soc. Am.*, 1984, **76**, p. 318–320.
- TRAUNMÜLLER, H. y ERIKSSON, A.: «Acoustic effects of variation in vocal effort by men, women, and children». *J. Acoustic. Soc. Am.*, 2000, **107**, p. 3438–3451.
- WARREN, R.M.: «A basis for judgments of sensory intensity». *American Journal of Psychology*, 1958, **71**, pp. 675–687.
- WENZEL E.M., KISTLER D.J., ARRUDA M. y F.L., WHIGHTMAN: «Localization using nonindividualized head-related transfers functions». *J. Acoustic. Soc. Am.*, 1993, **94**, pp. 111–123.
- ZAHORIK, P.: «Experiments in Auditory Distance Perception». *Tesis Doctoral, Universidad de Winsconsin*, 1998.
- : «Assessing auditory distance perception using virtual acoustics». *J. Acoustic. Soc. Am.*, 2002a, **111**, pp. 1832–1846.
- : «Auditory display of sound source distance». *Proceedings of the 8th International Conference on Auditory. Kyoto*, 2002b, pp. 239–243.
- ZAHORIK P., BRUNGART D.S. y A., BRONKHORST: «Auditory Distance Perception in Humans: A Summary of Past and Present Research». *Acta Acustica united with Acustica*, 2005, **91**, pp. 409–420.



## - Capítulo 4 -

---

---

### Enfoques analíticos de la espacialidad en la música electroacústica

---

#### § 4.1 INTRODUCCIÓN

Como bien indica Zattra (2005), el análisis de la música electroacústica se convierte en una tarea compleja y heterogénea debido a la cantidad de subgéneros que presenta, los cuales ponen el énfasis en determinada característica de la obra: el timbre, la altura, los sistemas de espacialización, el procesamiento en tiempo real, el soporte, la espacialidad misma que se le impregna a la obra, etc. Di Liscia (2005) indica que el principal obstáculo con el que se encuentra la disciplina del análisis musical es justamente la diversidad de estéticas y recursos compositivos que presentan las obras de música electroacústica. Esto hace impracticable la aplicación completa y orgánica de metodologías de análisis útiles para la música tradicional, en obras de música contemporánea. Por otra parte, el análisis de la música electroacústica plantea serias dificultades técnico metodológicas, que se incrementan a causa de que la espacialidad del sonido es difícil de cuantificar a partir de un análisis directo de la señal digital (Di Liscia, 2005). Sin algunos datos resulta difícil saber que quiso hacer el autor. Sin embargo y a pesar de esta dificultad, algunos autores han



logrado generar teorías de análisis de la espacialidad del sonido en la música electroacústica, logrando, en varios casos, acercarse lo suficiente a un análisis abarcativo de la cuestión.

En este capítulo se repasaran algunos enfoques analíticos del espacio que pueden ser utilizadas en una obra de música electroacústica. Si bien existen diversas teorías al respecto se realizará un recorte desarrollando solo algunas de estas, las cuales se eligieron por tener aspectos en común, pero a la vez por ser lo suficientemente autónomas como para tener una riqueza propia, además de que en general estas están basadas en características musicales y perceptivas del espacio.

En primera instancia, se pondrá un énfasis especial en la explicación de la teoría desarrollada por Kendall (2010), pues será la que se utilizará en esta tesis para el análisis de obras de música electroacústica. Luego se revisaran las teorías de análisis desarrolladas por Di Liscia (2000), Liut (2010) y Grela (2003).

## § 4.2 GARY KENDALL – ATRIBUTOS ESPACIALES Y JUEGOS PERCEPTIVOS

En Kendall (2010) se puede observar el desarrollo de una metodología que intenta generar herramientas para el análisis de la espacialidad de la música electroacústica a partir de la distinción de atributos espaciales y esquemas auditivos. Kendall se basa en los trabajos anteriores de Smalley y Rumsey. Del primero específicamente su tan conocida Espectromorfología (Smalley, 1986, 1997) y el desarrollo que hace de esta en cuestiones del espacio y la forma en la música electroacústica (Smalley, 2007). Kendall se nutre de las teorías

de Rumsey (2002) para establecer atributos y características espaciales de los sonidos y así generar un marco conceptual de análisis y producción.

Específicamente Kendall deja en claro que si bien su metodología se basa sensiblemente en la psicoacústica, los compositores electroacústicos no están limitados a trabajar con fuentes preexistentes, de hecho pueden generar sus propias fuentes de sonido, como así explorar acústicas inverosímiles. A partir de esta idea fundamental, Kendall intenta desarrollar una teoría que abarque esta doble característica del análisis de la música electroacústica.

Kendall (2010) parte de una segregación de los elementos que pueden desarrollarse dentro de una escena auditiva, y establece categorías para cada uno de ellos. Partiendo del nivel más bajo, la taxonomía de Kendall define a la FUENTE como unidad, el ENSAMBLE como un agrupamiento de varios elementos FUENTE, luego el RECINTO como un elemento con características acústicas propias que contiene a los elementos FUENTE y ENSAMBLE, y en el nivel más alto la ESCENA (Kendall toma y adapta esta clasificación de Rumsey, 2002).

Sobre cada uno de estos elementos, Kendall propone atributos espaciales que los caracterizan, los cuales divide en Atributos Dimensionales y Atributos Inmersivos.

### **Atributos dimensionales:**

**Ancho** (FUENTE / ENSAMBLE / RECINTO / ESCENA)

**Distancia** (FUENTE / ENSAMBLE / RECINTO / ESCENA)

**Profundidad** (FUENTE / ENSAMBLE / RECINTO / ESCENA)

**Dirección** (FUENTE / ENSAMBLE / RECINTO / ESCENA).

**Atributos inmersivos:****Envolvimiento** (RECINTO/relativo a la FUENTE)**Presencia**

Es interesante observar cómo Kendall define atributos de dirección a la categoría RECINTO o ESCENA, donde se podría pensar que al ser los niveles más altos no podrían poseer estas características. Sin embargo, dentro de la música electroacústica se puede fácilmente tener las características de un RECINTO en un lugar y otro RECINTO en otro lugar (para un ejemplo ver Capítulo 9, Sección 7, 120 Maquina Lirica”). En un nivel más alto, se podría tener diferentes ESCENAS. Quizás se podría plantear una categoría aun más alta que la de ESCENA, por ejemplo la categoría OBRA, la cual puede contener diversas ESCENAS.

También es interesante como Kendall plantea el atributo dimensional “profundidad” a la categoría FUENTE, aunque indica que es casi improbable que una FUENTE tenga profundidad. Una fuente puede tener profundidad sin duda, por ejemplo una FUENTE compleja (un camión, un tren, un piano), que no es un ENSAMBLE, dado que en esta categoría uno puede detectar diversas FUENTES. Como se observa claramente, cuando se comienza a pensar en los términos de Kendall, se pueden lograr combinaciones que la música electroacústica generalmente contiene, lo que muestra la efectividad que puede llegar a tener esta metodología.

En cuanto a los atributos inmersivos, Kendall no propone un desarrollo determinado sino que se limita a definirlos. Por ejemplo, el término “envolvimiento”

está relacionado a la experiencia de sentirse rodeado por fuentes sonoras o un campo difuso como la reverberación, aunque también se lo relaciona a la categoría FUENTE. En este caso es totalmente plausible sentir que se está “dentro de una FUENTE”. En cuanto a la definición del atributo inmersivo (“presencia”), Kendall no ahonda demasiado, y de hecho lo deja para futuros trabajos. Una vez planteadas las categorías y los atributos, tanto dimensionales como inmersivos, Kendall propone una serie de manipulaciones de estos atributos, que terminan definiendo juegos perceptivos con los agrupamientos y los esquemas espaciales, típicos de la música electroacústica. Para esto, primero realiza una descomposición de la palabra “fuente” en cuatro marcos de referencia, que ayudan a pensar los juegos perceptivos de una manera más sencilla: Señal Fuente, Fuente Conceptual, Imagen Fuente y Esquema FUENTE.

**Señal-fuente:** es la señal acústica o una representación de una señal acústica (un archivo de audio grabado en algún soporte, que luego se reproduce por un sistema de espacialización).

**Fuente conceptual:** el objeto que el oyente identifica como fuente, independientemente de sus atributos espaciales. En definitiva, aquellos sonidos con los que el oyente tiene un vínculo de significado (por ejemplo, la voz es una Fuente Conceptual con un nivel de relación con la especie humana casi inigualable).

**Imagen-fuente:** la fuente que tiene atributos espaciales en una escena auditiva, de la cual podemos definir una dirección, una elevación, una distancia, etc, etc. (fuente perceptiva). Es decir, la imagen de la fuente que percibimos efectivamente por los procesos de auralización a la que se ha sometido la señal fuente.

**Esquema espacial FUENTE:** Este es el esquema espacial del oyente, el cual de alguna manera agrupa todos los marcos conceptuales de la palabra “fuente”.

Una vez establecidas estas definiciones, es más fácil entender los juegos artísticos perceptivos que se pueden dar según esta taxonomía.

Las diversas relaciones espaciales entre Fuente/s Conceptual/es e Imagen/es Fuente producen juegos perceptuales de disrupción que en general son los que tienen un especial interés en el juego artístico-estético en las composiciones electroacústicas.

En general, se tiene una noción de cómo se ubica espacialmente una Fuente Conceptual. Por ejemplo, sabemos que la Fuente Conceptual “voz” proviene de la Imagen Fuente que podríamos denominar “persona”. Pero si esa misma voz, proviene de “la misma persona” desde diferentes localizaciones, se produce una ruptura perceptiva importante <sup>1</sup>. Estos son los juegos perceptivos a los que Kendall refiere, y que en general son muy comunes dentro de la música electroacústica. Kendall define diferentes relaciones básicas entre Fuente Conceptual e Imagen Fuente:

**-Una fuente conceptual, una fuente percibida** (esquema FUENTE). Esta podría ser, sin embargo, diferente a las fuentes conceptuales de origen.

Por ejemplo, si una fuente conceptual que tiene un ancho determinado y característico se representa en la obra con una imagen fuente que tiene otras características de ancho manifiestamente distintas.

---

<sup>1</sup>Las comillas en “la misma persona” denotan, justamente, lo paradójico de la situación.

- Múltiples fuentes conceptuales, una fuente percibida** (frontera entre FUENTE Y ENSAMBLE)
- Una fuente conceptual, múltiples fuentes percibidas** (frontera entre FUENTE Y ENSAMBLE)
- Múltiples fuente conceptuales, múltiples fuentes percibidas** (esquema ENSAMBLE). Estas podrían ser, sin embargo, diferentes en cantidad y calidad a las fuentes conceptuales de origen.

Según Kendall, los juegos perceptivos que parten de la manipulación de atributos espaciales, están basados en dos tipos de técnicas:

**Técnicas del tipo 1:** Estas técnicas perturban la identidad de la Señal Fuente ya que la que dividen en partes, en el tiempo, en su espectro o de ambas formas, para realizar algún tipo de procesamiento a cada porción de la señal y luego volver a ensamblarla. Kendall pone como ejemplos de estas técnicas a la “granulación” (separación temporal; Roads, 1996) y el “Phase Vocoding” (transformación espectral; Moorer, 1978). El resultado puede ser la misma señal original “reconstruída”, o la construcción de una nueva señal.

**Técnicas del tipo 2:** técnicas que involucran la ruptura de la formación de eventos perceptivos a través de la manipulación de la información binaural. Esta técnica afecta los atributos espaciales a lo largo del eje izquierda-derecha (manipulación de ILD e ITD; ver Capítulo 1, Localización horizontal de fuentes sonoras).

Kendall también define un tipo de juego perceptivo que involucra las relaciones entre los esquemas auditivos espaciales. Por ejemplo desarrolla la idea del esquema auditivo de CONTENCIÓN o CONTENEDOR. Se trata de un elemento con un espacio interno que contiene otro elemento, que también posee un espacio interno que contiene un objeto sonoro, el cual es transformado por las características de este objeto contenedor.

Una sala es una especie de contenedor, que puede contener tanto al oyente como a los objetos sonoros. Una sala transforma al sonido de modo diferente que otros contenedores, en parte debido a que el oyente se encuentra dentro de ella.

Kendall indica que en la música electroacústica se puede manipular las expectativas normales de estos esquemas auditivos. Esa manipulación de esquemas espaciales auditivos tiene el efecto de dirigir la atención del oyente hacia un contenido resaltado por relaciones inusuales o inesperadas.

Se puede considerar la posibilidad del oyente de estar dentro de un contenedor que no sea una sala. También existe la posibilidad de que un oyente perciba el sonido de una fuente proviniendo de una dirección determinada y el sonido de la sala que contiene a la fuente llegando desde otra dirección.

Como se indicó arriba, en el análisis que se realiza en esta Tesis sobre la obra “120 Máquina Lírica” del grupo Buenos Aires Sonora, existe un juego perceptivo con la noción de contenedor (ver Capítulo 9, Sección 7).

### § 4.3 OSCAR PABLO DI LISCIA – ANÁLISIS DEL TRATAMIENTO ESTÉTICO DE LOS PROCEDIMIENTOS CONSTRUCTIVOS DEL ESPACIO

Di Liscia (2000) ataca la problemática del tratamiento espacial de la música electroacústica y su posterior análisis desde el punto de vista del tratamiento mismo del espacio. Así en una primera instancia indica que como punto de partida para discutir el tratamiento estético de los recursos espaciales se debe determinar la concepción en el tratamiento espacial del sonido:

*”Ya que la concepción espacial que se quiere determinar en una obra está íntimamente ligada a la poética misma de la obra, a partir de allí se pueden establecer los recursos técnicos y su vinculación con otros aspectos de la estructura musical.”* Di Liscia (2005)

A partir de estas definiciones que vinculan la poética espacial con las demás estructuras de la música electroacústica Di Liscia (2000) establece cuatro categorías para el tratamiento espacial de los sonidos: Real, Neutra, Hiperreal y Virtual, donde los compositores eligen desde un primer momento qué características espaciales se quiere imprimir a la obra.

Un tratamiento espacial Realista, implica una simulación estricta de la espacialización del sonido, intentando igualar las características del mundo real. Una espacialidad Neutra minimiza los procedimientos que tienden a conferir a las estructuras sonoras una cualidad espacial. Las fuentes están fijas en el lugar donde los parlantes están y utilizan como ámbito el ambiente donde se reproducen. El tratamiento espacial Hiperreal se asocia a la realidad, pe-



ro su finalidad es ser más contundente y verosímil que esta. Por último, el tratamiento Virtual de la espacialidad de los sonidos en una obra de música electroacústica se realiza de una manera en la cual no sería posible que se desarrolle en realidad. En esta última categoría caerían las múltiples disrupciones entre “fuente conceptual” e “imagen-fuente” presentadas en el enfoque de Kendall.

Según Di Liscia (2000), estas concepciones pueden asociarse directamente a funciones del discurso musical, sobre todo en obras que tengan un carácter referencial. Una vez planteado el tipo de espacialidad que establece una obra, Di Liscia indica la necesidad de establecer puntos de partida analíticos, por lo cual se adopta en su trabajo la posición del analista y sus objetivos (según lo establecido en Nattiez, 1990), y por lo cual se puede abordar la espacialidad de la obra a partir de diferentes puntos de vista. Así se puede establecer un punto de partida poiesico, uno neutro y otro estésico. El análisis poiesico describe las intensiones propias del compositor, el análisis neutro establece solamente el análisis de las estructuras propias de la obra y el análisis estésico estudia la forma en que se percibe la obra. A pesar de que cada análisis presenta su atractivo, sin dudas la combinación de estos produce niveles más profundos de análisis. De esta interrelación se obtienen seis nuevos puntos de partida analíticos (Nattiez, 1990):

1)**Análisis neutro:** El analista pone el acento en la descripción de los elementos constituyentes de la obra musical.

2)**Del nivel neutro al poiesico (Poiesis inductiva):** el analista se basa en datos de la obra y procede luego a comprobar qué relación tienen con la

génesis de la misma.

3)**Del poiésico al neutro (Poiesis externa):** la situación inversa a la anterior, basándose en datos acerca de la producción de la obra (Documentos, publicaciones del compositor, etc.) el analista se dirige hacia la descripción de los rasgos de la misma.

4)**Del neutro hacia el estésico (Estesis inductiva):** basándose en datos de la obra, el analista procede luego a intentar explicar cómo es percibida por un grupo de oyentes. Corrientemente se basa también en teorías acerca de la percepción.

5)**Del estésico al neutro (Estesis externa):** El analista toma datos de un grupo de oyentes acerca de cómo se percibe la obra y busca a posteriori en las estructuras inmanentes de ésta las razones de esta percepción.

6)**Combinación de los tres:** Se utilizan bosquejos para guiar el análisis neutro y luego se especifica cómo la obra debe ser percibida y ejecutada.

Por último, Di Liscia (2000) establece posibles etapas a recorrer al abordar el análisis del espacio en una obra de música electroacústica:

1-Determinación de los parámetros y/o nociones sobre las que se basa principalmente la organización de la obra a analizar a partir de: documentación del autor, los participantes de la producción de la obra u otros estrechamente vinculados hayan dejado. Di Liscia indica que si bien la concepción y percepción de una obra es holística, los autores suelen elaborar más detalladamente las organizaciones de determinados parámetros por sobre otros, que tienen una elaboración secundaria y complementaria.

2-La detección y análisis de los procedimientos y dispositivos de tratamiento espacial del sonido que se ponen en juego en la obra. Esto se basa tanto en las teorías psicoacústicas como en las técnicas de procesamiento digital de señal.

3-La determinación del rol que juega la espacialidad del sonido en la obra en relación con las demás organizaciones de parámetros y/o nociones sonoras.

Es interesante como Di Liscia establece diversos conceptos operativos que son utilizados en esta etapa y que representan el foco del análisis en si, además de establecer un vínculo con la taxonomía propuesta por Kendall (2010):

**-Tratamiento espacial de sonido:** todo procedimiento técnico que confiera a una determinada señal acústica una característica de espacialidad.

**-Ámbito:** la característica de espacialidad por la que una señal acústica parece haberse generado en un determinado espacio físico real o virtual más allá de la fuente que la haya generado o de su eventual localización en un punto de este espacio.

**-Localización:** la característica de espacialidad por la que una señal acústica parece provenir de una fuente real o virtual situada en un determinado punto de un espacio físico real o virtual. La fuente sonora puede estar en posición fija, o moverse en varias direcciones y con velocidades variadas.

**-Indicio:** es un determinado rasgo de una señal acústica que es tomada como base por el sistema auditivo para conjeturar el ámbito, la localización y /o la característica de movimiento de una fuente sonora virtual o real.

#### § 4.4 MARTÍN LIUT – ANÁLISIS ESPACIAL DE LA MÚSICA A PARTIR DE LA PRAXIS COMPOSITIVA

En Liut (2010) se presenta un enfoque analítico producto de la praxis o sea que en este caso el análisis proviene directamente del proceso de composición, es establecido y discutido por los compositores mismos de la obra. En este caso, las composiciones que propone analizar Liut (2010) son las del colectivo de artistas Buenos Aires Sonora, los cuales realizan presentaciones en el espacio público. El Capítulo 9 de esta tesis está dedicado al análisis profundo de algunas de sus obras. Sin embargo, el desarrollo de este tipo de análisis presenta una generalidad que le permite ser utilizada en otros tipos de obra. Cada uno de los elementos que se analizan aparece en prácticamente todas las obras que son reproducidas o ejecutadas en un espacio.

Liut establece entonces una clasificación, en donde partiendo del proceso compositivo se promueve una secuencia de temas a considerar, que si bien no son nuevos para la música, si lo son en cuanto plantean vínculos en la relación obra-espectador. Se puede establecer entonces seis puntos centrales de esta taxonomía: Temporalidad, Espacio, Materialidad, Dispositivos, Presentación y Recepción.

**Temporalidad:** este concepto se establece a partir de las diferentes duraciones que puede tener una obra de este estilo.

- *Intervenciones:* acotadas en el tiempo. Pueden tener la extensión de una obra musical, un concierto hasta abarcar una jornada completa, pero están acotadas temporalmente, tienen un comienzo y un fin. Pueden ser montajes

elaborados previamente a la situación o performances complejas realizadas en vivo y en directo. El término intervención refuerza el carácter efímero y acotado dentro de espacios que no están previstos para el arte. Se diferencia del uso del mismo término por parte de las artes visuales en que las intervenciones sonoras no dejan huellas sobre el objeto o el espacio intervenido.

- *Instalaciones*: extensión de dos o más jornadas, llegando a un mes o un mes y medio. Por definición, involucran o bien montajes o diseños de interactividad previos. Entre sus posibilidades está la de la presencia permanente durante el tiempo en el que la instalación está montada. La instalación sonora también observa una problemática específica en el campo de lo sonoro, ya que el tiempo de observación por parte del espectador es central a la hora de definir la organización o el discurso que se proyectará en un espacio determinado físicamente, pero no temporalmente (el público puede entrar y salir cuando quiere). La idea de una composición orgánica entra claramente en crisis ante esta nueva situación.

**Espacio:** el primer aspecto es el material, concreto, el lugar en el que ocurren las obras. Hasta la fecha BAS ha presentado obras en espacios públicos al aire libre y en instituciones asociadas al arte pero utilizando espacios no habituales para la música. En ambos casos, las obras se caracterizan por su cualidad “site-specific”. Esto es, están pensadas para una ubicación concreta y única; por lo tanto no son, en principio, trasladables. Hay aquí tres cuestiones diferentes para desglosar, las que pueden sumarse o no, según el caso:

- *Trabajo con el espacio simbólico-histórico*
- *Transformación de objetos en instrumentos*

- *Trabajo con la arquitectura aural*

Vinculado a estos dos últimos puntos, el aspecto siguiente a ponderar es lo que llamamos:

**Materialidad:** concretamente es lo que suena. Lo que suena es tanto la arquitectura aural de un espacio determinado (y que por lo tanto le da el color final a lo que, por ejemplo, se emita a través de algún sistema de amplificación), como el repertorio de timbres que puede ofrecer un objeto a ser intervenido (un puente, una escalera, una escultura) o diversos materiales compuestos o montados específicamente para una obra.

**Dispositivos:** a partir de la constatación y el análisis de la materialidad se pasa a la etapa de diseño que se inicia con el orgánico de la obra. Esto define lo que será la cadena electroacústica y de otros soportes artísticos necesarios para presentar la obra.

**Presentación:** el modo en que la obra se presenta al público. Los tres grandes métodos que hemos trabajado hasta hoy son:

a.- *Montajes*

b.- *Performance en vivo*

c.- *Interacción*

**Recepción:** por último, una temática que, salidos de la situación de concierto tradicional, también se plantea como parte del proceso de composición, es la

situación del propio público de la obra. Como síntesis basta recordar que en este tipo de obras, la situación de escucha totalmente contemplativa y estática, propia del concierto tradicional, es solo una de las muchas posibilidades de las que dispone el espectador.

Lo interesante de esta taxonomía es lo abarcativo del procedimiento: contempla obras actuales de arte sonoro y puede ser aplicada a la ejecución de obras clásicas, pero además realiza un cruce desde la producción de la obra, pasando por los sistemas de reproducción, espacios simbólicos e históricos de la música, hasta llegar a la percepción del público de la obra.

#### § 4.5 DANTE GRELA – LA MÚSICA COMO UN ARTE TEMPORAL Y ESPACIAL

Dante Grela (2003) indica que todo evento sonoro queda definido a través de dos coordenadas, el tiempo y el espacio, por lo que considera que la música debiera considerarse como un arte temporal espacial, indicando que la percepción registra la espacialidad de los eventos sonoro, o su conjunto, debiendo indudablemente revisar la clasificación de la música solo como arte temporal. A partir de esto, Grela indica que puede apreciarse una espacialidad real y virtual de la música. Di Liscia (2005) aclara que la espacialidad real es la localización y movimientos de la fuente sonora en un espacio físico por medios tanto instrumentales como electroacústicos. Por el contrario la espacialidad virtual se da por la sensación de espacio o movimiento que generan determi-

nadas organizaciones de parámetros, particularmente la altura y la intensidad. Di Liscia también indica que en general se relaciona la altura con la dimensión vertical y la intensidad con la distancia, aunque como se vio en los primeros tres capítulos de esta tesis no son las únicas pistas necesarias para emular estas dimensiones.

Grela (2003) considera también a la textura y el timbre (en cuanto al espectro) como elementos que aportan a la espacialidad de la música.

Volviendo a la discriminación entre espacialidad Real y Virtual, Grela realiza una enumeración y clasificación de configuraciones que aparecen cuando se localiza o se mueve fuentes en un determinado espacio. Estas configuraciones provienen de la utilización del espacio como un elemento estético de la obra. Así se pueden tener configuraciones fijas y móviles. Las configuraciones fijas son aquellas en donde los atributos del sonido no cambian de ubicación. Es interesante pensar que aun con sonidos fijos puede lograrse la ilusión de movimiento (por ejemplo instrumentistas que “se pasan” sonidos en diferentes puntos del espacio).

Las configuraciones móviles son aquellas en donde las fuentes, y los fenómenos sonoros asociados a ellas cambian su localización en función del tiempo. A partir de esta diferenciación, Grela establece una serie de cambios y configuraciones espaciales que se pueden clasificar de esta manera:

- Desde una configuración fija a otra (también fija, pero diferente).
- Desde una configuración móvil a una fija.
- Desde una configuración fija a una móvil.
- Desde una configuración móvil a otra móvil, diferente.



De esta manera se establece una metodología para el análisis de las relaciones entre fuentes fijas y móviles. En todos los casos, los procesos de cambio entre las diferentes configuraciones espaciales, pueden darse gradualmente (por modulación) o bruscamente (cambio directo). Los desplazamientos de los eventos sonoros o las fuentes sonoras dentro del ámbito de audición pueden realizarse por modulación o cambio directo, y se pueden clasificar en Unitarios (un único tipo de estructura sonora sometida a desplazamientos espaciales) o Múltiples (coexistencia de dos o más estructuras con distintos tipos de desplazamientos).

Grela (2003) también tiene en cuenta en su teoría otras cualidades del movimiento, tales como sentido del giro, tipo de curva, velocidad del desplazamiento, distancia, magnitud y distribución del ámbito total deben ser también tenidos en cuenta y se pueden cambiar de manera gradual o directa.

## REFERENCIAS BIBLIOGRÁFICAS

- DiLISCIA, O.P.: «El espacio de la Imaginación». *Revista del Instituto Superior de Música de la Universidad Nacional del Litoral*, 2000, **7**.
- : «Los modos de vínculo de la concepción espacial del sonido con la poiesis de la música electroacústica». *Tesis Doctoral, Universidad Nacional de Rosario, Facultad de Humanidades y Arte.*, 2005.
- GRELA, D.: «La consideración analítica del espacio en las formas sonoras». *Inedito*, 2003.
- KENDALL, G.: «La interpretación de la espacialización electroacústica: atributos espaciales y esquemas auditivos». *En Música y espacio: ciencia, tecnología y estética. Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 241–259.
- LIUT, M.: «Creación artística y reflexión teórica en la Universidad Pública. La experiencia del grupo Buenos Aires Sonora». *Revista LIS*, 2010, **3(5)**, pp. 163–170.
- MOORER, J.A.: «The use of the Phase Vocoder in Computer Music Applications». *JAES*, 1978, **26(2)**, pp. 42–45.
- NATTIEZ, J.J.: *Music and discourse: Toward a semiology of music*. Princeton University Press, New Jersey, 1990.
- ROADS, C.: *The Computer Music Tutorial*. MIT Press, Massachusetts, 1996.
- RUMSEY, F.: «Spatial Quality Evaluation for reproduced Sound: Terminology, Meaning, and a Scene-based Paradigm». *Journal of the Audio Engineering Society*, 2002, **50(9)**, pp. 651–666.
- SMALLEY, D.: «Spectromorphology and structuring processes». *The language of electroacoustic music*, 1986, pp. 61–93.
- : «Spectromorphology: explaining sound-shapes». *Organised Sound*, 1997, **2(2)**, pp. 107–126.
- : «Space-Form and acousmatic image». *Organised Sound*, 2007, **12(1)**, pp. 35–58.
- ZATTRA, L.: «Analysis and analyses of Electroacoustic Music». *Proceedings of the sound and music computing, Salerno, Italy*, 2005.



## - Capítulo 5 -

---

### Técnicas y sistemas de espacialización

---

Un compositor de música electroacústica puede elegir diversas maneras de estructurar o representar el espacio de su obra. Según Di Liscia (2005) se puede elegir un fuerte efecto espacial emulando un ambiente mediante el uso de reverberación, sin que las fuentes sonoras se perciban en una ubicación determinada o se muevan dentro de este espacio. Por el contrario se puede elegir un efecto inverso donde la localización o el movimiento de las fuentes sonora cumplan un rol importante. Por último, se puede combinar ambas técnicas para lograr un control y una sensación sonora espacial más completa.

En cada uno de estos casos, sería útil para el compositor entender cómo funcionan y se desempeñan las diferentes técnicas y sistemas de espacialización, tanto para aprovecharlos y utilizarlos como para poder rediseñarlas o generar nuevas formas de uso y así lograr una mayor eficiencia en la estructuración del contenido sonoro espacial de la obra.

En cuanto a los sistemas que priorizan la localización o el movimiento de fuentes sonoras, estos se basan en tres técnicas diferentes: panorámico de intensidad, los basados en la reconstrucción binaural, a partir de grabaciones binaurales y los sistemas que se basan en la síntesis de campo sonoro. Con

respecto a las técnicas que emulan las características acústicas de un espacio determinado, sin buscar la localización certera de las fuentes sonoras, podemos encontrar las técnicas que utilizan respuestas impulso de ambientes (room impulse response), las técnicas que calculan estas respuestas impulso mediante el modelado de las características de un ambiente, y como estas respuestas impulso pueden ser utilizadas para traspasar las características acústicas de un determinado ambiente a otros sonidos previamente grabados.

El fin de este capítulo será entonces realizar una revisión de estas diferentes técnicas de espacialización como descripción previa de algunos de las técnicas y sistemas utilizados en las obras analizadas en esta tesis.

## § 5.1 TÉCNICAS Y SISTEMAS

A la hora de elegir un sistema de reproducción y espacialización de audio se puede hacer hincapié en varios aspectos que en definitiva tienen que ver directamente con querer o no querer copiar las características reales de una determinada o determinadas fuentes sonoras. Para esto siempre se debe tener en cuenta varias características del sistema, como lo pueden ser la respuesta en frecuencia, los niveles de intensidad y sonoridad, el rango dinámico, el patrón espacial, las características de reverberación. Un sistema de reproducción puede estar compuesto por un solo parlante o por decenas de ellos, pero en todo caso siempre se intentará utilizar el sistema más adecuado para lo que se desea reproducir. Por supuesto que siempre se busca la mejor calidad de sonido posible y esto responde a las características que arriba se detallan, pero también se debe tener en cuenta, y mucho, que tipo de señal queremos

reproducir y de qué manera.

Por lo tanto este capítulo se centrará en cómo funcionan básicamente algunos de los sistemas más utilizados y también se aclarará que beneficios y desventajas presentan a la hora de ser utilizados en composiciones. Para esta tesis es fundamental realizar la explicación clara de estos sistemas porque al desarrollar el trabajo de campo realizado en esta, se explicará cómo fueron utilizados estos sistemas tanto para el diseño de programas específicos para la espacialización virtual del sonido, como para el desarrollo de obras de música electroacústica.

David Malham (2010) indica que para entender la forma en que los compositores electroacústicos utilizan el espacio sonoro en sus obras es apropiado considerar las tecnologías que estos utilizan para concretar sus propias ideas. En general las técnicas de espacialización se basan en teorías perceptivas y acústicas. Por ejemplo la técnica de panorámico de intensidad se basa en las teorías de escucha en el plano horizontal, ILD e ITD, las técnicas binaurales en las teorías que postulan la forma en que el torso, cabeza y pinna filtran los sonidos para detectar la localización de una fuente en cualquier plano y las técnicas de síntesis basan su funcionamiento en la imitación del frente de onda de un sonido o la emulación de cómo se comporta la presión sonora en un ambiente específico.

Si bien uno de los objetivos de los sistemas de reproducción espacial de sonido es conseguir reconstruir un ambiente sonoro alrededor del oyente, de tal forma que este no sea capaz de distinguir el ambiente ficticio del real (Bleda Pérez, 2009), podemos decir que esto es casi imposible de realizar, debido a que no solamente la información sonora permite la inmersión en un estado de reali-

dad virtual. Además la música electroacústica no necesariamente persigue este hecho. Kendall (2010) indica que las percepciones espaciales son definidas en parte por el contenido de las fuentes sonoras (Mason, 2002; Mason et al, 2005) y los compositores electroacústicos no están limitados a fuentes acústicas pre-existentes. Los compositores electroacústicos no tienen solamente la libertad de diseñar sonidos que dan soporte específico a efectos de espacialización, sino que pueden también explorar ámbitos acústicos que pueden derribar la validez ecológica (Kendall, 2010). En resumen, cada compositor puede crear el espacio que más desee, o que mejor represente a su idea compositiva.

Sin embargo, es importante conocer las posibilidades técnicas de las cuales el compositor puede servirse a la hora de emular las cualidades acústicas de un ambiente o proporcionar la ubicación espacial de los sonidos y el movimiento de estos en un espacio determinado, por que justamente la apropiación de estas técnicas y herramientas permitirán al compositor no limitar su idea de tratamiento espacial. Puede esto compararse a las técnicas que un pintor debe aprender para generar las texturas y colores que desea. La apropiación de las técnicas, si bien no hacen al artista, permite plasmar las ideas de una forma más precisa.

Como se indicó en la introducción, si bien existen una cantidad importante de sistemas de reproducción espacial capaces de ubicar y mover fuentes en un espacio determinado, en general estos se basan en la aplicación de tres técnicas diferentes:

- 1) Sistemas basados en el panorámico de intensidad.
- 2) Reconstrucción binaural mediante grabaciones binaurales y técnicas tran-

saurales.

### 3) Síntesis de campo sonoro.

Dentro de los primeros se desarrollará la utilización de dos canales estéreo por ser uno de los más conocidos y utilizados (a pesar de sus limitaciones), los sistemas cuadrafónicos y octofónicos, los sistemas surround, que utilizan esta técnica de panorámico de intensidad en parte de su desarrollo y el sistema VBAP que presenta una extensión de esta técnica.

Para los sistemas que utilizan el segundo tipo de técnicas de espacialización se hará un resumen de cómo generar grabaciones binaurales para crear escenas más realistas. Como un ejemplo de los sistemas que utilizan las síntesis del campo sonoro se revisaran los sistemas Ambisonics y se realizará un breve comentario de una técnica que está creciendo en los últimos años, Wave Field Synthesis (Síntesis de campo de Onda).

Por otro lado, si bien existen muchos ejemplos posibles, se revisaran dos métodos para obtener y modelar las características acústicas de un recinto mediante la medición o modelado de la respuesta impulso de los mismos. Se comentaran las técnicas modernas de grabación de respuesta impulso monoaural y binaural, y la técnica de la imagen para obtener una respuesta impulso sintetizada a partir de características volumétricas de una sala cuadrada.

## § 5.2 PANORÁMICO DE INTENSIDAD

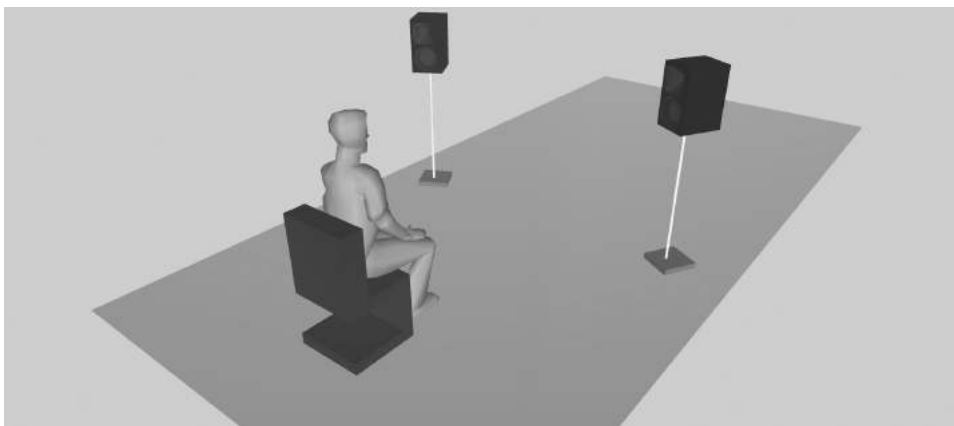
Mediante esta técnica se intenta que el oyente perciba una fuente sonora en una ubicación donde en realidad no se encuentra (fuente fantasma, del inglés,



phantom source). El llamado panorámico de intensidad (del inglés, intensity panning) se basa en la diferencia interaural de intensidad o ILD (Di Liscia, 2005).

La forma de lograr que la fuente fantasma se ubique en una posición angular determinada es justamente haciendo que la señal sea distribuida entre pares de parlantes, de forma tal que la intensidad este escalada de forma proporcional en cada uno de estos. Este escalamiento puede estar regido por varias funciones matemáticas como la lineal, la ley del seno, del coseno y la ley de la tangente (también llamada de energía constante o panorámico coseno/seno). Para una revisión exhaustiva de la aplicación práctica del panorámico de intensidad ver Di Liscia (2010).

La configuración más conocida es justamente el par estéreo (Figura 5.1) donde el sujeto se ubica entre dos parlantes, y la fuente fantasma puede localizarse entre estos dos.

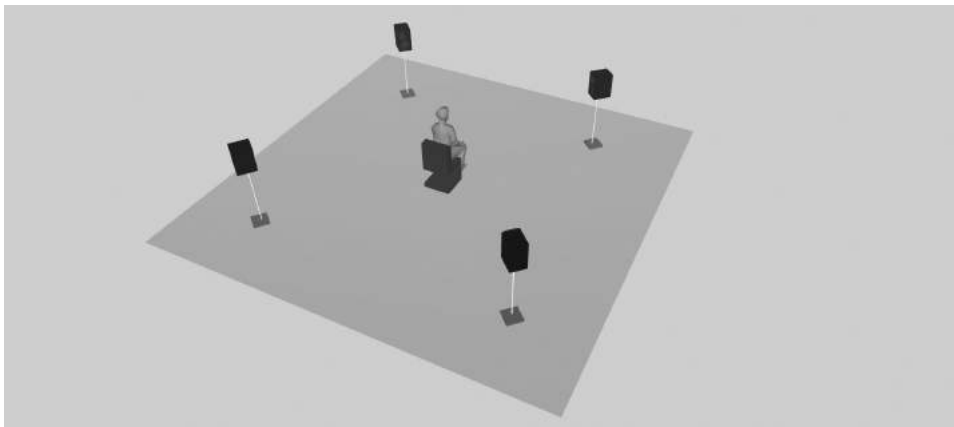


**Figura 5.1:** *Ejemplo de un sistema estéreo*

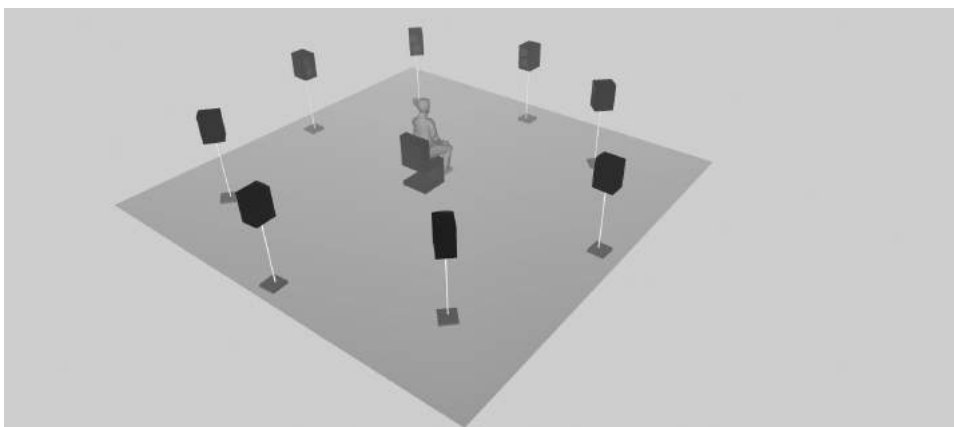
Dentro de la música electroacústica se pueden encontrar otras configuraciones de parlantes como la cuadrafonía (Figura 5.2) o la octofonía (Figura 5.3) que utilizan la técnica de panorámico de intensidad eligiendo entre que pares se encuentra la o las fuentes sonoras. Una de las ventajas de utilizar este tipo de

arreglos es que se gana la posibilidad de ubicar fuentes por detrás del oyente, expandiendo así las posibilidades.

Lo interesante de la técnica de panorámico de intensidad es que en realidad uno puede utilizar la cantidad de parlantes que desee y aplicarla sin problemas, como es el caso de una de las obras que se analizará más adelante (Capítulo 9, Sección 5, Oí(r) el ruido), donde se utiliza un arreglo de 32 parlantes (Subia Valdez, 2010).



**Figura 5.2:** *Ejemplo de sistema Cuadrafónico*



**Figura 5.3:** *Ejemplo de sistema Octofónico*

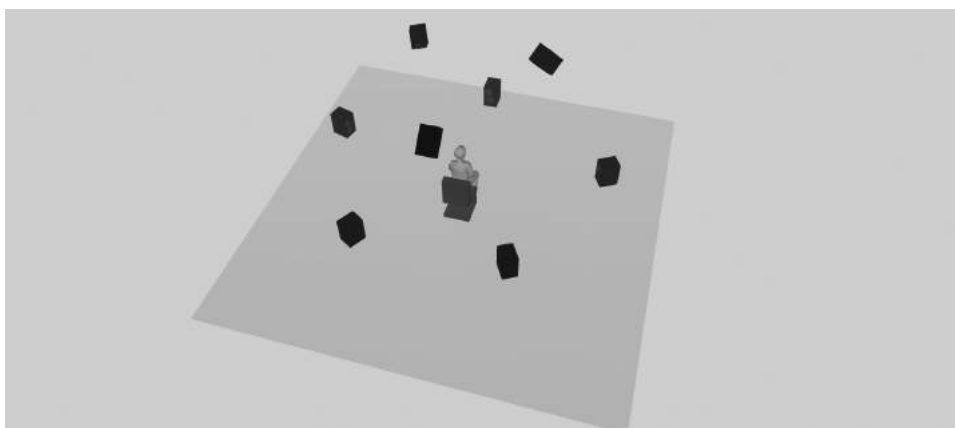
Una de las desventajas obvias de la técnica del panorámico de intensidad es que en muchos casos no se logra una imagen homogénea, o sea, en diferentes ubicaciones la fuente virtual se comporta y se percibe de forma diferente. Por

ejemplo cuando la fuente virtual está en la misma posición que uno de los parlantes, se logra una imagen puntual y nítida, pero cuando la fuente virtual está compuesta por la energía que emana de dos parlantes, la percepción se vuelve más borrosa, más aun cuando la separación de estos es muy amplia (Di Liscia, 2005).

### § 5.3 PANORÁMICO DE INTENSIDAD BASADO EN VECTORES - VBAP (VECTOR BASE AMPLITUDE PANNING)

Esta técnica fue presentada por Pulkki en 1997 y se trata de un método de reconstrucción de fuentes virtuales que emplea únicamente la teoría del panorámico de intensidad pero ampliándola a escenas de tres dimensiones (Pulkki y Karjalainen, 2001).

El funcionamiento de VBAP se basa en la utilización de ternas de parlantes ubicados en los vértices de triángulos (Figura 5.4)



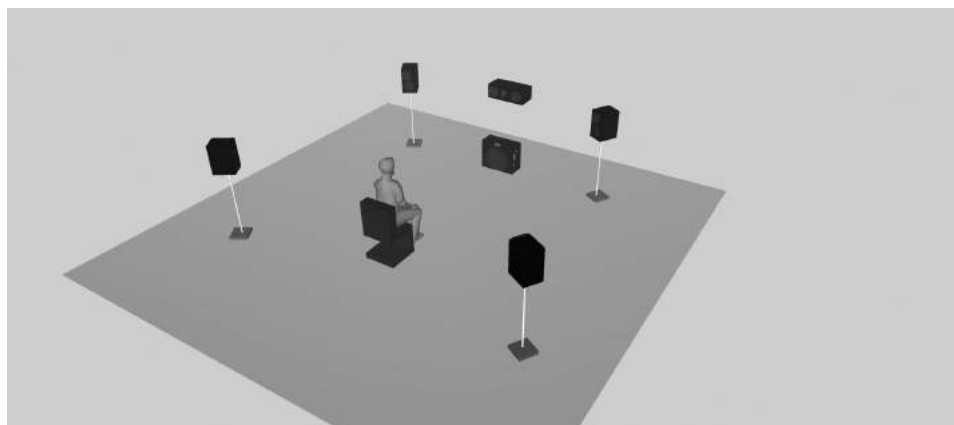
**Figura 5.4:** *Ejemplo de un sistema VBAP*

Según Bleda Pérez (2009) La síntesis de la imagen fantasma se logra en dos fases: primero se determina en que triangulo se encuentra la fuente, para posteriormente calcular la ganancia que debe aplicarse a cada uno de los

tres parlantes ubicados en ese triángulo. Las ecuaciones que se utilizan para calcular la intensidad de cada parlante pueden ser revisadas en Pulkki (1997, 1999, 2001a y 2001b). Un ejemplo concreto de utilización de dichas ecuaciones se puede encontrar en (Di Liscia, 2010).

## § 5.4 SISTEMAS SURROUND

Estos sistemas, muy conocidos en la industria audiovisual, son generalmente la solución comercial al problema de generar un campo sonoro más inmersivo, pero sin embargo no dejan de ser la ampliación de un sistema estéreo mediante el agregado de parlantes por detrás del oyente, sistemas de subwoofers que remarcan las frecuencias graves y el agregado de reverberación. Así la industria audiovisual a patentado los sistemas 5.1, 6.1, 7.1 y hasta 22.2, donde el primer dígito señala la cantidad de parlantes utilizados que envuelven al oyente, con el agregado de parlantes opcionales (el numero que aparece luego del punto) cuyo contenido espectral se restringe a un 10 % del espectro, y que se utilizan específicamente para enfatizar frecuencias graves (Figura 5.5).



**Figura 5.5:** *Ejemplo de sistema surround 5punto1*

En general se trata de un par de parlantes estéreos ubicados en el frente a la

izquierda y derecha del oyente, con una dirección de  $30^\circ$ , con el agregado de un parlante central que estabiliza la posición de los sonidos en los diálogos (este agregado proviene del uso de estos sistemas en las grandes pantallas de cine). Se agregan además parlantes traseros que no se utilizan para localizar sonido como lo hace el estéreo de frente, sino que radian reflexiones y sonido difuso, lo que agrega la sensación de estar más inmerso en un ambiente virtual. También se dice que al poder generar reflexiones laterales, la percepción de distancia mejora notablemente (recordemos que la relación D/R es una de las pistas reconocidas de la percepción auditiva de distancia – Capítulo 3, localización en distancia de fuentes sonora). Estos parlantes traseros se ubican a  $110^\circ$  y a  $-110^\circ$  del oyente. Al estar separados entre sí a más de  $60^\circ$  no pueden usarse para generar fuentes virtuales. Aunque con el paso del tiempo se han agregado parlantes para mejorar esto, pero no se utilizan generalmente más que para generar ambientes.

Podemos deducir entonces, que estos sistemas tienen más en cuenta la idea de inmersión en un determinado ambiente que la localización de las fuentes sonoras, más allá del par estéreo frontal. Por supuesto que el hecho de generar un determinado ambiente puede ser buscado por compositores de música electroacústica, pero no de esta manera. Este tipo de sistemas van de la mano con la proyección de películas, imágenes que ayudan a modelar los espacios sonoros aportando información visual, y que muchas veces capturan la ubicación de las fuentes sonoras.

Para una revisión más exhaustiva de estos sistemas se puede consultar a Bleda Perez (2009) y a Cura (2010).

## § 5.5 RECONSTRUCCIÓN BINAURAL

La reconstrucción binaural, en donde se intenta generar y reproducir señales que contengan las características de ILD e ITD y filtros del torso, cabeza y pinna, es generalmente aplicada a la escucha con auriculares.

Al utilizar auriculares podemos elegir la sola lateralización de los sonidos o la emulación de todo el espacio en tres dimensiones mediante la utilización de las funciones de transferencia relativas a la cabeza o HRTF (del inglés, Head Related Transfer Function).

La lateralización mediante auriculares (Moore, 2012 – Capítulo: Space Perception) se basa en la utilización de la técnica de panorámicos de intensidad en auriculares para emular la localización en acimut de una fuente sonora, además de la posibilidad de agregar diferentes tipos de filtrado y reverberación para la emulación de distancia y tamaño de recintos. Si bien esta técnica es una de las más utilizadas dentro de la práctica musical comercial presenta bastantes inconvenientes a la hora de construir un diseño espacial semejante a la percepción de fuentes reales. Uno de los problemas más evidentes es que los sonidos se escuchan dentro de la cabeza, o sea no se puede generar una sensación de que los sonidos se están moviendo en un espacio real, sino que solamente escuchamos que se mueve entre los auriculares. Otro problema es que muchas veces no se logra una emulación eficaz de una sala, debido a que un algoritmo que genera un tipo de reverberación realista necesita información sobre la localización de cada uno de los rebotes que produce la fuente en las diferentes superficies, y esto a menudo refleja una relación entre la cantidad de procesamiento necesario y el producto final bastante pobre. Por último, no se

puede emular la elevación de una fuente sonora, y mucho menos la ubicación detrás de la cabeza.

Una técnica que salva estos inconvenientes se basa en la utilización de grabaciones binaurales. Este tipo de grabación se realiza mediante la utilización de pequeños micrófonos omnidireccionales, y con respuesta en frecuencia plana en todo el rango audible, ubicados en la entrada del canal auditivo, de una persona o de un maniquí diseñado para tal fin). Al grabar un sonido ubicado a determinada distancia y dirección de un oyente que tiene puestos estos micrófonos, se logra obtener la función de transferencia relativa a la cabeza o mejor conocida como HRTF (del inglés, Head Related Transfer Function) la cual contiene toda la información del filtrado que se produce en el espectro de frecuencias del sonido cuando este “choca” con el torso, la cabeza y la pinna. Entonces el resultado de esta grabación no es ni más ni menos que un sonido que contiene toda la información de localización en acimut, elevación y distancia. Si se reproduce esta grabación con auriculares (de ser auriculares intraurales, el efecto es aun mejor) el oyente tendrá la sensación de estar escuchando un sonido por fuera de la cabeza y con el acimut, la elevación y la distancia que tenía el sonido grabado, además de recoger cada una de las características de la sala donde se reprodujo el estímulo.

Otra forma de poder obtener información binaural para dotar a los sonidos de estas características es utilizando las respuestas impulso relativas a la cabeza o mejor conocidas como HRIR (del inglés, Head Related Impulse Response). La obtención de estas respuestas a impulso se realiza utilizando micrófonos binaurales, como en el caso de las HRTF, pero con técnicas y estímulos que permitan obtener una señal impulsiva que contenga todas las características

espaciales de una ubicación espacial dentro de un recinto determinado (Por ejemplo grabando un barrido exponencial de entre 20Hz y 20Khz de 10 segundos - Farina, 2000). La forma y contenido de las HRIR varían con la posición en el espacio del estímulo y con las características acústicas del entorno donde esta se reproduce. Una vez obtenidas las HRIR de las posiciones que se desean emular, se puede traspasar las características espaciales de la ubicación espacial y el ambiente a cualquier otro sonido mediante la técnica de convolución, la cual se revisara más adelante. Esta es quizás una de las mejores técnicas para obtener información de un entorno acústico y la posición de una fuente sonora, que sin embargo presenta algunas ventajas y desventajas. La principal desventaja de esta técnica es que debido a que las características antropomórficas de cada ser humano son diferentes, las grabaciones binaurales funcionan de manera efectiva solo de forma individual. Por lo tanto una grabación binaural funciona solo para la persona a la cual se le realizó dicha grabación. Sin embargo al escuchar estas grabaciones de forma no individual se pueden percibir ciertas características del ambiente, principalmente en acimut. Las posiciones frente a los oyentes y la elevación, como la ubicación delante atrás, se ven claramente perjudicadas, justamente porque los principales cambios antropomórficos de una persona se da en la conformación de la cara y el tamaño y forma de la pinna.

Otro inconveniente de la utilización de esta técnica, es que para que funcione de manera totalmente efectiva se debería contar con un sistema que detecte los movimientos de la cabeza del oyente que tiene puestos los auriculares y calcule el ángulo de giro de la misma, para luego aplicar la correspondiente HRIR. De no contar con este sistema, toda la escena giraría al mover la ca-



beza, rompiendo la percepción de la escena auditiva.

Para una revisión más exhaustiva sobre este tema consultar a Brown y Duda (1997), Huopaniemi y Riederer (1998), Cheng y Wakefield (2001) y Cetta (2007).

## § 5.6 SÍNTESIS DEL CAMPO SONORO

En la síntesis del campo sonoro se pretende recomponer el campo sonoro por completo en toda la sala, y no solamente la localización de las fuentes sonoras o una sensación de ambiente acústico (Bleda Perez, 2009). De esta manera la escucha no depende de la posición del oyente, ni si mueve o no la cabeza, sino que el espacio sonoro es sintetizado en su totalidad.

Existen por lo menos dos técnicas que utilizan la síntesis del campo sonoro como herramienta para generar sensación de espacios sonoros y localización de fuentes sonora: Ambisonics y la Síntesis de Campo de Onda, más conocida como WFS (del inglés, Wave Field Synthesis). A continuación se realizará un resumen de estas dos técnicas de síntesis del campo sonoro.

### 5.6.1. Ambisonics

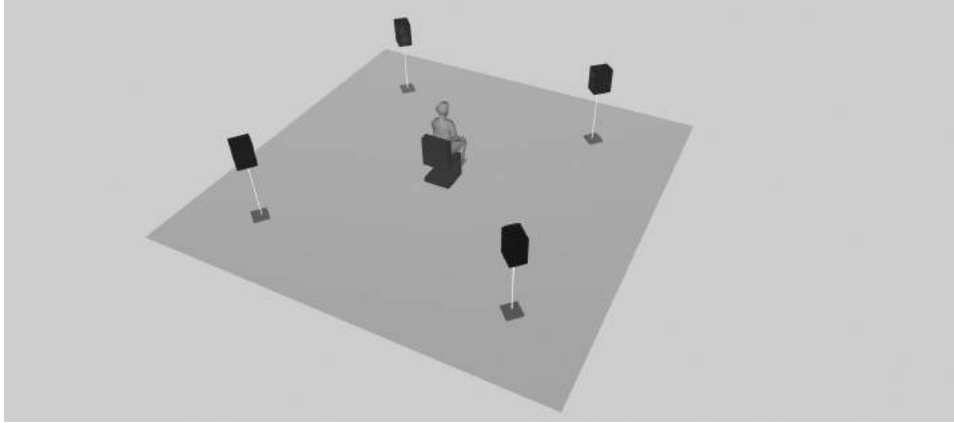
David Malham, un gran estudioso de los sistemas Ambisonics, expone que Ambisonics es una solución tecnológica en dos partes al problema de codificar direcciones de fuentes sonoras, desde la grabación, y reproducirlas (decodificarlas) luego en un sistema práctico de parlantes, logrando que un oyente ubicado en el lugar adecuado perciba estas fuentes en el mismo lugar y dirección en que se grabaron (Malham, 2010). Como bien indica Bleda Pérez

(2009), esta técnica de sonido espacial se fundamenta en distintas técnicas de grabación microfónica, ya que utiliza un arreglo de 4 micrófonos para reconstruir el campo sonoro en sus tres dimensiones (micrófonos SoundField). En este arreglo se utiliza un micrófono omnidireccional que registra la presión sonora (cuya señal se denomina W), y tres micrófonos bidireccionales que registran gradientes de presión en los tres ejes, cuyas señales se denominan X, Y y Z (Gerzon, 1973, 1975).

El conjunto de señales grabadas con este array de micrófonos es denominado formato B, y puede presentar diferentes órdenes de reproducción, ya que las cuatro señales grabadas con el micrófono pueden decodificarse para cubrir un espacio en los 360° del plano horizontal (sistemas pantofónicos) o en toda la superficie de una esfera completa (sistemas perifónicos). Los sistemas de primer orden, los cuales están basados en diversos trabajos anteriores (Felgett, 1972; Gerzon, 1973; Cooper y Shiga, 1972; Gibson et al, 1972) solo utiliza tres de estas señales (W, X e Y) para representar la ubicación de sonidos de manera perifónica. El cuarto canal (Z) se utiliza para la ubicación de fuentes sonoras en el ámbito perifónico (Malham, 2010).

Para la reproducción pantofónica se utilizan 4 parlantes (Figura 5.6) necesitando de un cubo de 8 parlantes para reproducir también la elevación de las fuentes sonoras (Figura 5.7).

Es posible también sintetizar toda una escena sin necesidad de utilizar el array de micrófonos descrito anteriormente. Esto se realiza mediante la simulación del funcionamiento de los micrófonos, que en el formato B responden a las siguientes ecuaciones:



**Figura 5.6:** *Ejemplo de sistema surround 5punto1*

$$W(t) = s(t)$$

$$X(t) = s(t)\cos(\text{acimut})\cos(\text{elevacion})$$

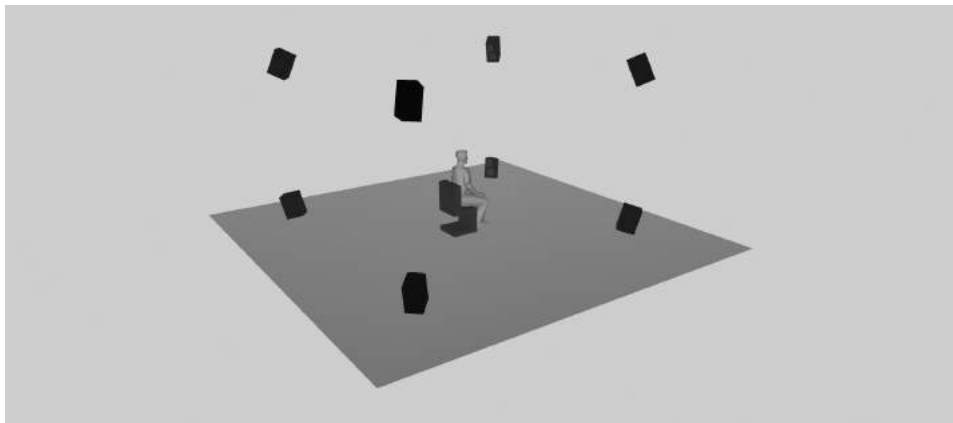
$$Y(t) = s(t)\text{sen}(\text{acimut})\cos(\text{elevacion})$$

$$Z(t) = s(t)\text{sen}(\text{acimut})$$

Donde  $s(t)$  es la señal que deseamos espacializar,  $\text{acimut}$  y  $\text{elevación}$  son los ángulos de localización en el plano vertical y horizontal que deseamos que tenga la fuente, y  $W(t)$ ,  $X(t)$ ,  $Y(t)$  y  $Z(t)$ , son las señales obtenidas. De esta manera tendríamos las cuatro señales necesarias para alimentar un decodificador Ambisonics. En el Capítulo 8, Sección 4 (GUISPA), se repasará un programa informático para manejo de fuente sonoras en un espacio virtual, el cual utiliza un codificador y decodificador Ambisonics para ubicar fuentes en el espacio. Este software fue utilizado para la performance de la obra “120 Máquina Lírica” la cual se analiza también en esta Tesis (Capítulo 9, Sección 7, 120 Máquina Lírica).

El uso y funcionalidad de Ambisonics puede extenderse a grados más altos, lo que mejora las condiciones de escucha ya que el llamado *sweet spot* (lo

que sería el punto de escucha ideal) se amplía directamente al subir el grado. Para una revisión más exhaustiva de este sistema de espacialización se puede consultar Malham (2010) y Bleda Pérez (2009).



**Figura 5.7:** *Ejemplo de sistema surround 5punto1*

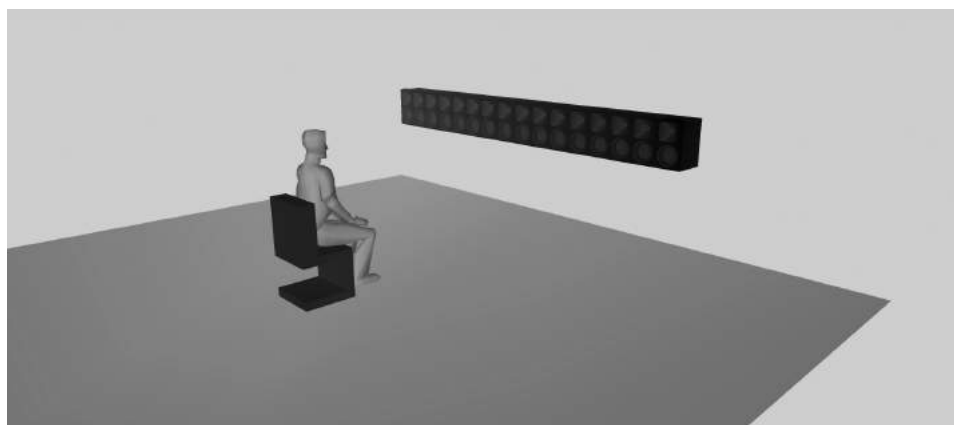
### 5.6.2. WFS – Wave Field Synthesis: síntesis de campo de onda

Uno de los sistemas de síntesis del campo sonoro que más ha crecido en los últimos años es el llamado Wave Field Synthesis (WFS - Figura 5.8). Según Bleda Pérez (2009) WFS es una técnica de reproducción análoga a la holografía en óptica.

Este autor declara que el concepto de “cortina acústica” es quizás el más adecuado para describir el funcionamiento. Este concepto se basa en la idea de que si tenemos una cortina con infinidad de micrófonos (con una buena sensibilidad, respuesta en frecuencia plana, etc), esta podría grabar exactamente como llega el frente de onda de una fuente sonora. De manera inversa, si la señal que llega a cada uno de estos micrófonos fuera reproducida por parlantes que se encuentren en la misma posición en la que estaban los micrófonos, se podría reconstruir de manera exacta el frente de onda del sonido grabado. Por

supuesto que es imposible tener infinitos micrófonos e infinitos parlantes, pero esta teoría ha podido ser puesta en funcionamiento con un número menor de parlantes y micrófonos, con muy buenos resultados. Estos conceptos fueron descritos por Berkhout (1987 y 1988), aunque empezó a implementarse ya en la década de 1990, con el nombre de WFS (Berkhout et al , 1992, 1993).

Para mayor información sobre este tipo de sistemas se puede revisar la excelente Tesis Doctoral de Bleda Perez (2009) la cual desarrolla en profundidad las herramientas teóricas y matemáticas para el funcionamiento de WFS, presentando además muchas mejoras y aplicaciones al sistema.



**Figura 5.8:** *Sistema básico de WFS*

## § 5.7 EMULACIÓN DE LAS CARACTERÍSTICAS DE UN AMBIENTE ACÚSTICO MEDIANTE GRABACIÓN Y MODELADO DE RESPUESTAS IMPULSO

Si solamente se desea obtener información sobre las cualidades acústicas de un ambiente determinado, como la reverberación, se puede trabajar con las técnicas de grabación de respuestas impulso. Las respuestas impulso con señales obtenidas mediante determinada forma de grabación que permiten guardar las cualidades temporales y espaciales de un determinado espacio. Estas res-

puestas impulso pueden ser luego utilizadas para transferir estas cualidades acústicas a otros sonidos, obteniendo de esta manera un estímulo sonoro que parece estar inmerso en este ambiente grabado.

Para obtener la información acústica de un ambiente se debe poder grabar la respuesta impulso utilizando un estímulo que contenga información en todo el espectro audible, para justamente obtener el comportamiento del ambiente en todo el espectro. Además se debe contar con un micrófono que posea una respuesta en frecuencias que abarque todo el espectro audible y la cual debe ser lo más plana posible. Si bien han existido muchas técnicas para grabar la respuesta impulso de una sala, quizás una de las más efectivas sea la técnica Swept-Sine (Farina, 2000), en la cual se utiliza como estímulo una señal senoidal que barre en frecuencia a todo el espectro audible (entre los 22Hz y los 22Khz, en unos 15 segundos – Farina, 2003). Según Farina, La señal obtenida es luego procesada mediante deconvolución para obtener la respuesta impulso. La deconvolución es obtenida mediante una convolución lineal con un filtrado inverso apropiado (Farina, 2000).

Como se indicó anteriormente esta respuesta impulso puede luego ser utilizada para plasmar las características del ambiente en otras señales, las cuales generalmente fueron grabadas en cámara anecoica para que no tengan ningún tipo de reverberación propia. El traspaso de estas características se realiza mediante la técnica de convolución, la cual será desarrollada en el Capítulo 8, Sección 2, Gtconvolve, en el cual se presenta un programa informático para realizar convolución múltiple.

Otra forma de obtener respuestas impulsos es modelándolas mediante un programa de simulación. Así podemos encontrarnos con varios métodos que, me-

dian­te los pa­rá­me­tros ar­qui­tec­to­ní­cos de la sala (an­cho, al­to y lar­go, ti­po de ma­te­ri­ales con los que está con­strui­da, etc), puede cal­cu­lar una res­pues­ta im­pul­so bas­tan­te apro­xi­ma­da a la re­ali­dad. Las téc­ni­cas de mo­de­la­do de am­bie­n­tes para ob­te­ner res­pues­tas im­pul­so, cuyas ca­rac­te­rís­ti­cas pue­dan ser trans­fe­ri­bles a otros soni­dos, se pue­den di­vi­dir en tres téc­ni­cas:

- Las téc­ni­cas ba­sa­das en el tra­za­do de rayos (ray tracing).
- Las téc­ni­cas ba­sa­das en el cál­cu­lo de com­por­ta­mien­tos de las on­das so­no­ras (wave based – re­solu­ción de la ecua­ción de la on­da).
- Las téc­ni­cas ba­sa­das en mé­to­dos es­ta­dís­ti­cos.

En ge­ne­ral un mo­de­la­do ade­cu­ado que abar­que todo el ran­go au­di­ti­vo, re­quiere una com­bi­na­ción de estas téc­ni­cas. Por ejem­plo, me­diante las téc­ni­cas ba­sa­das en on­das se pue­de mo­de­lar la res­pues­ta del am­bie­nte a ba­jas fre­cuen­cias de ma­ne­ra efec­ti­va y con téc­ni­cas ba­sa­das en el tra­za­do de rayos se pue­de mo­de­lar la res­pues­ta a al­tas fre­cuen­cias (Savioja, 1999).

Si lo que se de­sea es re­ali­zar estos cál­cu­los en tiem­po real, por ejem­plo para apli­ca­cio­nes de re­ali­dad vir­tual, se debe te­ner en cuenta que el pro­ce­sa­mien­to se vuel­ve más pe­sa­do, y por lo tan­to al­gunas de estas téc­ni­cas pue­den ser in­ade­cu­adas.

A esta al­tu­ra vale la pe­na acla­rar que la apro­xi­ma­ción más sim­ple al cál­cu­lo en tiem­po real de la res­pues­ta im­pul­so de una sala es asu­mir que la fuente es un pun­to om­ni­di­rec­cio­nal, ya que tí­pi­ca­mente la di­rec­ti­vi­dad de la fuente so­no­ra afec­ta a las fre­cuen­cias que com­ponen dicha fuente (Moore, 1990 - Ca­pí­tu­lo: Rooms).

También es necesario tener en cuenta la propagación del sonido en el aire, la cual también es modelada, así como también la energía del sonido directo (sensible a la distancia como vimos en el Capítulo 3, localización en distancia de fuentes sonoras) y las reflexiones tempranas, la atenuación de las frecuencias, los límites del entorno elegido y el agregado de reverberación densa (Savioja, 1999; Huopaniemi, 1999).

En la auralización en tiempo real, para simplificar se modelan solo el sonido directo y las primeras reflexiones de forma Individual y luego las reverberación densa mediante una estructura recursiva de filtros digitales.

Un ejemplo interesante de modelos que pueden ser utilizados para obtener respuesta a impulsos de una sala simple con un costo de procesamiento relativamente bajo, es el método de la imagen (McGovern, 2008), el cual se puede aplicar para el cálculo del sonido directo y las primeras reflexiones.

En general, luego de este cálculo se puede añadir la reverberación densa mediante una estructura de filtros recursivos (Savioja, 1999).

Recientemente han adquirido interés métodos más exigentes computacionalmente basados en ondas, como el método de elementos finitos (FEM), métodos de elementos de contorno (BEM) y el método de diferencias finitas en el dominio del tiempo (FDTD). Como indicamos antes, estas técnicas son adecuadas para la simulación de bajas frecuencias solamente. Los métodos de modelado estadístico, como el análisis estadístico de energía (SEA), son principalmente aplicada en predicción de niveles de ruido en sistemas acoplados, en donde la transmisión del sonido por estructuras es un factor importante. Estos métodos no son adecuados para la auralización por que en general no modelan el comportamiento temporal de un campo sonoro.





## REFERENCIAS BIBLIOGRÁFICAS

- BERKHOUT, A.J.: *Applied Seismic Wave Theory*. Elsevier Science, USA, 1987.
- : «A holographic approach to acoustic control». *J. Audio Eng.Soc.*, 1988, **36(12)**, pp. 977–995.
- BERKHOUT A.J, VOGEL P. y DEVRIES D.: «Use of wave field synthesis for natural reinforced sound». *En Proc. of the 92nd Audio Eng. Soc.*, 1992.
- : «Acoustic control by wave field synthesis». *J. Acoust. Soc. Am.*, 1993, **93(5)**, pp. 2764–2778.
- BLADAPÉREZ, S.: «Contribuciones a la implementación de sistemas de Wavefield Synthesis». *Tesis Doctoral, Universidad Politécnica de Valencia*, 2009.
- BROWN, C.P. y DUDA, R.O.: «An efficient HRTF model for 3D sound». *Workshop on Applications of Signal Processing to Audio and Acoustics New York '97*, 1997.
- CETTA, P.: *Un modelo para la simulación del espacio en música*. Editorial de la Universidad Católica Argentina, Buenos Aires, Argentina, 2007.
- CHENG, C.I. y WAKEFIELD, G.H.: «Moving sound synthesis for binaural electroacoustics music using interpolated Head Related (HRTFs)». *Computer Music Journal*, 2001, **25(4)**, pp. 57–80.
- COOPER, D.H. y SHIGA, T.: «Discrete Matrix Multi-channel Stereo». *Journal of the Audio Engineering Society*, 1972, **20(5)**, pp. 346–360.
- CURA, M.: «Sistemas de sonido multicanal para la industria audiovisual». *En Música y espacio: ciencia, tecnología y estética, Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 203–226.
- DILISCIA, O.P.: «Los modos de vínculo de la concepción espacial del sonido con la poiesis de la música electroacústica». *Tesis Doctoral, Universidad Nacional de Rosario, Facultad de Humanidades y Arte.*, 2005.
- : «Técnicas de localización espacial de sonido con altoparlantes usando indicios de intensidad y tiempo». *En Música y espacio: ciencia, tecnología y estética, Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 99–126.
- FARINA, A.: «Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique». *AES convention 108*, 2000.
- FARINA, A. y AYALON, R.: «Recording concert hall acoustics for posterity». *24th International Conference: Multichannel Audio, The New Reality*, 2003.
- FELGETT, P.B.: «Directional Information in Reproduced Sound». *Wireless World*, 1972, **78**, pp. 413–417.
- GERSON, M.A.: «Periphony: With-height sound reproduction». *J. Audio Eng. Soc.*, 1973, **21**, pp. 2–10.
- : «Ambisonics: Part two: Studio techniques». *Studio Sound*, 1975, pp. 24–30.
- GIBSON J.J., CHRISTENSEN R.M. y A.L.R., LIMBERG: «Compatible FM Broadcasting of Panoramic Sound». *Journal of the Audio Engineering Society*, 1972, **20**, pp. 816–822.
- HUOPANIEMI, J.: «Virtual acoustics and 3D-sound in multimedia signal processing». *Tesis de doctorado, Universidad Tecnológica de Helsinki*, 1999.
- HUOPANIEMI, J. y RIEDERER, K.A.J.: «Measuring and modeling the effect of source distance in Head-Related Transfer Functions». *En Proceedings of the ICA/ASA '98*, 1998, pp. 20–26.

- KENDALL, G.: «La interpretación de la espacialización electroacústica: atributos espaciales y esquemas auditivos». *En Música y espacio: ciencia, tecnología y estética. Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 241–259.
- MALHAM, D.: «El espacio acústico tridimensional y su simulación por medio de Ambisonics». *En Música y espacio: ciencia, tecnología y estética, Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 161–202.
- MASON, R.: «Elicitation and measurement of auditory spatial attributes in reproduced sound». *Tesis de doctorado, Universidad de Surrey*, 2002.
- MASON R., BROOKES T. y F., RUMSEY: «Frequency dependency of the relationship between perceived auditory source width and the interaural cross-correlation coefficient for time-invariant stimuli». *J. Acoust. Soc. Am.*, 2005, **117(3)**, p. 1337.
- MCGOVERN, S.G.: «Fast image method for impulse response calculations of box-shaped rooms». *Applied Acoustics*, 2008, **70(1)**, pp. 182–189.
- MOORE, B.C.J.: *An introduction to the psychology of hearing*. Esmerald, Bingley, 2012.
- MOORE, F.R.: *Elements of Computer Music*. Prentice Hall, USA, 1990.
- PULKKI, V.: «Virtual Sound Source Positioning Using Vector Base Amplitude Panning». *J. Audio Eng. Soc.*, 1997, **45(6)**, pp. 456–466.
- : «Uniform spreading of amplitude panned virtual sources». *Applications of Signal Processing to Audio and Acoustics, 1999 IEEE Workshop on. IEEE*, 1999, pp. 187–190.
- : «Localization of amplitude-panned virtual Sources II: two and three-dimensional panning». *J. Audio Eng. Soc.*, 2001a, **49(9)**, pp. 753–767.
- : *Spatial sound generation and perception by amplitude panning techniques*. Helsinki University of Technology, Helsinki, 2001b.
- PULKKI, V. y M., KARJALAINEN: «Localization of amplitude-panned virtual Sources I: stereophonic panning». *J. Audio Eng. Soc.*, 2001, **49(9)**, pp. 739–752.
- SAVIOJA, L.: «Modeling techniques for virtual acoustics». *Tesis de doctorado, Universidad Tecnológica de Helsinki*, 1999.
- SUBÍAVALDEZ, J.R.: «Real Time Custom Spatialisation Patch for a large dimensión and non conventional area». *Proceeding of PDcon 2011, Berlín*, 2011.

---

SEGUNDA PARTE:  
TRABAJO DE CAMPO Y RESULTADOS



## - Capítulo 6 -

---

---

### El rol de la visión en la percepción auditiva de distancia

---

#### § 6.1 INTRODUCCIÓN

Con el fin de obtener una imagen coherente del mundo exterior, nuestro sistema nervioso debe procesar información proveniente de las diferentes modalidades sensoriales. El cerebro combina esta información multimodal para mejorar la detección, localización y discriminación de objetos y el tiempo de reacción. Al navegar en un ambiente, los humanos usan principalmente la información proveniente de las modalidades visual y auditiva. En general, estos sistemas trabajan juntos para facilitar la identificación y localización de eventos y objetos del mundo exterior.

Investigaciones previas señalan la influencia de la visión en la precisión de los juicios de localización auditiva (para una revisión, ver King, 2009). Por ejemplo, aunque el ángulo de incidencia de un sonido puede ser determinado solamente mediante pistas auditivas, la performance es más efectiva cuando se recibe, además, información visual (Jackson, 1953; Shelton y Searle, 1980; Stein et al, 1989). Un número menor de trabajos han reportado un efecto inverso (Simon y Craft, 1970; Perrott et al, 1990; Vroomen y de Gelder, 2000).

Estos autores mostraron que la información auditiva espacial puede también mejorar la respuesta a un objetivo visual, incluyendo la discriminación angular, reducir los tiempos de reacción y mejorar la localización.

A menudo, las modalidades auditiva y visual aportan información conflictiva, y el estudio de como esta discrepancia es resuelta ha atraído mucha atención en los últimos años (Kubovy y Van Valkenburg, 2001; Alais y Burr, 2004; King, 2009; Evans y Treisman, 2010). Un ejemplo conocido es el del “efecto ventrílocuo”, en el cual la localización de un estímulo visual produce un sesgo en el ángulo de localización de una fuente sonora (Radeau y Bertelson, 1974; Recanzone, 1998; Lewald, 2002). Mientras algunos investigadores han teorizado que este sesgo proviene de una captura completa de la señal auditiva por parte de la señal visual (Pick et al, 1969; Bertelson y Radeau, 1981; Warren et al, 1981), Alais y Burr (2004) han demostrado que este efecto puede ser explicado mediante un simple modelo de combinación óptima de las pistas espaciales, visuales y auditivas, donde cada modalidad es ponderada por la estimación inversa de su variabilidad. Debido a que nuestra habilidad para localizar estímulos es típicamente mejor con pistas visuales que nuestra habilidad para hacerlo mediante la información auditiva, cuando se produce un conflicto entre estas modalidades, la información visual tiende a sesgar la percepción del estímulo auditivo. Sin embargo, si el estímulo visual es confuso, o sea, difícil de localizar, la información visual puede ser más débil que la auditiva y la ilusión se revierte y la ubicación el objeto visual es desviada hacia el lugar donde se encuentra la fuente que emite el sonido (Alais y Burr, 2004). Si bien existen muchos estudios sobre localización multisensorial, la mayoría de ellos están dedicados a la localización en acimut y muy pocos se centran en el

efecto conjunto de la información visual y auditiva en la percepción auditiva de distancia (PAD). El objetivo de esta parte de la presente tesis es estudiar si la precisión en la PAD mejora en presencia de información visual.

## § 6.2 PERCEPCIÓN DE DISTANCIA AUDITIVA Y VISUAL

Los seres humanos, tanto para la modalidad visual como para la auditiva, utilizan un amplio rango de pistas para estimar la distancia de un objeto. Las escenas visuales contienen potencialmente muchas fuentes diferentes de información de profundidad (tanto monocular como binocular). Las fuentes de información de profundidad monoculares pueden clasificarse como pistas pictóricas (estáticas o dinámicas). Una imagen bidimensional puede proveer pistas de profundidad estáticas como el tamaño relativo, la interposición de los objetos, la declinación angular, distribución de luz y sombra, perspectiva, y muchas otras (Bülthoff et al, 1998; Loomis, 2001). Las pistas basadas en el movimiento son inducidas por los movimientos relativos entre el observador y los objetos (Ono et al, 1988) e incluyen paralaje móvil, efectos de profundidad kinéticos y oclusión dinámica. La visión estéreo permite percibir la distancia de un objeto en el campo cercano (Foley, 1980). La pista binocular más conocida es la llamada disparidad binocular, que se produce por una discrepancia entre las imágenes que llegan al ojo derecho e izquierdo cuando se observa un objeto (Qian, 1997; Blake y Sekuler, 2006; Sousa et al, 2010). Esta disparidad permite al cerebro estimar la profundidad relativa de los objetos con respecto al punto de fijación, un proceso conocido como estereopsis. La convergencia es otra pista binocular, que no es ni más ni menos que el giro que se produce



hacia el interior de los ojos para mantener la visión estereoscópica cuando un objeto se acerca, y que sirve como pista para el sujeto (Semmlow y Heerema, 1979).

Como vimos en el Capítulo 3, localización en distancia de fuentes sonoras, en la PAD se han reportado y estudiado ciertas pistas, las cuales se consideran de las más importantes. La intensidad es una de las pistas primarias para PAD (Coleman, 1962), debido a que esta decrece cuando la distancia a la fuente aumenta (6dB por el doble de la distancia en campo libre). Otra importante pista de la distancia de una fuente sonora es la relación de energía del sonido directo y el reverberante (D/R – Bekesy, 1938; Mershon y King, 1975; Mershon y Bowers, 1979; Butler et al, 1980; Nielsen, 1993; Bronkhorst y Houtgast, 1999; Zahorik, 2002).

En ambientes reverberantes, la intensidad del sonido directo de una fuente decrece 6dB cada vez que la distancia se duplica, mientras que la energía de las reflexiones que arriban más tarde se mantiene relativamente constante. La D/R es una pista muy interesante ya que no depende de la intensidad de la señal y puede ser efectiva en cualquier ambiente en donde haya reflexiones (hasta en un ambiente al aire libre tenemos una reflexión proveniente del piso). Otras pistas involucradas en la PAD son: el contenido espectral, la familiaridad con la fuente sonora, las pistas binaurales y dinámicas (ver Zahorik et al, 2005 para una revisión). En general se indica que la distancia a un objeto es percibida con mayor precisión y menor variabilidad en la modalidad visual, que en la modalidad auditiva (Loomis et al, 1998). De hecho, muchos estudios han reportado que aun en presencia de múltiples pistas auditivas, la PAD suele ser pobre (Coleman, 1962; Middlebrooks y Green, 1991; Wenzel et

al, 1993; Zahorik, 2002).

Investigaciones previas indican que, para la modalidad auditiva, la distancia a una fuente sonora es sobrestimada cuando la fuente se localiza a menos de 2m, mientras que es substancial y progresivamente subestimada para distancias mayores a estas (Zahorik et al, 2005). En muchos casos, la función que relaciona la percepción de los sujetos con la distancia física a la fuente se ajusta bien a una función de potencia con un exponente considerablemente menor a uno, correspondiente a una no linealidad compresiva entre la distancia física y la percibida. Otra característica común en la medición de la PAD es la alta variabilidad de la respuesta, tanto entre sujetos como intra sujetos, o sea la variabilidad entre distintas respuestas de un sujeto (Loomis et al, 1998; Zahorik 2002; Zahorik et al, 2005).

En contraste, ha sido reportado que la Percepción Visual de Distancia (PVD) es bastante exacta para objetivos visuales que se encuentran hasta los 20 metros de distancia, si múltiples pistas visuales están disponibles para el observador (Thomson, 1983; Elliott, 1986; Loomis et al, 1992; Fukusima et al, 1997). Sorpresivamente, cuando el objeto visual es presentado en la oscuridad (con escasas pistas visuales), los resultados de PVD son muy similares a los obtenidos en experimentos de PAD: la distancia es sobrestimada cuando el objetivo se encuentra antes de los dos metros y es subestimada cuando la distancia es más grande que 3 metros. Además, bajo estas condiciones la variabilidad de la respuesta se incrementa significativamente (Loomis et al, 1998). Una cuestión importante que se quiere tratar aquí es si la precisión en la PAD podría ser mejorada mediante la presencia de claves visuales, como sucede en las tareas de PVD (Philbeck y Loomis, 1997).

### § 6.3 EFECTO DE LA INFORMACIÓN VISUAL EN LA PAD

Muchos estudios multisensoriales que integran la modalidad visual y auditiva se han enfocado en la localización de sonidos en el eje horizontal. Incluso cuando la localización en acimut y en distancia se basa en pistas visuales y auditivas diferentes (con diferentes tipos de variabilidad) en PAD han sido reportados sesgos similares. Un trabajo pionero en el estudio del rol de la visión en PAD fue publicado por Gardner (1968). Su dispositivo experimental consistía de cinco parlantes alineados, ubicados a diferentes distancias y apuntando al sujeto experimental, dentro de una cámara anecoica. Los parlantes fueron ubicados al nivel de los ojos de los sujetos, de tal manera que estos solo podían observar el primer parlante durante el experimento. Los resultados mostraron que aunque el estímulo solo fue reproducido por el parlante más lejano, todos los participantes, sin excepción, reportaron que el sonido provenía del parlante más cercano (el único que los sujetos podían ver). Gardner llamó a este efecto como “efecto de proximidad de la imagen” (del inglés, *proximity image effect*), y demostró que este opera sobre una gran variedad de intensidades, y que no depende de la distancia del parlante más lejano (Gardner, 1968).

En 1980, Mershon et al estudiaron el mismo fenómeno pero bajo condiciones donde el entorno era semi reverberante. Ellos concluyeron que el efecto de proximidad de la imagen opera similarmente tanto en entornos anecoicos como en reverberantes. Aún más, reportaron que la distancia a la fuente sonora puede ser sobrestimada o subestimada, dependiendo de la posición del objetivo visual (en este caso un parlante falso – Mershon et al, 1980).

Con el fin de reexaminar este efecto, Zahorik (2001) realizó experimentos utilizando un dispositivo similar al de Gardner (1968), pero en un entorno semi reverberante. Bajo estas condiciones los oyentes poseían pistas auditivas adicionales para la localización de las fuentes sonoras basadas en las reflexiones del entorno. Para este experimento, a la mitad de los oyentes se les tapo los ojos antes de entrar a la sala donde se realizó la tarea (la cual no conocían previamente), y se mantuvieron así durante todo el experimento. La otra mitad de los participantes entraron a la sala con sus ojos destapados, y se mantuvieron así durante el experimento. De esta manera, el segundo grupo de oyentes podían ver tanto el ambiente en el que se encontraban como el dispositivo de cinco parlantes. Sin embargo, durante la tarea se mantuvieron sentados delante de la fila de parlantes, de tal manera que solo podían observar el parlante más cercano, como sucedía en los experimentos de Gardner (1968). Sus resultados mostraron que la presencia de múltiples pistas visuales (condición con los ojos destapados) incrementaba la precisión de la PAD y disminuía la variabilidad de los juicios, comparados con los resultados obtenidos bajo las mismas condiciones pero sin pistas visuales (condición donde los sujetos tenían los ojos tapados). Zahorik concluye que “este experimento provee evidencia que sugiere que el efecto de captura visual en distancia no es tan general como proponen los estudios pasados. Bajo las condiciones del presente experimento, el cual es rico en pistas auditivas de PAD, el efecto no fue observado”. Agrega además, que la presencia de múltiples pistas visuales disponibles durante el experimento, incrementa la precisión de los oyentes de percibir la distancia de una fuente sonora (Zahorik, 2001). Sin embargo, como en el experimento original realizado por Gardner, en el experimento de Zahorik los participantes

podían solamente observar el primer parlante del dispositivo durante la tarea perceptual. Por lo tanto, no se puede excluir que el efecto de proximidad de la imagen haya influenciado en la respuesta de los oyentes. De hecho, aunque la respuesta obtenida por Zahorik en la condición con múltiples pistas visuales es más precisa que en la oscuridad, para distancias mayores a 2 metros los oyentes subestimaron significativamente la distancia de la fuente sonora (para distancias a la fuente de 4 y 5 metros la respuesta nunca sobrepasó los 3 metros). En la PAD, tanto el sesgo como la variabilidad en la respuesta crecen cuando crece la distancia a la fuente sonora (Zahorik et al, 2005). Por lo tanto, se espera que la imagen del primer parlante pueda haber influido en la respuesta, más aun cuando esta se localiza a una gran distancia, mejorando las condiciones experimentales que tratan de prevenir los artefactos perceptuales potenciales causados por el efecto de proximidad de imagen se puede remover la subestimación de las respuestas de los oyentes cuando la fuente de sonido se ubica a distancias mayores de 2 metros.

También vale la pena señalar que Zahorik usa un dispositivo con cinco parlantes en línea, un arreglo que puede proveer pistas auditivas extras debido al filtrado que puede producir la sombra acústica de cada parlante (por ejemplo, la atenuación en el contenido de altas frecuencias de los parlantes que son tapados por el primer parlante). La meta de este experimento es estudiar si la PAD puede ser mejorada mediante la presencia de pistas visuales que otorguen información de distancia. Para este propósito, se realizaron experimentos en un entorno real y semi reverberante, en la oscuridad (en presencia y ausencia de pistas visuales mínimas), usando un parlante móvil que permite reproducir estímulos sonoros en varias distancias. De esta manera, las pistas acústicas

potenciales producidas por el filtrado de la señal mencionado, encontrado en los dispositivos utilizados por Zahorik y Gardner, son removidos.

## § 6.4 MÉTODOS GENERALES

La meta de este trabajo es estudiar si la precisión en la PAD puede ser mejorada mediante la presencia de información visual. Para esto, se realizaron tres experimentos diferentes (ver Figura 6.1). Primero, se realizó un experimento de PVD que actuó como control del dispositivo, y que fue tomado como base para el uso de pistas visuales de referencia en los experimentos subsecuentes. Luego, fueron realizados experimentos de PAD con dos grupos de sujetos, en la oscuridad, tanto en presencia como en ausencia de pistas visuales mínimas. En el tercer y último experimento, se estudia como el conocimiento a priori de un ambiente específico afecta a la PAD.

Experimentos	Condición
1: Percepción de Distancia Visual Sin conocimiento previo de la sala	i (sin pistas visuales) ii (dos pistas visuales) iii (cuatro pistas visuales)
2A: Grupo A: Percepción Auditiva de Distancia Sin conocimiento previo de la sala	i (sin pistas visuales) ii (dos pistas visuales) iii (cuatro pistas visuales)
2B: Grupo B: Percepción Auditiva de Distancia Sin conocimiento previo de la sala	iii (cuatro pistas visuales) i (sin pistas visuales)
4: Percepción Auditiva de Distancia Con conocimiento previo de la sala	i (sin pistas visuales)

**Figura 6.1:** Resumen de todos los experimentos y condiciones, listados en el orden en que fueron realizados.

## § 6.5 ENTORNO EXPERIMENTAL

Todos los experimentos fueron realizados en un entorno semireverberante de 12x7x3 metros (largo x ancho x alto) cuyas paredes están recubiertas de paneles absorbentes (espuma de poliuretano piramidal acústica de 50mm) el piso alfombrado y el techo es de paneles acústicos de fibra de vidrio. El tiempo de reverberación promedio de la sala es de 0.49 segundos en el lugar donde se ubica el sujeto (T30 a 1kHz medido mediante el método MLS). El ruido de fondo de la sala, medido en el lugar donde se ubica el sujeto, sin ningún tipo de estímulo sonando es de 19dBa (medido con un sonómetro RION NL-32).

## § 6.6 PARTICIPANTES

Un total de 32 voluntarios (24 hombres y 8 mujeres) participaron en los experimentos. Los voluntarios eran estudiantes de grado y graduados de la Licenciatura en Composición con Medios Electroacústicos. El rango de las edades fue de los 21 a los 42 años (edad media, 27,68 años). A pesar de que no se realizaron medidas de agudeza visual y sensibilidad auditiva, todos los participantes informaron una visión normal (o corregida) y una audición normal. Cuatro de los participantes utilizaron lentes correctivos. Ninguno conocía la sala a priori ni se les indico las dimensiones de la misma (salvo en el experimento 3). Todos los participantes hicieron el experimento de forma consciente y firmaron una forma por escrito indicando esto.

Antes de entrar a la sala donde se realizó el experimento, cada participante fue instruido sobre la tarea que debía realizar. Las instrucciones estaban di-

señadas de tal manera que enfatizaban que los reportes debían estar basados en la distancia aparente de la fuente, en vez de tratar de estimar de forma objetiva la distancia exacta de la misma (ver Carlson, 1977). Una vez instruidos, a los participantes se les tapo los ojos y se los condujo dentro de la sala, donde se los sentó en una silla posicionada en el punto cero de inicio de la tarea. Durante el experimento, la luz de la sala se mantuvo apagada, y los sujetos tenían los ojos destapados.

### § 6.7 PISTAS VISUALES

Las pistas visuales de referencia fueron montadas en postes de metal de 1.2 metros de alto, y localizadas a 2, 4, 6 y 8 metros del participante, formando una línea paralela (20cm a la derecha) al riel por donde se mueve el parlante (ver la Figura 6.2C). Estas pistas visuales consistían en un par de LEDs rojos (estándar de 3mm, 18mca) orientados verticalmente y separados 4 cm (figura Figura 6.2D). Antes de entrar en la sala, todos los participantes fueron informados de las condiciones en que harían el experimento.

Fueron utilizadas tres diferentes condiciones de pistas visuales de referencia:

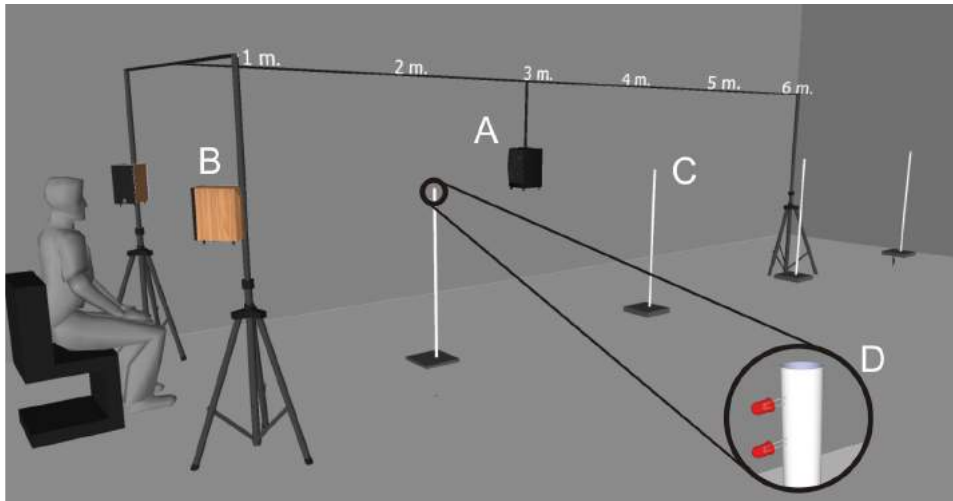
Condición i: Sin pistas visuales;

Condición ii: Con dos pistas visuales localizadas a 4 y 8 metros;

Condición iii: Con 4 pistas visuales localizadas a 2, 4, 6 y 8 metros.

Para las condiciones ii y iii, los LEDs se mantuvieron encendidos durante todo el experimento. En todas las condiciones descritas los participantes tuvieron





**Figura 6.2:** Modelo en tres dimensiones del dispositivo experimental. (a) Parlante móvil; (b) sistema de enmascaramiento; (c) pistas visuales usadas en el experimento; (d) las pistas visuales, formadas por pares de LEDs rojos localizados verticalmente y separados 4cm.

sus ojos abiertos y la posición de la cabeza no fue fijada. Mediante el uso de este diseño experimental, se intentó proveer una variedad de pistas visuales de distancia monoculares (paralaje de movimiento y acomodación) y binoculares (disparidad binocular y convergencia).

### § 6.8 ESTÍMULOS VISUALES

Los estímulos visuales empleados en el experimento 1 consistieron de un par de LEDs verdes (3mm, 18mca) localizados verticalmente, separados 4cm entre sí y ubicados a la altura de los ojos de los sujetos. Estas pistas estuvieron ubicadas a seis distancias (D) diferentes:  $D = 1, 2, 3, 4, 5$  y 6 metros.

### § 6.9 ESTÍMULOS AUDITIVOS

Los estímulos auditivos empleados en los experimentos 2 y 3 consistieron de ráfagas de ruido blanco de 500ms de duración (ancho de banda medido entre 50HZ y 20Khz  $\pm 2$  Hz) con una envolvente de amplitud al inicio y al final de

aproximadamente 50ms. Las señales fueron generadas en Matlab y reproducidas mediante una placa de sonido a una frecuencia de muestreo de 44.1Khz con una resolución de 24 bits. El nivel de presión sonora de los estímulos fue de 70dBA, medido en la posición donde el oyente realizaría la tarea y con la fuente sonora ubicada a 1m (sonómetro RION NL-32). Por esto, la intensidad fue una importante pista durante el experimento. Entre prueba y prueba, se presentaba un sonido enmascarante en los parlantes ubicados a los lados del participante (figura 1b). El contenido espectral, la duración y la intensidad de este sonido (12 segundos, 70dBA medidos en la ubicación del participante) fueron ajustados de manera que enmascarara completamente el sonido producido por el desplazamiento del parlante de prueba sobre el riel. Dos segundos después de terminar este sonido enmascarante, se presentaba el estímulo auditivo a través del parlante de prueba.

El dispositivo experimental fue diseñado y construido en el LAPSo (Laboratorio de Acústica y Percepción Sonora de la Universidad Nacional de Quilmes) y consistió de un parlante móvil (figura 1A, Genelec 8020B bi-amplificado 50W) localizado en el frente de los participantes, a 1.2 metros del suelo (aproximadamente la altura a la que se encontraban los oídos de los participantes sentados) que puede moverse libremente a lo largo de un riel de metal de 6 metros de largo. Este sistema permitió reproducir estímulos sonoros a diferentes distancias del participante. El dispositivo posee además un sistema de enmascaramiento de dos parlantes fijos (Edifier R1000TCN 25W) localizados a ambos lados del participante y apuntando a su oídos (figura 1b). Tanto el parlante móvil como los parlantes de enmascaramiento fueron controlados por una placa de sonido estéreo (Presonus AUDIOBOX- 2 out 2 in). El parlante

móvil se movía en línea recta al sujeto experimental y por el centro y en forma paralela a dos paredes de la sala.

Bajo estas condiciones los oyentes contaban al menos con dos importantes pistas de PAD: la intensidad y la relación de energía entre el sonido directo y el sonido reverberante (D/R). Debido a las distancias utilizadas en este experimento (de 1 a 6 metros) las pistas binaurales y espectrales de PAD no estaban disponibles de forma significativa (ver Zahorik et al 2005).

### § 6.10 PROCEDIMIENTO GENERAL

Las distancias a las que se evaluó la percepción de distancia de los estímulos visuales y auditivos fueron: D= 1, 2, 3, 4, 5 y 6 metros. Para todos los experimentos el procedimiento consistió en la presentación del estímulo a una de las seis distancias pidiendo a los participantes que realizaran un juicio indicando la distancia egocéntrica aparente del objetivo visual (LEDs) o el auditivo (parlante). Los juicios de distancia fueron realizados en forma verbal, usando una escala en metros. Los estímulos fueron presentados tres veces para cada una de las 6 distancias formando un total de 18 juicios por bloque. La presentación de los estímulos fue aleatoria. Solamente una respuesta verbal fue realizada en cada presentación, y el participante no recibió información sobre lo correcto o incorrecto de su respuesta.

### § 6.11 EXPERIMENTO 1

El propósito de este experimento fue testear la idoneidad del sistema de pistas visuales que fue utilizado para los demás experimentos. Para esto, se realizó un

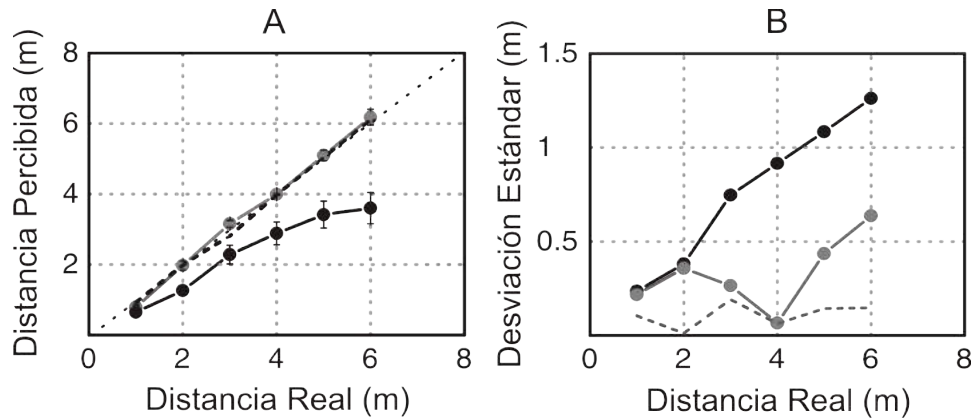
experimento de percepción visual de distancia (PVD) en presencia y ausencia de pistas visuales. Las diferentes condiciones de pistas visuales fueron utilizadas en experimentos separados, con descansos de 10 minutos entre cada condición, en el siguiente orden: condición i, condición ii, condición iii (ver Figura 6.1). Ninguno de los participantes (cinco hombres y tres mujeres) utilizó lentes correctivos durante el experimento.

### 6.11.1. Resultados

En la Figura 6.3A se puede observar la media de los juicios de distancias realizados por los participantes como una función de la distancia real sin pistas visuales (condición i). La barra de error representa 1 SEM (del inglés, Standart Error of Mean) ( $n=8$ ). Una función de potencia de la forma  $Y = aX^b$  fue ajustada a los datos obtenidos ( $R^2 = 0,92$ ), usando un método de mínimos cuadrados ( $a = 0.81 \pm 0.12$  y  $b = 0.87 \pm 0.1$ ). Como puede observarse claramente en esta figura, en ausencia de pistas visuales la distancia a la fuente real fue subestimada para distancias mayores a 1 metro. Este resultado concuerda con lo reportado en estudios anteriores, donde se ha reportado que la restricción de las pistas visuales de distancia producidas por presentar el objetivo a evaluar en la oscuridad produce errores perceptuales.

Sin embargo, este resultado no concuerda con estudios previos, los cuales reportan una sobrestimación de la respuesta para distancias menores a 3 metros (Gogel, 1961; Philbeck y Loomis, 1997).

En contraste con lo expresado arriba, cuando el experimento fue realizado en presencia de dos pistas visuales localizadas a 2 y 4 metros del participante (condición II, Figura 6.3A) las respuestas fueron muy precisas ( $r = 0.98$ ;  $a =$



**Figura 6.3:** Experimento 1: Percepción Visual de Distancia (sin conocimiento previo de la sala). (A) Valor medio de los juicios de distancia como una función de la distancia real obtenida sin pistas visuales (condición i,  $n=8m$ , círculos negros), con dos pistas visuales localizadas a 4 y 8 metros del participante (condición ii,  $n=8$ , círculos grises), y cuatro pistas visuales localizadas a 2, 4, 6 y 8 metros del participante (condición iii,  $n=8$ , línea negra discontinua). (B) Desviación estándar de los juicios de distancia como una función de la distancia de la fuente para: condición i (círculos negros), condición ii (círculos grises) y condición iii (línea negra discontinua).

$0.95 \pm 0.06$ ; exponente  $b = 1.04 \pm 0.03$ ). Esto se refleja como un incremento significativo en el valor del exponente  $b$  de la función de potencia ajustada a los resultados, en comparación a los datos obtenidos bajo la condición i (2-sample t-tests;  $t = 2.37$ ;  $p = 0.032$ ). La respuesta obtenida en presencia de cuatro pistas visuales (figura 2a) localizadas a 2, 4, 6 y 8 metros del participante (condición iii) fue muy similar a la obtenida bajo la condición ii ( $r = 0.99$ ;  $a = 0.91 \pm 0.24$  y un exponente  $b = 1.06 \pm 0.08$ ).

La respuesta bajo la condición iii fue significativamente más lineal que la obtenida en la condición sin pistas visuales (condición i). Esto se refleja como un incremento significativo en el valor del exponente  $b$  de la función de potencia ajustada a los datos ( $t = 3.22$ ;  $p = 0.01$ ). Fueron calculados análisis de varianza (one-way ANOVA) para cada una de las distancias (D) con cada una de las condiciones como factor. El promedio de la distancia aparente fue significativamente diferente a través de las condiciones para todas las distancias

( $F_{3,21} > 4, p < 0,03$ ). Un análisis posterior (test Tukey-Kramer) mostró que las distancias estimadas bajo las condiciones ii y iii fueron significativamente más grandes que las estimadas bajo la condición i para todas las distancias testeadas, con excepción de  $D = 1$  metro, donde solamente las estimaciones para la condición iii fueron significativamente más grande que las de las condición i.

La Figura 6.3B muestra la variabilidad de los datos para las condiciones i, ii y iii. Se puede observar que cuando el experimento se realizó sin pistas visuales la variabilidad de la respuesta fue significativamente mayor que la obtenida con pistas visuales (excepto cuando se compara con la condición ii, donde esto ocurre solamente para distancias mayores a 2 metros). Este último resultado es consistente con estudios pasados (Philbeck y Loomis, 1997), donde se obtuvo menor variabilidad en la respuesta bajo condiciones con múltiples pistas visuales que la obtenida en ausencia de las mismas.

Además, bajo la condición i (sin pistas visuales) la variabilidad de la respuesta se incrementa cuando la distancia al objetivo aumenta, mientras que en presencia de pistas visuales la variabilidad de la respuesta se mantiene relativamente constante para todas las distancias testeadas. Además, de acuerdo con investigaciones previas, la restricción de pistas visuales causa errores perceptuales, mientras la respuesta fue muy precisa cuando las pistas visuales están disponibles. Sin embargo, es importante mencionar que en este experimento las pistas visuales fueron efectivas solamente cuando se indicó a los participantes la distancia real de las pistas visuales de referencia (par de LEDs). Cuando las mismas pistas fueron utilizadas sin proveer esta información a los participantes, la respuesta observada fue muy similar a la obtenida sin pistas

visuales (datos no mostrados). Sin embargo, cuando los participantes fueron informados sobre la distancia de las pistas visuales, estos realizaron juicios muy precisos de la distancia del objetivo relativo a estas referencias. Esto es consistente con la hipótesis de que si la distancia a un objeto (referencia) es conocida, todas las otras distancias pueden ser derivadas de las disparidades relativas con respecto a estas referencias (Brenner y van Damme, 1999; Sousa et al, 2010).

Los resultados obtenidos en presencia de pistas visuales (condición ii y iii) fueron similares a los reportados en estudios previos bajo condiciones donde se dispone de abundantes pistas visuales, sugiriendo que el sistema de pistas visuales utilizadas en este experimento puede ser usado como pista visual de referencia efectiva para los siguientes experimento de PAD.

## § 6.12 EXPERIMENTO 2

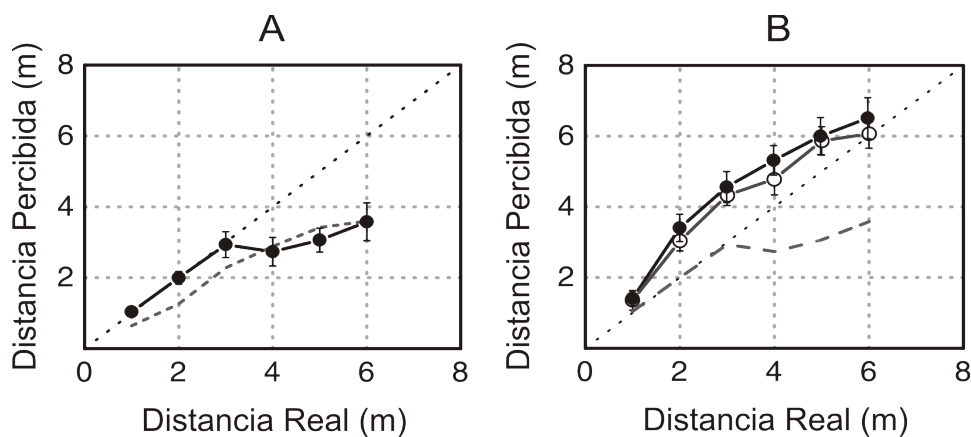
Este experimento fue realizado para examinar si la presencia de pistas visuales puede mejorar la PAD. Un total de 16 voluntarios (13 hombres y 3 mujeres) participaron de este experimento. Dos de los participantes utilizaron lentes correctivos.

Los participantes fueron divididos en dos grupos. Para el primer grupo de participantes (grupo A,  $n = 8$ ) las pistas visuales fueron usadas en la misma secuencia que en el experimento 1: condición i (sin pistas visuales), condición ii (dos pistas visuales localizadas a 4 y 8 metros) y condición iii (cuatro pistas ubicadas a 2, 4, 6 y 8 metros). Para el segundo grupo de participantes (grupo B,  $n = 8$ ), se testeó la condición iii primero y luego la condición i. De la misma

forma que en el experimento 1, todos los participantes fueron informados sobre la distancia física de las pistas visuales de referencia antes de entrar en la sala donde se realizó el experimento.

### 6.12.1. Resultados

En la Figura 6.4A se muestra la media de los juicios subjetivos de distancia obtenidos en la condición sin pistas visuales como una función de la distancia física (grupo A,  $n = 8$ , condición i; ver sección 2). Cada punto corresponde a la media geométrica de las tres presentaciones y de todos los participantes. Una función de potencia de la forma  $Y = aX^b$  fue utilizada para ajustar los datos utilizando un método de cuadrados mínimos ( $R^2 = 0.78$ ;  $a = 1.33 \pm 0.2$  y el exponente  $b = 0.55 \pm 0.1$ ). Estos valores son muy similares a los obtenidos en un meta análisis de resultados de 21 estudios previos de PAD ( $a = 1.32$  con un exponente  $b = 0.54$ ; Zahorik et al, 2005).



**Figura 6.4:** Experimento 2 – grupo A: Percepción Auditiva de Distancia (sin conocimiento previo de la sala). (A) Valor medio de los juicios de distancia como una función de la distancia real obtenida sin pistas visuales (condición 1,  $n=8$ , círculos negros). Como referencia se agregó, mediante una línea gris discontinua, los resultados obtenidos en el experimento 1 (condición i para visión). (B) Valor medio de los juicios de distancia como una función de la distancia real obtenida con dos (condición ii,  $n=8$ , círculos blancos) y cuatro pistas visuales (condición i, para percepción auditiva de distancia).



Vale la pena señalar que en este experimento, para fuentes sonoras más lejanas que 3 metros, las distancias percibidas fueron substancial y progresivamente subestimadas. Sin embargo, la respuesta fue muy precisa para las tres distancias más cercanas al participante. Con el fin de comparar la respuesta entre las modalidades visual y auditiva bajo condiciones similares, en la Figura 6.4 se muestran los resultados obtenidos en el experimento 1 sin pistas visuales (línea gris de trazos cortos, condición i, del experimento de PVD). Esta comparación no presentó diferencias significativas entre ambos grupos (condición i para experimentos 1 y 2) para ninguna las distancias (one-way ANOVA F1,  $114 < 4, p > 0,05$ ), con excepción en los dos metros ( $F1, 14 = 10, p = 0,007$ ). En la Figura 6.4B se muestra la media de los juicios subjetivos de distancia como una función de la distancia real de la fuente, obtenida bajo la condición ii (dos pistas visuales localizadas a 4 y 8 metros) para el mismo grupo de participantes ( $n = 8$ , grupo A, ver Figura 6.1). Una función de potencia de la forma  $Y = aX^b$  fue ajustada a estos datos usando el método de cuadrados mínimos ( $r = 0.91$ ;  $a = 1.84 \pm 0.22$  con un exponente  $b = 0.71 \pm 0.08$ ). La respuesta fue significativamente menos compresiva que la respuesta obtenida sin pistas visuales. Esto se refleja como un incremento significativo en el valor del exponente b de la función de potencia (2-sample t -test;  $t = 2.61$ ;  $p = 0.02$ ). En la misma figura, los círculos negros corresponden a la media de los juicios subjetivos de distancia como una función de la distancia real de la fuente sonora bajo la condición iii (cuatro pistas visuales ubicadas a 2, 4, 6 y 8 metros) para el mismo grupo de participantes ( $r = 0.91$ ;  $a = 2.01 \pm 0.24$  con un exponente  $b = 0.68 \pm 0.08$ ). La respuesta fue muy similar a la obtenida bajo la condición ii. Los valores de ajuste de la función de potencia

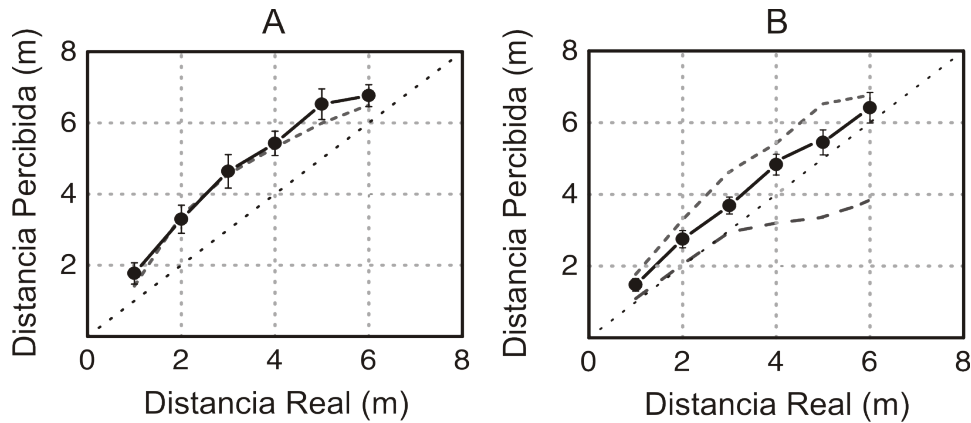
para las dos condiciones no fueron significativamente diferentes ( $t = 0.48$ ;  $p = 0.63$  para a y  $t = 0.43$ ;  $p = 0.66$  para el exponente b). Como en el caso previo, la respuesta fue más extendida que la obtenida sin pistas visuales lo que se ve reflejado como un significativo incremento en el valor del exponente b, comparado con el obtenido bajo la condición i ( $t = 2.45$ ;  $p = 0.028$ ).

Nuestros resultados muestran que en presencia de pistas visuales, la distancia a la fuente sonora es sobrestimada. Un análisis de varianza fue realizado para cada distancia (D) con la condición como factor. La distancia aparente promedio fue significativamente diferente a través de todas las distancias ( $F_{2, 21} > 6, p < 0,008$ ), con la excepción de  $D = 1$  metro. Un análisis a posterior (para  $D > 1$  metro) muestra que la distancia estimada bajo las condiciones ii y iii fueron significativamente más grandes que las estimadas para la condición i en todos los casos.

En la Figura 6.5 se muestran los resultados obtenidos para el grupo B. La Figura 6.5A ilustra la media de los juicios subjetivos de distancia como una función de la distancia a la fuente sonora obtenidos bajo la condición iii (cuatro pistas visuales) para los ocho participantes del grupo B, los cuales no habían participado en experimentos de ADP ( $r = 0.98$ ,  $a = 2.01 \pm 0.16$  con un exponente  $b = 0.68 \pm 0.05$ ). Como referencia, en la misma figura se observa, en una línea gris de trazos cortos, la media de los juicios subjetivos de distancia bajo la condición iii para el grupo A. Este experimento fue realizado para excluir la posibilidad de que los resultados obtenidos en presencia de pistas visuales para el grupo A no se hayan producido por un efecto de aprendizaje (dado que las tareas fueron realizadas en forma secuencial). Bajo estas condiciones la respuesta de los participantes no fue significativamente

diferente que la obtenida para el grupo A en las mismas condiciones ( $t = 0.48$ ;  $p = 0.63$  para a, y  $t = 0.43$ ;  $p = 0.66$  para el exponente b). Además, un análisis de varianza para cada condición por separado no muestra diferencias significativas a través de los grupos ( $F(1, 14) < 1, p > 0.3$ ). Es importante notar que la respuesta fue significativamente menos compresiva que la respuesta sin pistas visuales (grupo A bajo la condición i;  $t = 2.41, p = 0.031$ ).

En la Figura 6.5B (círculos negros) se muestran los resultados obtenidos sin pistas visuales (condición i) para los participantes que previamente habían participado en el experimento con cuatro pistas visuales (grupo B). Sorpresivamente, la respuesta fue más precisa que la obtenida por el grupo A bajo la misma condición ( $r = 0.95$ ;  $a = 1.57 \pm 0.16$  con el exponente  $b = 0.79 \pm 0.05$ ). Esto se ve reflejado por el significativo incremento en el valor del exponente b ( $t = 1.93$ ;  $p = 0.0005$ ). Como referencia, en la figura 4b se muestra, en línea de trazos gris, la media de los juicios subjetivos de distancia para la condición i para el grupo A. Más aun, la respuesta en ausencia de pistas visuales para el grupo B no difiere significativamente de los resultados obtenidos previamente en presencia de pistas visuales (condición iii) en ninguna de las distancias testeadas (ANOVA  $F(1, 14) > 4, p > 0.06$ ). Además, la respuesta bajo esta condición fue ligeramente menos compresiva que la obtenida en presencia de cuatro pistas visuales (condición iii) para el mismo grupo de participantes (grupo B), aunque esta diferencia no es estadísticamente significativa ( $t = 1.07$ ;  $p = 0.3$ ).



**Figura 6.5:** Experimento 2 – grupo B: Percepción Auditiva de Distancia (sin conocimiento previo de la sala). (A) Valor medio de los juicios de distancia como una función de la distancia real obtenida con cuatro pistas visuales (condición iii,  $n=8$ , círculos negros). Como referencia se agregó, mediante una línea gris clara discontinua, los resultados obtenidos para el grupo bajo las mismas condiciones. (B) Valor medio de los juicios de distancia como una función de la distancia real obtenida sin pistas visuales (condición i,  $n=8$ , círculos negros). La línea gris clara discontinua corresponde a la respuesta obtenida con cuatro pistas visuales (condición iii) por los participantes del grupo A. Como referencia se agregó, mediante una línea gris oscura discontinua, la respuesta obtenida sin pistas visuales (condición i) por los participantes del grupo A.

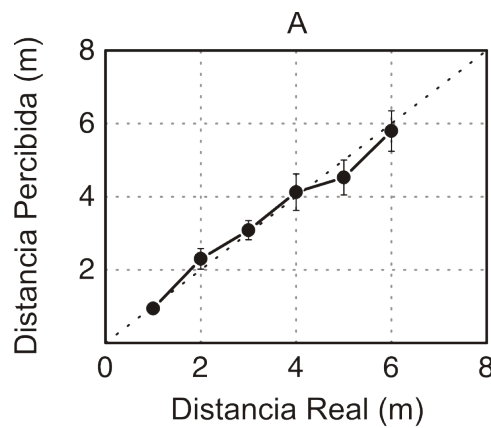
### § 6.13 EXPERIMENTO 3

El propósito de este experimento fue evaluar si el conocimiento previo de un ambiente específico afecta la PAD. A los participantes (7 hombres y 1 mujer) se les permitió conocer la sala donde se realizaría el experimento, con las luces prendidas, durante 5 minutos, antes de realizar la tarea. Sin embargo, estos no fueron informados sobre las dimensiones de la sala. Luego de esta inspección visual los participantes fueron llevados a una sala adyacente donde fueron instruidos sobre lo que debían realizar durante el experimento. Dos de los participantes utilizaron lentes correctivos.

Este experimento fue realizado con los mismos procedimientos que el experimento anterior, pero utilizando solamente la condición i (sin pistas visuales).

### 6.13.1. Resultados

En la Figura 6.6 se observa la media de los juicios subjetivos de distancia obtenidos mediante reportes verbales sin pistas visuales (condición i) como una función de la distancia real de la fuente sonora. Sorpresivamente, la respuesta obtenida bajo estas condiciones, fue muy precisa en todas las distancias testeadas.

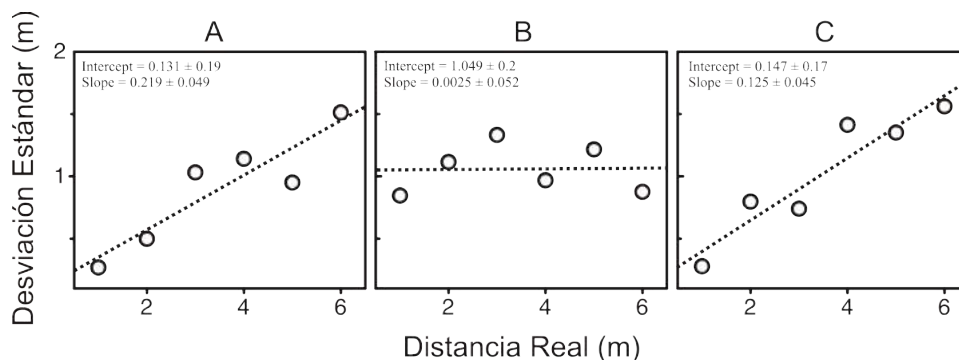


**Figura 6.6:** Experimento 3: Percepción Auditiva de Distancia (con conocimiento previo de la sala). (A) Valor medio de los juicios de distancia como una función de la distancia real obtenida con sin pistas visuales para participantes que pudieron ver la sala antes del experimento (condición i,  $n=8$ , círculos negros).

La gran precisión observada para esta condición se refleja en los valores de los parámetros (cercaos a 1) de la función de potencia ajustada ( $R^2 = 0.98$ ;  $a = 1.14 \pm 0.12$  con un exponente  $b = 0.89 \pm 0.06$ ). Como comparación, la precisión observada para los participantes del grupo A, en el experimento 1, para misma condición, fue substancialmente más baja (experimento 2, figura 3a, exponente  $b = 0.55 \pm 0.1$ ) ( $e = 4.14$ ;  $p = 0.0015$ ). Como agregado, las respuestas obtenidas en el experimento 3 fueron más precisas que las del experimento 2, para todas las condiciones testeadas (grupo A, condición ii y iii, exponente  $b = 0.71 \pm 0.08$  y exponente  $b = 0.68 \pm 0.08$ , respectivamente; y grupo B, condición iii y i, exponente  $b = 0.71 \pm 0.05$  y exponente  $b = 0.79$

$\pm 0.02$ , respectivamente), aunque, con excepción del grupo A, la condición iii ( $t = 2.4$ ;  $p = 0.034$ ), el incremento en el valor del exponente no fue estadísticamente significativo.

Finalmente, en la Figura 6.7 se muestra la variabilidad de los juicios de distancia para los siguientes datos: (a) experimento 2, grupo A, condición i ( $n = 8$ ); (b) experimento 2, grupo B, condición iii ( $n = 8$ ); y experimento 3, condición i ( $n = 8$ ) (ver Figura 6.1). De forma similar a lo reportado en el experimento 1, la variabilidad de la respuesta en ausencia de pistas visuales (Figura 6.7A y Figura 6.7C) se incrementa cuando la distancia a la fuente se incrementa, mientras, en presencia de pistas visuales, la variabilidad se mantiene relativamente constante para todas las distancias testeadas (Figura 6.7B).



**Figura 6.7:** Desviación Estándar de los juicios de distancia como una función de la distancia de la fuente. (A) Sin pistas visuales (experimento 2, condición i, grupo A,  $n=8$ ). (B) Con cuatro pistas visuales ubicadas a 2, 4, 6 y 8 metros (experimento 2, condición iii, grupo B,  $n=8$ ). (C) sin pistas visuales, para participantes que conocieron la sala antes de realizar el experimento (experimento 3, condición i,  $n=8$ ). Las líneas punteadas en (A), (B) y (C) corresponden a los ajustes lineales de los datos. Los parámetros de la función lineal ajustada se muestran en la esquina superior izquierda en cada una de las figuras.

## § 6.14 DISCUSIÓN

Los resultados obtenidos soportan la hipótesis original de que la presencia de información visual afecta la percepción auditiva de distancia. Para la condi-

ción sin pistas visuales, se observó un comportamiento similar al reportado en muchos estudios previos (ver Bekesy, 1949; Cochran et al, 1968; Simpson y Stanton, 1973; Bronkhorst y Houtgast, 1999; Zahorik, 2001, 2002; y Zahorik et al, 2005, para una revisión): la respuesta se incrementa linealmente cuando la fuente está ubicada en un rango cercano, pero converge a un cierto límite cuando la distancia de la fuente se incrementa más allá de los 3 metros. Cuando el experimento fue realizado en presencia de pistas visuales (pares de LEDs), se observaron al menos tres efectos: (1) la distancia a la fuente sonora fue sobrestimada en cinco de las seis distancias testeadas ( $D = 2, 3, 4, 5$  y  $6$  metros); (2) la respuesta de los participantes fue menos compresiva que las obtenidas sin pistas visuales y (3) la variabilidad de la respuesta fue relativamente constante para todas las distancias testeadas, mientras que, a oscuras, su valor aumenta linealmente cuando se incrementa la distancia a la fuente sonora. Además, los resultados muestran que la información visual obtenida por los participantes durante el experimento, puede ser guardada en la memoria para ser utilizada, minutos más tarde, como una referencia espacial en experimentos realizados sin pistas visuales. De hecho, se obtuvo una respuesta muy precisa en la oscuridad cuando se les permitió a los participantes inspeccionar visualmente el ambiente (condición con todas las pistas visuales disponibles) antes de realizar el experimento.

### § 6.15 RELACIÓN CON LOS RESULTADOS PASADOS

En el experimento 2, se estudió como la PAD es afectada por la presencia de pistas visuales. Considerando la escases de estudios sobre este tema, se reali-

zará una comparación con los experimentos realizados por Zahorik (2001). La metodología utilizada en este trabajo se asemeja a la utilizada por Zahorik: (a) los experimentos fueron realizados en un ambiente semireverberante; (b) los estímulos fueron presentados en un rango de seis distancias diferentes, un procedimiento que facilita los juicios de distancia auditiva (Mershon y Bowers, 1979); y (c) el experimento fue realizado en presencia y ausencia de pistas visuales.

Como agregado a los resultados reportados, los datos obtenidos bajo la condición i (experimento 2, grupo A) se ajustaron a una función de potencia de la forma  $Y = aX^b$  con un exponente significativamente menor a 1 ( $b = 0.55$ ). Los parámetros de ajuste para el experimento en ausencia de pistas visuales son muy similares a los reportados en un meta análisis de 21 estudios previos de PAD (Zahorik et al, 2005). De esta manera, este resultado otorga una base confiable y precisa de los juicios de distancia auditiva para evaluar las contribuciones de las pistas visuales en la precisión de la PAD (ver Zahorik, 2001). En la condición (i) los participantes subestimaron la distancia de la fuente sonora cuando esta se ubicó a distancias mayores a 3 metros (figura 3a). Este resultado también es consistente en los reportados en estudios previos (ver Zahorik et al, 2005, para una revisión). Sorpresivamente, cuando el experimento fue realizado sin pistas visuales, los estímulos auditivo y visual fueron localizados por los participantes a distancias similares (figura 3a, ver la línea gris para comparar). Resultados similares fueron reportados por Philbeck y Loomis (1997) y Loomis et al (1998), donde fueron obtenidas respuestas similares para estas dos modalidades cuando los experimentos fueron realizados en la oscuridad.



Cuando fueron utilizadas pistas visuales (condiciones ii y iii) durante el experimento, la subestimación observada en la condición (i) fue significativamente revertida (comparar figuras 3b y 4a). De hecho, la distancia a la fuente sonora fue sobrestimada en distancias mayores a 1 metro (figuras 3b y 3c).

Como agregado, la respuesta de los participantes fue significativamente menos compresiva, lo que se ve reflejado en un incremento significativo del exponente  $b$ . Estos resultados son consistentes con los reportados por Zahorik (2001), donde la presencia de la información visual causa un significativo incremento en el exponente de la función de potencia utilizada para ajustar los datos. Gran parte de las diferencias entre las dos condiciones se da cuando la fuente se localiza a  $D= 4, 5$  y  $6$  metros. Por ejemplo, para la condición (i), cuando la fuente sonora fue ubicada en el rango de distancias entre  $3$  y  $6$  metros, la respuesta tendió a agruparse en un rango de  $85$  cm (entre los  $2.90$  y los  $3.75$  metros), mientras que en presencia de pistas visuales (condición ii) los datos tendieron a agruparse en un rango mucho más amplio de  $220$ cm (entre  $4.81$  y los  $7.01$  metros).

La mayor diferencia entre los resultados presentados en este trabajo y los reportados por Zahorik es que para el experimento 2 (condición ii y iii) los participantes sobrestimaron la distancia de la fuente en todas las distancias testeadas, mientras que en el experimento realizado por Zahorik (2001) todas las distancias fueron subestimadas. Esta discrepancia puede haber sido causada por importantes diferencias metodológicas entre los dos experimentos. En el dispositivo experimental utilizado en este trabajo, no puede verse el parlante en ninguna de las condiciones testeadas, mientras que en el experimento realizado por Zahorik (bajo condiciones de pistas visuales disponibles) los par-

participantes podían ver el primer parlante (ubicado a 1 metro) y la sala durante la tarea. La subestimación reportada por Zahorik podría entonces deberse a un efecto de atracción producido por el parlante más cercano. Aunque en el estudio de Zahorik (2001) el efecto de captura visual en distancia no están evidente como en los trabajos de Gardner (1968) y Mershon (1980), este efecto no puede ser totalmente excluido y existe la posibilidad de que la imagen del parlante más cercano (que los participantes pueden ver todo el tiempo) pueda haber causado una atracción en la ubicación aparente de la fuente sonora.

En el diseño experimental utilizado en este trabajo se tomaron precauciones adicionales con el propósito de minimizar un efecto de atracción de las pistas visuales sobre la ubicación aparente de la fuente sonora. En primer lugar, utilizamos pistas visuales pequeñas ubicadas fuera de la línea media, y se mantuvieron encendidas durante el experimento. Esta última condición, evita posibles efectos temporales, como la sincronización entre los estímulos visuales y los estímulos auditivos, que podrían proporcionar sesgos adicionales en la ubicación de la fuente sonora. El hecho de que para el grupo B las respuestas no fueron significativamente diferentes entre las dos condiciones (si pistas visuales y con cuatro pistas visuales) indica que si existe una atracción sobre la localización de la fuente sonora, esta es mínima.

Otra diferencia con las investigaciones anteriores fue observada en el comportamiento de la variabilidad de las respuestas. Trabajos previos de PVD (Philbeck y Loomis, 1997) y PAD (Zahorik, 2001) apoyan la idea de que la presencia de información visual provoca un decremento en la variabilidad de las respuestas, comparándolo con la condición sin pistas visuales. Sin embargo, en contraste con los datos reportados por Zahorik (2001), en el presente

trabajo no se observó una reducción en la variabilidad de la respuesta en presencia de pistas visuales (ver figura 6). Esta diferencia puede deberse a diferencias metodológicas entre los dos experimentos. En el experimento realizado por Zahorik (bajo condiciones visuales) los participantes disponían de la información visual proveniente de ver el arreglo de parlantes (y la distancia entre ellos) y el entorno donde se desarrollaba la tarea durante el experimento, mientras que en el experimento realizado aquí (bajo las condiciones ii y iii) el número y complejidad de las pistas visuales fue menor.

Sorpresivamente, cuando se probó la condición (i) (sin pistas visuales) minutos después del experimento con pistas visuales (grupo B, condición iii) la respuesta fue significativamente más precisa que la obtenida por el grupo A bajo las mismas condiciones. Este resultado sugiere que la información generada por las pistas visuales durante la primera parte del experimento puede ser utilizada, minutos después, como una referencia visual para el experimento sin pistas visuales.

En acuerdo con esto, Loomis et al (1992) reportaron que cuando se remueve un feedback visual después de una exposición inicial de un objetivo visual, una imagen espacial de este objeto se mantiene en la memoria. Como referencia adicional a este fenómeno, los experimentos de percepción de distancia visual, que usan como método de respuesta la caminata directa del participante hacia el objetivo visual, indican que un participante puede usar información visual como un guía para realizar subsecuentes tareas de localización en la oscuridad. Estos experimentos han mostrado que los participantes pueden caminar a ciegas hacia los objetivos con gran precisión, luego de verlos durante un periodo de adaptación previo (Thomson, 1983; Laurent y Cavallo, 1985; Elliott, 1986;

Loomis et al, 1992, 1998; Fukusima et al, 1997).

Algo similar a esto sucede el experimento 3, donde los datos obtenidos por los sujetos en un reconocimiento previo de la sala sirven como una referencia espacial para realizar, minutos más tarde, la tarea de percepción auditiva de distancia en la oscuridad. Sorpresivamente, bajo estas condiciones la respuesta fue mucho más precisa que la obtenida por el grupo B en el experimento 2 en la condición (i). Tal vez, la diferencia entre las dos respuestas está dada por el número y la complejidad de las pistas visuales disponibles en cada condición. Mientras en el experimento 2, los participantes del grupo B solo tenían un máximo de 4 pares de LEDs como referencias visuales, los sujetos del experimento 3 tuvieron a disposición todas las pistas visuales que podían obtener de la sala iluminada. Los resultados obtenidos aquí muestran que la PAD de fuentes sonoras en la oscuridad es más precisa cuando los participantes conocen el ambiente antes del experimento. Esto sugiere que algún tipo de representación espacial del ambiente puede ser recordado durante el experimento de manera tal que incrementa la precisión en la localización de distancia de una fuente sonora.

## REFERENCIAS BIBLIOGRÁFICAS

- ALAIS, D. y BURR, D.: «The ventriloquist effect results from near-optimal bimodal integration». *Current Biology*, 2004, **14**, pp. 257–262.
- BATTAGLIA P.W., JACOBS R.A. y R.N., ASLIN: «Bayesian integration of visual and auditory signals for spatial localization». *Journal of the Optical Society of America*, 2003, **20**, pp. 1391–1397.
- BEGAULT, D.R.: «Perceptual effects of synthetic reverberation on three-dimensional audio systems». *Journal of the Audio Engineering Society*, 1992, **40**, pp. 895–904.
- BERTELSON, P. y RADEAU, M.: «Cross-modal bias and perceptual fusion with auditory-visual spatial discordance». *Perception and Psychophysics*, 1981, **29**, pp. 578–584.
- BÉKÉSY, G.V.: «Über die Entstehung der Entfernungsempfindung beim Hören (On the origin of the sensation of distance in hearing)». *Akustische Zeitschrift*, 1938, **3**, pp. 21–31.
- : «The moon illusion and similar auditory phenomena». *American Journal of Psychology*, 1949, **62**, pp. 540–552.
- BLAKE, R. y SEKULER, R.: *Perception*. McGraw-Hill, New York, 2006.
- BÜLTHOFF I., BÜLTHOFF H. y P., SINHA: «Top-down influences on stereoscopic depth-perception». *Nature Neuroscience*, 1998, **1**, pp. 254–257.
- BRENNER, R. y VANDAMME, W.: «“Perceived distance, shape and size». *Vision Research*, 1999, **39**, pp. 975–986.
- BRONKHORST, A.W. y HOUTGAST, T.: «Auditory distance perception in rooms». *Nature*, 1999, **397**, pp. 517–520.
- BUTLER R.A., LEVY E.T. y W.D., NEFF: «Apparent distance of sounds recorded in echoic and anechoic chambers». *The Journal of Experimental Psychology*, 1980, **6**, pp. 745–750.
- CABRERA D., CAPRA A. FARINA A., AZZALI A. y P., MARTIGNON: «Perceived room size and source distance in five simulated concert auditoria». *12th International Congress on Sound and Vibration Lisbon*, 2006.
- CABRERA D., KWAKE H.J., JEONGB D. y J.Y., KIMD: «Auditory room size perception for modeled and measured rooms». *Internoise Congress and Exposition on Noise Control Engineering*, 2005.
- CARLSON, V.R.: «Instructions and perceptual constancy judgment». *En Stability and Constancy in Visual Perception: Mechanisms and Processes Ed. W Epstein*, 1977, pp. 217–254.
- COCHRAN P., THROOP J. y W.E., SIMPSON: «Estimation of distance of a source of sound». *American Journal of Psychology*, 1968, **81**, pp. 198–206.
- COLEMAN, P.D.: «Failure to localize the source distance of an unfamiliar sound». *J. Acoust. Soc. Am.*, 1962, **34**, pp. 345–346.
- ELLIOTT, D.: «Continuous visual information may be important after all: A failure to replicate Thomson». *Journal of Experimental Psychology*, 1986, **12**, pp. 388–391.
- EVANS, K.K. y TREISMAN, A.: «Natural cross-modal mappings between visual and auditory features». *Journal of Vision*, 2010, **10**, pp. 1–12.
- FOLEY, J.M.: «Binocular distance perception». *J. Psychological Review*, 1980, **87**, pp. 411–434.

- FUKUSIMA S.S., LOOMIS J.M. y J.A., DASILVA: «Visual perception of egocentric distance as assessed by triangulation». *Journal of Experimental Psychology: Human Perception and Performance*, 1997, **23**, pp. 86–100.
- GARDNER, M.B.: «Proximity image effect in sound localization». *J. Acoustic. Soc. Am.*, 1968, **43**, p. 163.
- GOGEL, W.C.: «Convergence as a cue to absolute distance». *Journal of Psychology*, 1961, **52**, pp. 287–301.
- HAMEED S., VALDE K., PAKARINEN J. y V., PULKKI: «Psychoacoustic cues in room size perception». *En Proceedings of Convention, 116th Audio Engineering Society Convention, Berlin*, 2004.
- JACKSON, C.V.: «Visual factors in auditory localization». *Quarterly Journal of Experimental Psychology*, 1953, **5**, pp. 52–65.
- JONES, B.: «Visual facilitation of auditory localization in school children: A signal detection analysis». *Perception and Psychophysics*, 1975, **11**, pp. 217–220.
- KING, A.J.: «Visual facilitation of auditory localization in school children: A signal detection analysis». *Philosophical Transactions of the Royal Society of London*, 2009, **364**, pp. 331–339.
- KUBOVY, M. y VANVALKENBURG, D.: «Auditory and visual objects». *Cognition*, 2001, **80**, pp. 97–126.
- LAURENT, M. y CAVALLO, V.: «Role des modalités de prise d'informations visuelles dans un pointage locomoteur (The role of visual input modality in a locomotor pointing task)». *L'Année Psychologique*, 1985, **85**, pp. 41–48.
- LEWALD, J.: «Rapid adaptation to auditory visual spatial disparity». *Learning and Memory*, 2002, **9**, pp. 268–278.
- LOOMIS, J.M.: «Looking down is looking up». *Nature News and Views*, 2001, **414**, pp. 155–156.
- LOOMIS J.M., FUJITA N., DASILVA J.A. y S.S., FUKUSIMA: «Visual space-perception and visually directed action». *Journal of Experimental Psychology: Human Perception and Performance*, 1992, **18**, pp. 906–921.
- LOOMIS J.M., PHILBECK J.W., KLATZKY R.L. y R.G., GOLLEDGE: «Assessing auditory distance perception using perceptually directed action». *Perception and Psychophysics*, 1998, **60(6)**, pp. 966–980.
- MCGRATH R., WALDMANN T. y M., FERNSTROM: «Listening to rooms and objects». *En Proceedings of the 16th Audio Engineering Society International Conference, Rovaniemi, Finland*, 1996.
- MERSHON, D.H. y BOWERS, J.N.: «Absolute and relative cues for the auditory perception of egocentric distance». *Perception*, 1979, **8**, pp. 311–322.
- MERSHON, D.H. y KING, L.: «Intensity and reverberation as factors in the auditory perception of egocentric distance». *Perception and Psychophysics*, 1975, **18**, pp. 409–415.
- MERSHON D.H., AMERSON T.L., DESAULNIERS D.H. y KIEFER, S.A.: «Visual capture in auditory distance perception: proximity image effect reconsidered». *Journal of Auditory Research*, 1980, **20**, pp. 129–136.
- MERSHON D.H., LITTLE A.D. MCMURTRY P.L., BALLENGER W.L. y J.L., BUCHANAN: «Effects of room reflectance and background noise on perceived auditory distance». *Perception*, 1989, **18**, pp. 403–416.

- MIDDLEBROOKS, J.C. y GREEN, D.M.: «Sound Localization by Human Listeners». *Annual Review of Psychology*, 1991, **42**, pp. 135–159.
- NIELSEN, S.H.: «Auditory distance perception in different rooms». *Journal of the Audio Engineering Society*, 1993, **41**, pp. 755–770.
- ONO H., OHMI M., ROGERS B.J. y M.E., ONO: «Dynamic occlusion and motion parallax in depth perception». *Perception*, 1988, **17**, pp. 255–266.
- PERROTT D.R., BROWN K., SABERI K. y T.Z., STRYBEL: «Auditory psychomotor coordination and visual search performance». *Perception and Psychophysics*, 1990, **48**, pp. 214–226.
- PHILBECK, J.W. y LOOMIS, J.M.: «Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions». *Journal of Experimental Psychology: Human Perception and Performance*, 1997, **23**, pp. 72–85.
- PICK H.L., WARREN D.H. y J.C., HAY: «Sensory conflict in judgements of spatial direction». *Perception and Psychophysics*, 1969, **6**, pp. 203–205.
- QIAN, N.: «Binocular disparity review and the perception of depth». *Neuron*, 1997, **18**, pp. 359–368.
- RADEAU, M. y BERTELSON, P.: «The after-effects of ventriloquism». *Quarterly Journal of Experimental Psychology*, 1974, **26**, pp. 63–71.
- RECANZONE, G.H.: «Rapidly induced auditory plasticity: the ventriloquism aftereffect». *Proceedings of the National Academy of Sciences of the USA*, 1998, **95**, pp. 869–875.
- SANDVAD, J.: «Auditory perception of reverberant surrounding». *J. Acoustic. Soc. Am.*, 1999, **105**, p. 1193.
- SEMMLow, J.J. y HEEREMA, D.: «The role of accommodative convergence at the limit of fusional vergence». *Investigative Ophthalmology and Visual Science*, 1979, **18**, pp. 970–976.
- SHEELINE, C.W.: «An investigation of the effects of direct and reverberant signal interaction on auditory distance perception». *Tesis Doctoral Universidad de Standford*, 1984.
- SHELTON, B.R. y SEARLE, C.L.: «The influence of vision on the absolute identification of sound-source position». *Perception and Psychophysics*, 1980, **28**, pp. 589–596.
- SHINN-CUNNINGHAM, B.G.: «Learning reverberation: Considerations for spatial auditory displays». *En Proceedings of the 2000 International Conference on Auditory Displays, Atlanta*, 2000.
- SIMON, J.R. y CRAFT, J.L.: «Effects of an irrelevant auditory stimulus on visual choice reaction time». *Journal of Experimental Psychology*, 1970, **86**, pp. 272–274.
- SIMPSON, W.E. y STANTON, L.D.: «Head movement does not facilitate perception of the distance of a source of sound». *American Journal of Psychology*, 1973, **86**, pp. 151–159.
- SOUSA R., BRENNER E. y J.B.J., SMEETS: «A new binocular cue for absolute distance: Disparity relative to the most distant structure». *Vision Research*, 2010, **50**, pp. 1786–1792.
- STEIN B.E., HUNEYCUTT W.S., MEREDITH M.A. y L., MCDADE: «Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli». *Journal of Cognitive Neuroscience*, 1989, **1**, pp. 12–24.
- THOMSON, J.A.: «Is continuous visual monitoring necessary in visually guided locomotion?» *Journal of Experimental Psychology: Human Perception and Performance*, 1983, **9**, pp. 427–443.
- THURLOW, W.R. y JACK, C.E.: «Certain determinants of the ventriloquism effect». *Perceptual and Motor Skills*, 1973, **36**, pp. 1171–1184.

- VROOMEN, J. y DEGELDER, B.: «Sound enhances visual perception: Cross-modal effects of auditory organization on vision». *Journal of Experimental Psychology: Human Perception and Performance*, 2000, **26**, pp. 1583–1590.
- WAGENAARS, W.M.: «Localization of sound in a room with reflecting walls». *Journal of the Audio Engineering Society*, 1990, **38**, pp. 99–110.
- WALLACE M.T., HAIRSTON W.D., PERRAULT T.J.JR y B.E., STEIN: «Visual experience is necessary for the development of multisensory integration». *Journal of Neuroscience*, 2004, **24**, pp. 9580–9584.
- WARREN, D.H.: «The role of visual-auditory ‘compellingness’ in the ventriloquism effect: implications for transitivity among the spatial senses». *Cognitive Psychology*, 1973, **1**, pp. 114–133.
- WARREN D.H., WELCH R.B. y T.J., MCCARTHY: «The role of visual-auditory ‘compellingness’ in the ventriloquism effect: implications for transitivity among the spatial senses». *Perception and Psychophysics*, 1981, **30**, pp. 557–564.
- WENZEL E.M., KISTLER D.J., ARRUDA M. y F.L., WHIGHTMAN: «Localization using nonindividualized head-related transfers functions». *J. Acoustic. Soc. Am.*, 1993, **94**, pp. 111–123.
- ZAHORIK, P.: «Estimating sound source distance with and without vision». *Optometry and Vision Science*, 2001, **78**, pp. 270–275.
- : «Assessing auditory distance perception using virtual acoustics». *J. Acoustic. Soc. Am.*, 2002, **111**, pp. 1832–1846.
- ZAHORIK P., BRUNGART D.S. y A., BRONKHORST: «Auditory Distance Perception in Humans: A Summary of Past and Present Research». *Acta Acustica united with Acustica*, 2005, **91**, pp. 409–420.





## - Capítulo 7 -

---

---

### El rol del contenido espectral en la percepción auditiva de distancia

---

#### § 7.1 INTRODUCCIÓN

Como vimos en el Capítulo 3 (Localización en distancia de fuentes sonoras), para percibir la distancia a una fuente sonora un oyente puede utilizar una gran variedad de indicios, tanto acústicos como no acústicos. Uno de los primeros indicios en ser estudiados por parte de los investigadores fueron los cambios espectrales que sufre la señal en su interacción con el ambiente al viajar desde la fuente sonora hasta el lugar donde se encuentra el oyente. Sin embargo, como veremos a continuación el número de estudios que han tratado este tema ha sido escaso.

Existen por lo menos dos fenómenos relacionados con cambios en la distancia de la fuente que provocan cambios sistemáticos en el espectro de frecuencias de un sonido: el filtrado que sufre el sonido inducido por las propiedades de absorción del aire y el aumento relativo de las frecuencias bajas producidas en el campo cercano (efecto de proximidad).

Con respecto al efecto de la absorción del aire sobre el espectro de un sonido,

Nyborg y Mintzer (1955) han reportado que para una temperatura de  $20^{\circ}$  C, y una humedad relativa del 50 %, el coeficiente de absorción para un tono de 10Khz es 20 veces mayor que para un tono de 1Khz. De este modo, en su viaje desde la fuente al oyente, un sonido de banda ancha pierde mayor cantidad de energía en la zona de frecuencias agudas del espectro que en la zona de frecuencias graves. Esto hace que los sonidos provenientes de fuentes cercanas suenen más “brillantes” que los provenientes de fuentes lejanas. Es importante destacar que esta pérdida es moderada (entre 3 y 4dB por cada 100 metros a 4Khz – Ingard, 1953) y por ende puede ser percibida solamente para cambios en la distancia de la fuente relativamente grandes (mayor a 15 m según Blauert, 1997). Sin embargo, Coleman en 1968 demostró que la cantidad relativa de energía en la zona aguda del espectro puede afectar la PAD aun en distancias menores. Los experimentos de Coleman (1968) muestran como la percepción auditiva de distancia de seis sujetos fue dependiente de la frecuencia de corte de un filtro pasa alto aplicado a un pulso cuadrado (las frecuencias de corte fueron 7680Hz y 10560Hz). Coleman utilizó el método de presentaciones repetidas para sus experimentos. Los estímulos fueron presentados a partir de 2,4 hasta 8,40 metros en pasos de 60 cm. Sus resultados muestran que al aumentar la frecuencia de corte del filtro pasa alto (aumentando la cantidad relativa de frecuencias altas en el estímulo auditivo) la distancia aparente de la fuente sonora fue significativamente menor que los estímulos cuya frecuencia de corte fue menor. Coleman propuso que este efecto se debe a que el recorte en frecuencias agudas sobre el estímulo sonoro emularía el efecto del sonido atravesando grandes distancias. Resultados similares fueron obtenidos por Little, et al en 1992. Estos autores observaron

además que el nivel relativo de alta frecuencia de un sonido afectó la distancia aparente de la fuente solamente cuando los oyentes pudieron comparar entre varios estímulos filtrados y con diferentes frecuencias de corte. De este modo, demostraron que, al igual que la intensidad, los cambios espectrales producidos por la absorción del aire funcionan como un indicio relativo de distancia. Por otro lado, algunos estudios han indicado que, cuando la fuente se encuentra a menos de 2 metros del oyente (campo cercano), el cambio espectral que sufre el sonido en relación a la distancia de la fuente es totalmente inverso al que se observa en campo lejano. Békésy (1938) reportó que, en el campo cercano, el contenido de baja frecuencia de un estímulo de banda ancha aumenta cuando la distancia a la fuente sonora disminuye. Este resultado llevó a Békésy (1938) a suponer que este efecto podía ser una buena pista de PAD en distancias pequeñas. Coleman (1962), adhiere a esta hipótesis, aunque en su estudio no aporta resultados convincentes para probarla. Blauert (1997) sin embargo, indica que este efecto se da de forma muy tenue en el mejor de los casos y que por lo tanto no puede ser considerado como una pista influyente de PAD.

Dentro del campo de la percepción espacial auditiva hay diversos estudios que determinan como influye el ancho de banda de un estímulo sonoro en la localización de la fuente en el plano horizontal y vertical. Sin embargo, todavía no sabemos con claridad de que manera las cualidades espectrales de la señal sonora en sí misma afectan nuestra capacidad para estimar la distancia a una fuente sonora.

Como veremos a continuación, la localización angular del sonido en los planos horizontal y vertical puede ser influenciada de manera importante por el con-

tenido espectral de los sonidos. Los indicios que participan en la localización angular en estos dos planos han sido tratados en los Capítulo 1 (Localización horizontal de fuentes sonoras) y Capítulo 2 (Localización vertical de fuentes sonoras). A continuación se procederá a realizar una revisión de los resultados más relevantes en el estudio de la relación entre el ancho de banda del sonido y la localización angular en los planos horizontal y vertical.

## § 7.2 RELACIÓN ENTRE EL ANCHO DE BANDA DE UN SONIDO Y LA LOCALIZACIÓN HORIZONTAL

La teoría dúplex de la localización binaural de sonidos en el plano horizontal ha sido confirmada por numerosos estudios psicofísicos y fisiológicos. Sin embargo hay un punto importante a tener en cuenta. En los experimentos de Rayleigh (1907) los estímulos utilizados fueron tonos puros. En la naturaleza, la mayoría de los sonidos que escuchamos son espectralmente complejos y rara vez estamos expuestos a estímulos sinusoidales puros. Como vimos, Stevens y Newman (1936) y Mills (1958, 1960) encontraron que la precisión en la localización depende de la frecuencia del estímulo, ya que los tonos de frecuencias graves son localizados en base a la información interaural de fase (ITD) y los tonos agudos en base a la diferencia interaural de amplitud (ILD). Agregaron además que la localización en frecuencias del rango ubicado entre los dos tipos de localización es imprecisa. La diferencia entre estos dos estudios reside en que Stevens y Newman reportan que el rango medio, el cual supuestamente refleja una pobre localización, se encuentra entre 2000Hz y 4000Hz, mientras

que Mills ubica este rango entre los 1250Hz y los 2000Hz. En otro caso, en donde los estímulos fueron presentados mediante auriculares, los oyentes utilizaron la diferencia de fase hasta alrededor de 1000Hz y 1300Hz (Licklider et al, 1950; Zwislocki y Feldman, 1956). Esto indica que existen ambigüedades entre el uso del ILD y el ITD en algunos rangos de frecuencias.

Cuando aparece algún tipo de ambigüedad en la información que puede proporcionar el ITD y el ILD, los indicios espectrales pueden servir para resolverlos (Mills, 1972; Middlebrooks y Green, 1991). Estos indicios espectrales, los cuales pueden ser tanto monoaurales como binaurales, se producen debido al filtrado que sufre la señal sonora al hacer contacto con diferentes partes del cuerpo, principalmente la cabeza, los hombros y el pabellón auricular. Esta información es utilizada por el sistema nervioso para estimar la procedencia del sonido en el plano vertical y son utilizados además para reconocer si la fuente sonora se encuentra delante o detrás del oyente (Middlebrooks, 1992; Slattery y Middlebrooks, 1994).

Estas pistas espectrales de localización pueden generarse a frecuencias por encima de los 4KHz, en donde se pueden producir valores de ILD por encima de los 20dB (Hebrank y Wright, 1974). Con respecto al ancho de banda del estímulo en si (y no a como este es afectado por nuestro cuerpo), existen numerosos trabajos que estudiaron como influye esto en la localización espacial de una fuente sonora en acimut. En principio varios estudios han demostrado que la eficacia en la localización mejora cuando el ancho de banda del estímulo utilizado es mayor (Roffler y Butler, 1968a; Hebrank y Wright, 1974; Butler y Planert, 1976; Butler, 1986; King y Oldfield, 1997; Basso, 2006; Moore, 2012). Por ejemplo, Butler (1986) reportó una importante reducción en errores de

localización horizontal asociados con el incremento en el ancho de banda de los estímulos que estaba utilizando en sus experimentos, especialmente en aéreas donde resulta más difícil realizar la localización (cono de confusión, adelante y detrás de la cabeza). Butler concluye que la localización binaural se hizo más precisa cuando los sonidos presentaban un ancho de banda de por lo menos 2Khz, y además la precisión aumentaba cuando el ancho de banda lo hacía. Moore (2012) indica que para sonidos con un ancho de banda pequeño, como una senoide o ruido de banda angosta, el ITD y el ILD pueden ser ambiguos en algunas zonas. Esta ambigüedad puede resolverse mediante el incremento del ancho de banda del estímulo de manera que se produzcan más líneas de información independiente para el sistema auditivo (Jeffress, 1972; Stern y Trahiotis, 1995).

### § 7.3 RELACIÓN ENTRE EL ANCHO DE BANDA DE UN SONIDO Y LA LOCALIZACIÓN VERTICAL

Como vimos en el Capítulo 2 (Localización vertical de fuentes sonoras) la localización angular en el plano vertical depende de indicios espectrales. Obviamente, la señal sonora deberá tener el ancho de banda espectral necesario para que este filtrado genere información relevante sobre la ubicación de la fuente en el plano vertical. Como veremos a continuación varios estudios han avalado esta hipótesis.

Uno de los primeros trabajos experimentales que midieron la efectividad en la localización de sonidos en el plano vertical fue el realizado por Pratt en 1930. Este reportó que cuando se les pidió a los oyentes ubicar ráfagas de tonos

puros en el eje vertical, estos tendían a ubicarlos en una escala vertical dependiente de la altura de los mismos (del inglés, “pitch”). Los oyentes percibían los sonidos agudos por encima de los sonidos graves aunque la ubicación de la fuente siempre era la misma. Resultados similares fueron reportados por Trimble (1934) y Roffler y Butler (1968b), aun cuando los estudios de estos últimos fueron realizados con personas con ceguera congénita y niños pequeños los cuales no tenían conciencia de nociones como “frecuencias altas” o “frecuencias bajas”. Una implicación de estos datos fue que los estímulos tonales tienen características espaciales intrínsecas. Sin embargo quedó claro que la habilidad para percibir la procedencia angular en el plano vertical, cuando se utilizan tonos puros como estímulo, es poco robusta. Es por esto que más adelante varios autores realizaron experimentos utilizando estímulos con diferentes anchos de banda. Roffler y Butler (1968a) investigaron la habilidad de los oyentes para localizar un sonido en el plano vertical utilizando diferentes tipos de estímulos diferenciados por su complejidad espectral. Los resultados mostraron que para localizar de forma efectiva un estímulo auditivo en el plano vertical estos debían poseer un contenido espectral complejo y debían incluir frecuencias por encima de los 7Khz.

Batteau (1966), Condamine (1965) y Thurlow y Runge (1967) han reportado que el ruido blanco se comportó como un buen estímulo para ser utilizado en experimentos de localización vertical. Los tres estudios coincidieron en indicar que los filtrados espectrales producidos por las características fisiológicas humanas son más efectivos en el contenido espectral uniforme que presenta el ruido blanco. Hebrank y Wright (1974) indican que es necesario que el ancho de banda de un estímulo contenga información entre los 4 y los 16Khz



para ser localizado correctamente en el plano vertical. Estudios posteriores a estos, sobre localización de fuentes sonoras en elevación han utilizado pistas espectrales monoaurales, manipulación de la elevación de la fuente sonora, diferentes fuentes sonoras de banda angosta, filtros pasa-banda, banda ancha, etc (Bloom, 1977; Hofman y Van Opstal, 1998; Jin et al, 2004; Langendijk y Bronkhorst 2002; Meijer, 1992; Sodnik et al, 2004; Sodnik et al, 2005; Watkins, 1978; Zibera y Zazula, 2003). Estudios que han usado sonidos de banda angosta (tonos o ruido de banda angosta) han demostrado que la elevación percibida varía con la frecuencia central del sonido independientemente de su elevación real (Hofman y Van Opstal, 1998). En los estudios donde se han utilizado bandas de ruido gaussiano (Langendijk y Bronkhorst 2002), se observó que las pistas espectrales en las bandas de frecuencia de 4-16Khz son esenciales para la correcta localización de sonidos de banda ancha y que las pistas más importantes de elevación están presentes en la banda que va desde 5.7 a 11.3Khz. Otro estudio que utilizó ruidos filtrados de banda ancha, con un filtro notch cuya frecuencia central estuvo ubicada en diferentes zonas del espectro, demostró que una fuerte caída en la zona de frecuencias agudas del espectro de señales presentadas monoauralmente, es suficiente para evocar la sensación de elevación de una fuente sonora (Bloom, 1977). En 2012, un estudio presentado por Iwaya et al, reportó que efectivamente se necesita estímulos con información espectral compleja (entre los 4 y 8Khz) para poder percibir sin problemas la elevación real de una fuente sonora.

Como observamos, existen numerosos trabajos realizados sobre la percepción de estímulos de diferentes anchos de banda y tonos puros tanto en acimut como en elevación. En ambos casos se ha reportado que los sonidos de banda

ancha son localizados con mayor precisión que los tonos puros. Sin embargo, la cosa es totalmente diferente en el estudio de la PAD y, como veremos a continuación, el número de trabajos que abordaron este tema es llamativamente escaso.

#### § 7.4 RELACIÓN ENTRE EL ANCHO DE BANDA ESPECTRAL Y LA PAD

Extrañamente son pocos los investigadores que han estudiado el rol del espectro en la PAD. Sin embargo, un trabajo reciente publicado por Kopco y Shinn-Cunningham (2011) aborda de lleno el estudio de la relación entre el ancho de banda espectral de un estímulo auditivo y la PAD. En este estudio fueron examinados los efectos de la frecuencia del estímulo y el ancho de banda espectral para sonidos reproducidos por fuentes sonoras virtuales ubicadas en el campo cercano (0.15 a 1.7 metros) en un entorno virtual reverberante. Las frecuencias centrales de los estímulos estaban entre los 300Hz y 5700Hz, y el ancho de banda variaba desde 200Hz hasta 5400Hz. Los resultados de este estudio mostraron que en un ambiente reverberante virtual, la PAD de fuentes cercanas fue dependiente del contenido espectral de los estímulos auditivos. La característica espectral que más influyó en la respuesta de los oyentes fue la presencia de frecuencias bajas en el estímulo auditivo, ya que cuando se reproducían estímulos que contenían energía en frecuencias alrededor de los 300Hz la ubicación de la fuente fue percibida de manera más precisa que al utilizar estímulos de alta frecuencia (centrados en 5,7KHz). Además los autores reportaron que el ancho de banda del estímulo auditivo no influyó sobre la PAD. En otro trabajo realizado por el mismo grupo (Kopco et al, 2012) se

estudió directamente la influencia del ancho de banda de los estímulos sobre la PAD. Al igual que el trabajo anterior se utilizó un entorno virtual y la PAD fue explorada entre 0.15 y 2 metros (campo cercano). En este trabajo se utilizaron estímulos de ruido blanco con diferentes tipos de filtrado: pasa bajo (300-3000Hz), pasa alto (3000-5700Hz) y pasa banda (300-5700Hz). Los resultados de este estudio no soportaron la hipótesis de que la PAD, en ambientes reverberantes, depende del contenido espectral del estímulo. Los autores sugieren que esto puede estar directamente relacionado con el uso de HRTFs no individuales para la creación de los estímulos.

Si bien ambos trabajos estudiaron el efecto del ancho de banda espectral de un estímulo auditivo sobre la PAD, lo hacen en el plano cercano (nunca pasando los 2 metros) y con entornos acústicos virtuales utilizando HRTF no individuales. El hecho de que se evalúe la respuesta de los oyentes a corta distancia no garantiza que el ancho de banda espectral no influya sobre la PAD en distancias mayores, debido a que aparentemente se utilizan pistas diferenciadas para la PAD en el campo cercano y lejano. Además, tampoco ayuda el hecho de haber utilizado HRTF no individuales debido a que es sabido que estas no llegan a ser totalmente efectivas para emular la distancia de una fuente sonora, aun menos en el plano frontal.

En el presente capítulo se mostraran los resultados obtenidos en estudios comparativos mediante experimentos psicoacústicos con el fin de estudiar la relación entre las cualidades espectrales de un estímulo auditivo y la PAD. Los experimentos fueron realizados en un entorno real y se utilizaron distancias mayores a las utilizadas en los trabajos previos (de 1 a 6 metros). Para estos experimentos se utilizó una gran variedad de estímulos auditivos de diferente

cualidad espectral, desde tonos puros, hasta ruido blanco y rosa, pasando por ruidos filtrados centrados a diferentes frecuencias y con diferentes anchos de banda, con el fin de establecer el rol del ancho de banda espectral en la PAD.

## § 7.5 MÉTODOS GENERALES

El objetivo de este estudio fue examinar cómo afecta la cualidad espectral de un estímulo sonoro a la PAD. Para esto se realizaron un total de 4 experimentos en donde se presentaron estímulos con diferentes anchos de banda y ubicados en diferentes zonas del espectro. De esta manera, además de estudiar la influencia del ancho de banda espectral en la PAD, se observó que zona del espectro otorga información relevante para percibir correctamente la distancia de una fuente sonora. En el experimento 1, se utilizaron estímulos auditivos con información en todo el rango audible, Ruido Blanco y Ruido Rosa, para observar la respuesta con toda la información espectral disponible. El experimento 2 evaluó la PAD de los oyentes mediante la presentación de tonos puros de diferentes frecuencias, utilizando así estímulos con contenido espectral muy simple. En el experimento 3 se utilizaron estímulos con grandes anchos de banda en la zona baja media y alta del espectro de frecuencias audible. En la última serie de experimentos (que conforman el experimento 4) se utilizaron estímulos de ruido rosa filtrado centrado a diferentes frecuencias y con diferentes anchos de banda.

## § 7.6 ENTORNO EXPERIMENTAL, SET UP Y PROCEDIMIENTO

Todos los experimentos fueron realizados en el mismo entorno, con el mismo dispositivo y el mismo procedimiento utilizados en el experimento 3 descrito en el Capítulo 6 de la presente tesis.

## § 7.7 PARTICIPANTES

Un total de 48 voluntarios (38 hombres y 10 mujeres) participaron en los experimentos. Los voluntarios fueron estudiantes de grado y graduados de la Licenciatura en Composición con Medios Electroacústicos. El rango de las edades fue de los 19 a los 46 años (edad media, 26,3 años). A pesar de que no se realizaron medidas de agudeza visual y sensibilidad auditiva, todos los participantes informaron una visión normal (o corregida) y una audición normal. Todos los participantes conocieron a priori la sala donde se realizó el experimento. Sin embargo, en ningún momento se les indicó a los oyentes las dimensiones de la misma (ver métodos del Experimento 4, Capítulo 6). Todos los participantes hicieron el experimento de forma consciente y firmaron una forma por escrito indicando su consentimiento para participar del mismo.

Antes de entrar a la sala donde se realizó el experimento, cada participante fue instruido sobre la tarea que debía realizar. Las instrucciones enfatizaban que los reportes debían estar basados en la distancia aparente de la fuente, en vez de tratar de estimar de forma objetiva la distancia exacta de la misma (ver Carlson, 1977). Una vez instruidos se los condujo a la sala de experimentación, se los sentó en una silla posicionada en el punto de distancia cero (Capítulo

6 - Figura 6.2) y se les taparon los ojos, para que no tuvieran información visual sobre la localización de la fuente sonora durante el experimento.

## § 7.8 ESTÍMULOS AUDITIVOS

En el experimento 1 se utilizaron estímulos de Ruido Blanco y Ruido Rosa. Para el Experimento 2 los estímulos fueron tonos puros de 500Hz, 1KHz, 2KHz y 4KHz. En el experimento 3 se utilizaron tres estímulos diferentes: Ruido Rosa filtrado con un filtro pasa bajos (20Hz-1KHz), Ruido Rosa filtrado con un filtro pasa banda (1-3KHz) y Ruido Rosa filtrado por un filtro pasa alto (3-20KHz). Para el experimento 4 se utilizaron estímulos de Ruido Rosa filtrados con filtros centrados en cuatro frecuencias diferentes (500Hz, 1KHz, 2KHz y 4KHz) y con anchos de banda de 1/12 de octava, 1/6 de octava, una octava y una octava y media. Todos los estímulos utilizados tuvieron una duración de 500ms, con una envolvente de amplitud al inicio y al final de 50ms.

Las señales fueron generadas en Matlab y reproducidas mediante una placa de sonido a una frecuencia de muestreo de 44.1Khz con una resolución de 24 bits. El nivel de presión sonora de los estímulos fue de 70dBA, medido en la posición del oyente con la fuente sonora ubicada a 1 m (sonómetro RION NL-32). De este modo, la intensidad fue una importante pista durante el experimento. Entre prueba y prueba, se presentó un sonido enmascarante en los parlantes ubicados a los lados del participante (ver métodos del Capítulo 6, Figura 6.2).

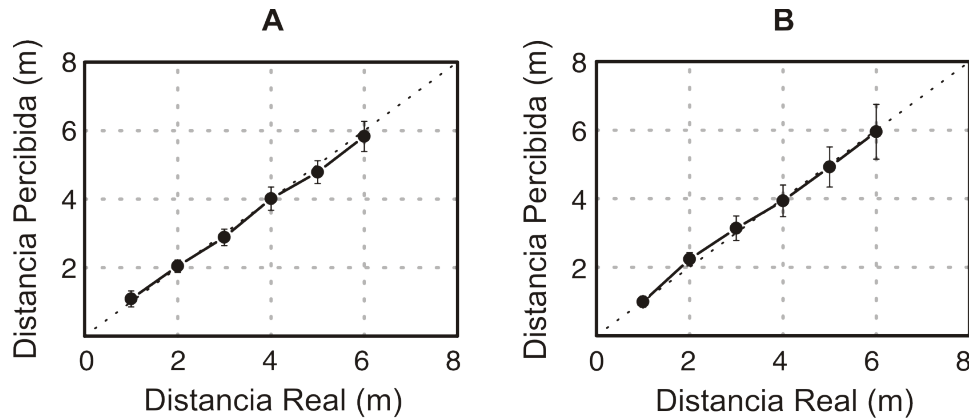
## § 7.9 EXPERIMENTO 1

Este experimento fue realizado para evaluar precisión de los oyentes para percibir la distancia a una fuente sonora cuando se utilizan estímulos auditivos que contienen información espectral en todo el rango audible (20Hz-20KHz). Para esto se estimuló con ruido blanco (n=8) y ruido rosa (n=8). En el caso del ruido blanco, la información espectral es plana por lo que contiene la misma energía en todo el espectro. En el caso del ruido rosa, la energía espectral decae exponencialmente. De este modo la señal posee la misma energía pero por octavas. Este último tipo de ruido fue utilizado para generar los estímulos usados en el Experimento 4, ya que de esta manera se obtiene la misma energía por octavas y no hace falta normalizar los estímulos.

### 7.9.1. Resultados

En la Figura 7.1 se muestran las medias de los juicios subjetivos de distancia obtenidos como una función de la distancia física utilizando como estímulos ruido blanco (n = 8; A) y ruido rosa (n = 8; B).

Cada punto corresponde a la media geométrica de las tres presentaciones y de todos los participantes. Se ajustó a estos resultados una función lineal (de la forma  $Y = a + sX$ ), para observar la pendiente de la respuesta obtenida (Ruido Blanco:  $R^2 = 0,99$ ;  $a = 0,14 \pm 0,05$  y la pendiente  $s = 0,94 \pm 0,01$ ; Ruido Rosa:  $R^2 = 0,98$ ;  $a = 0,07 \pm 0,11$  y la pendiente  $s = 1,00 \pm 0,04$ ) y una función de potencia (de la forma  $Y = a * X^b$ ) para observar la compresión de la misma (Ruido Blanco:  $R^2 = 0,99$ ;  $a = 1,04 \pm 0,04$  y el exponente  $b = 0,95 \pm 0,02$ ; Ruido Rosa:  $R^2 = 0,99$ ;  $a = 1,08 \pm 0,05$  y el exponente  $b$



**Figura 7.1:** (A)- *Ruido Blanco*,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido blanco; (B)- *Ruido Rosa*,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa.

$= 0,94 \pm 0,04$ ). Como se observa en los ajustes, tanto la pendiente (s) de la función lineal como el exponente (b) en la función de potencia (ambos cercanos a 1) muestran una relación lineal y precisa entre la posición real de la fuente y la distancia percibida por los oyentes para todas las distancias testeadas, indicando una buena performance para este tipo de estímulos en las condiciones experimentales utilizadas. Estos valores fueron similares a los obtenidos en el Capítulo 6 de esta tesis para las mismas condiciones, donde se utilizó ruido blanco como estímulo.

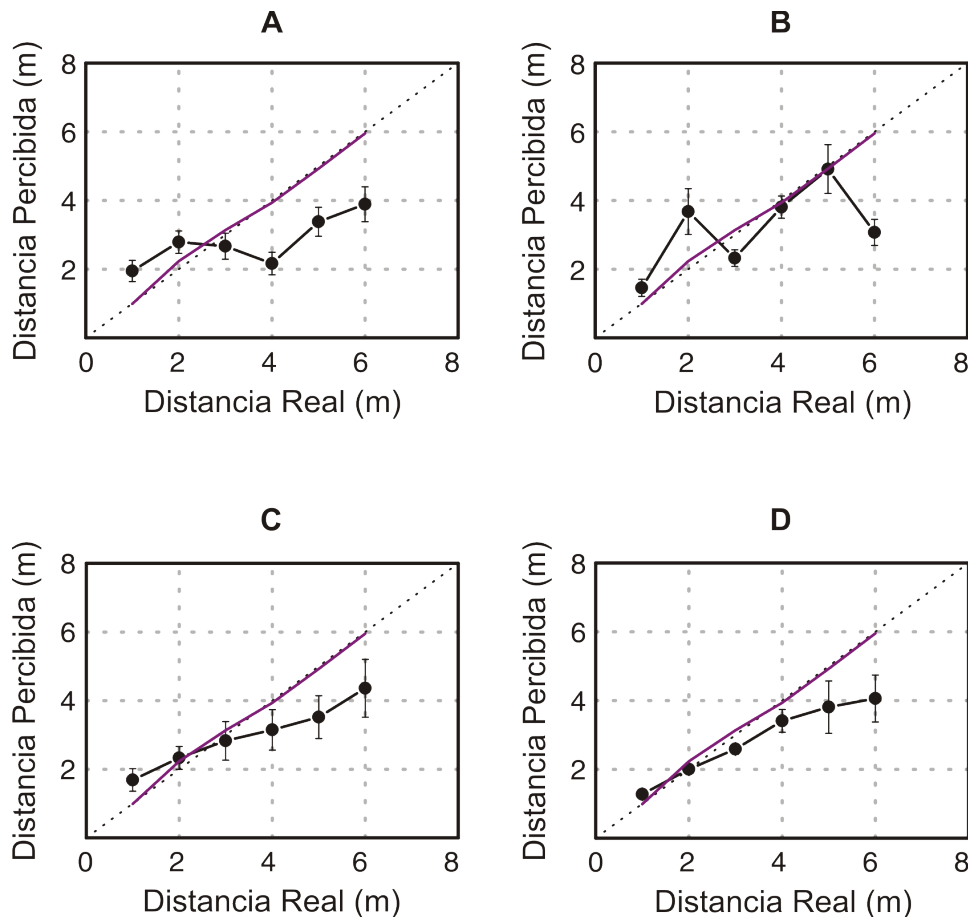
## § 7.10 EXPERIMENTO 2

El objetivo de este experimento fue evaluar la precisión de los oyentes para percibir la distancia a una fuente sonora cuando se utilizan estímulos auditivos de contenido espectral simple. De esta manera se eligió trabajar con cuatro tonos puros ubicados en diferentes lugares del espectro (500Hz, 1KHz, 2KHz y 4KHz), con el fin de determinar además si la zona del espectro donde se encuentra el estímulo afecta a la PAD.



## 7.10.1. Resultados

En la Figura 7.2 se muestran las medias de los juicios subjetivos de distancia obtenidos como una función de la distancia física ( $n = 8$ ) utilizando como estímulos tonos puros de 500Hz (A), 1KHz (B), 2KHz (C) y 4KHz (D).



**Figura 7.2:** (A) – Tono 500Hz,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo un tono puro de 500Hz; (B) – Tono 1KHz,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo un tono puro de 1KHz;– (C), Tono 2KHz,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo un tono puro de 2KHz; (D)– Tono 4KHz,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo un tono puro de 4KHz; En todas las figuras se agregó para comparación la respuesta obtenida con ruido rosa en el experimento 1 (línea púrpura).

En las cuatro figuras se agregó, para poder realizar una comparación, la respuesta obtenida para Ruido Rosa en el Experimento 1 (línea púrpura). A

simple vista puede observarse que la respuesta de los oyentes cuando se utilizan tonos puros como estímulos auditivos fue menos precisa que la obtenida con ruido rosa. Comparando las cuatro figuras puede observarse claramente como la respuesta mejora cuando el estímulo se encuentra por encima de 1KHz, siendo menos precisa la respuesta por debajo de este valor. Sin embargo, ni siquiera cuando se estimuló con un tono de 4 KHz la respuesta fue tan precisa como la obtenida tanto con ruido rosa como con ruido blanco. Los ajustes de una función lineal y de potencia para cada una de las curvas muestra claramente esta tendencia. Para el tono de 500Hz, no se obtiene un ajuste adecuado para ninguna de las funciones, y además los valores de pendiente y compresión se encuentran muy alejados a los obtenidos con Ruido Rosa y Ruido Blanco lo que denota una respuesta poco precisa (función lineal:  $R^2 = 0,55$ ;  $a = 1,70 \pm 0,45$  y la pendiente  $s = 0,31 \pm 0,11$ ; función de potencia:  $R^2 = 0,48$ ;  $a = 1,90 \pm 0,39$  y el exponente  $b = 0,33 \pm 0,14$ ). Para el tono de 1KHz, se observa la misma tendencia en los ajustes, aunque con un poco más de precisión que con el tono de 500Hz (función lineal:  $R^2 = 0,17$ ;  $a = 1,88 \pm 1,02$  y la pendiente  $s = 0,37 \pm 0,26$ ; función de potencia:  $R^2 = 0,26$ ;  $a = 2,08 \pm 0,77$  y el exponente  $b = 0,37 \pm 0,26$ ). Con el tono de 2KHz se obtiene una respuesta más precisa lo que se ve claramente reflejado en los valores del ajuste de ambas funciones. En este caso, los valores de pendiente ( $s$ ) y exponente ( $b$ ) son mayores que los obtenidos con los tonos puros anteriores, mostrando una mejora en la precisión por parte de los oyentes, aun cuando se observa una clara subestimación después de los 2 metros (función lineal:  $R^2 = 0,97$ ;  $a = 1,25 \pm 0,13$  y la pendiente  $s = 0,49 \pm 0,03$ ; función de potencia:  $R^2 = 0,95$ ;  $a = 1,58 \pm 0,13$  y el exponente  $b = 0,52 \pm 0,05$ ). Cuando se utilizó el

tono de 4KHz se observa una mejora aun más marcada en la respuesta de los oyentes lo que se ve reflejado una vez más en los valores del ajuste de las funciones lineal y de potencia (función lineal:  $R^2 = 0,96$ ;  $a = 0,84 \pm 0,18$  y la pendiente  $s = 0,57 \pm 0,04$ ; función de potencia:  $R^2 = 0,98$ ;  $a = 1,30 \pm 0,08$  y el exponente  $b = 0,65 \pm 0,04$ ). Estos valores muestran una mayor precisión en la respuesta de los oyentes que la obtenida con los demás tonos. Sin embargo, se observa una fuerte subestimación de la distancia real a la fuente en distancias mayores a 2 metros.

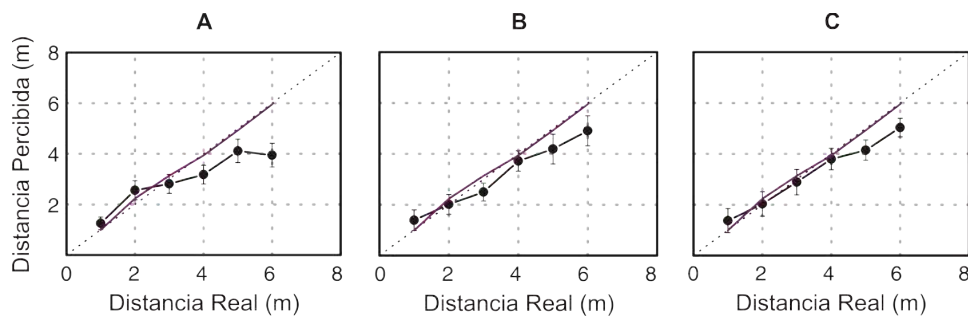
### § 7.11 EXPERIMENTO 3

En los experimentos anteriores se aprecia claramente como el contenido espectral del estímulo auditivo afecta de manera contundente la PAD. En el experimento 1 se observó una respuesta muy precisa cuando se utilizaron estímulos con información espectral en todo el rango de frecuencias mientras que en el experimento 2, se puede ver como la respuesta se vuelve menos precisa cuando se utilizan tonos puros. Los resultados de este experimento muestran además que la PAD se ve afectada por la ubicación en el espectro de los estímulos ya que los tonos mayores a 2 KHz mostraron ser más efectivos que los tonos de 500 y 1000 Hz.

En el siguiente experimento utilizamos estímulos formados por ruido rosa filtrado con amplios anchos de banda ubicados en diferentes zonas del espectro de frecuencias con el fin de estudiar de qué manera se ve afectada la PAD cuando se estimula con sonidos de banda ancha ubicados en diferentes zonas del espectro de frecuencias.

### 7.11.1. Resultados

En la Figura 7.3 se puede observar las medias de los juicios subjetivos de distancia obtenidos como una función de la distancia física ( $n = 8$ ) utilizando como estímulos Ruido Rosa filtrado entre 20Hz-1KHz (A), 1-3KHz (B) y 3-20KHz (C).



**Figura 7.3:** (A) – 20Hz-1KHz,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa filtrado entre 20Hz y 1KHz; (B) -1KHz-3KHz,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa filtrado entre 1KHz y 3KHz; (C)– 3KHz-20KHz,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa filtrado entre 3KHz y 20KHz; En todas las figuras se agregó para comparación la respuesta obtenida con ruido rosa en el experimento 1 (línea purpura).

En todas las figuras de este experimento se agregó, como comparación, la respuesta para Ruido Rosa del experimento 1 (línea purpura). Los resultados obtenidos muestran claramente que en los tres casos la respuesta fue más precisa que la obtenida con tonos puros y, por el contrario, menos precisa que la obtenida con ruido rosa y ruido blanco. Si bien a simple vista no puede apreciarse una diferencia significativa entre las tres respuestas, el ajuste de las funciones lineales y de potencia arroja resultados interesantes.

Cuando se utiliza como estímulos Ruido Rosa filtrado entre 20Hz-1KHz (Figura 7.3A) se obtienen valores de pendiente y compresión bastante bajos, indicando una baja precisión de los oyentes para percibir la distancia a la

fuente sonora (función lineal:  $R^2 = 0,89$ ;  $a = 0,89 \pm 0,28$  y la pendiente  $s = 0,58 \pm 0,08$ ; función de potencia:  $R^2 = 0,93$ ;  $a = 1,41 \pm 0,15$  y el exponente  $b = 0,60 \pm 0,08$ ). De este modo se observa que en la parte baja del espectro (de 20 a 1000 Hz) no basta un estímulo de banda ancha para percibir correctamente la distancia a la fuente. Las respuestas de los oyentes se vuelven más precisas y lineales al utilizar como estímulo Ruido Rosa filtrado entre 1KHz-3KHz (Figura 7.3B). Los valores de ajuste de las respuestas de los oyentes a este estímulo son los más cercanos a los obtenidos en el primer experimento (función lineal:  $R^2 = 0,97$ ;  $a = 0,55 \pm 0,19$  y la pendiente  $s = 0,72 \pm 0,05$ ; función de potencia:  $R^2 = 0,96$ ;  $a = 1,15 \pm 0,13$  y el exponente  $b = 0,79 \pm 0,08$ ). Por último, para el estímulo de Ruido Rosa filtrado entre 3KHz-20KHz (Figura 7.3C) se obtiene una respuesta similar a la anterior, con valores de pendiente y compresión de los ajustes muy cercanos a 1 (función lineal:  $R^2 = 0,98$ ;  $a = 0,66 \pm 0,14$  y la pendiente  $s = 0,72 \pm 0,03$ ; función de potencia:  $R^2 = 0,98$ ;  $a = 1,26 \pm 0,09$  y el exponente  $b = 0,76 \pm 0,04$ ). Los resultados obtenidos en este experimento, combinados con los del Experimento 2, muestran que estímulos con contenido espectral en el rango medio- agudo, específicamente por encima de 1KHz, son percibidos de forma más precisa por parte de los oyentes.

## § 7.12 EXPERIMENTO 4

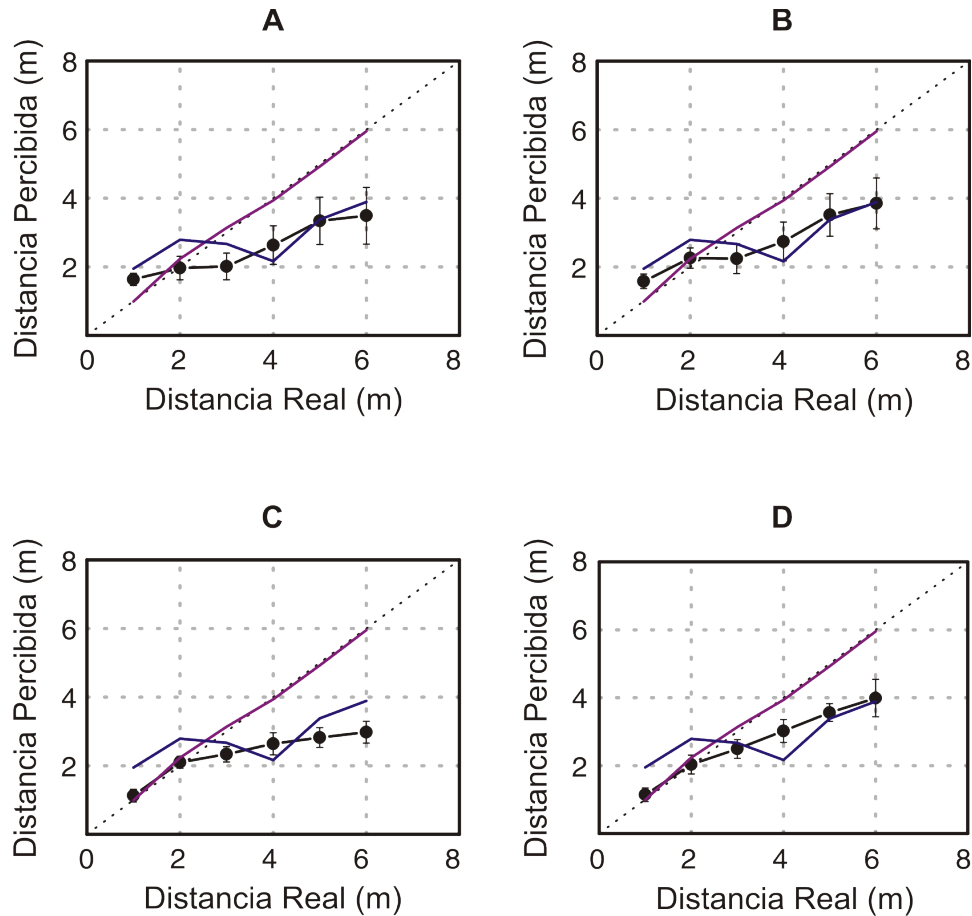
La información obtenida en los experimentos 2 y 3 fue utilizada para diseñar las diferentes instancias de este último experimento. En este se utilizan estímulos de ruido rosa filtrados con filtros centrados en 500Hz, 1KHz, 2KHz y 4KHz

y con diferentes anchos de banda, para poder determinar si esto afecta de alguna manera a la PAD y cambia con respecto a la zona del espectro donde se ubica. Se utilizan anchos de banda de 1/12 octava, 1/6 de octava, octava y octava y media.

### 7.12.1. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 500Hz

En la Figura 7.4 se puede observar las medias de los juicios subjetivos de distancia obtenidos como una función de la distancia física ( $n = 8$ ) utilizando como estímulos ruido rosa filtrado con un filtro centrado en 500Hz.

Los anchos de banda utilizados fueron: 1/12 de octava (A), 1/6 octava (B), una octava (C) y una octava y media (D). En todas las figuras se agregó, con fines comparativos, la respuesta obtenida para Ruido Rosa en el Experimento 1 (línea purpura), y la respuesta para un tono puro de 500Hz obtenida en el experimento 2 (línea azul). Los resultados muestran claramente como, cuando el ancho de banda de la señal aumenta, la respuesta mejora en casi todas las distancias testeadas. Sin embargo, ni siquiera cuando se utilizó el ancho de banda más grande (una octava y media Figura 7.4D) se obtuvo una precisión similar a la obtenida en el Experimento 1 con sonidos que poseen información en todo el espectro. Se realizó un ajuste de una función lineal (de la forma de la forma  $Y = a + s * X$ ) y una función de potencia de la forma ( $Y = a * X^b$ ) con el fin de obtener información sobre las pendientes y los niveles de comprensión de las curvas respectivamente. Tanto los valores de la pendiente ( $s$ ) como del exponente ( $b$ ) muestran claramente como, cuando el ancho de banda de la señal aumenta, la respuesta de los oyentes es más precisa.



**Figura 7.4:** *Ruido Rosa filtrado centrado en 500Hz con diferentes anchos de banda. (A)– 1/12 octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 500Hz con un ancho de banda de 1/12 de octava; (B) – 1/6 octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 500Hz con un ancho de banda de 1/6 de octava; (C)– Octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 500Hz con un ancho de banda de una octava; (D)– Octava y Media,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 500Hz con un ancho de banda de una octava y media. En todas las figuras se agregó para comparación la respuesta para ruido rosa (línea púrpura) y la respuesta para un tono de 500Hz (línea azul).*

La respuesta obtenida con Ruido Rosa filtrado con un ancho de banda de 1/12 de octava y centrado en 500Hz (Figura 7.4A) muestra valores de pendiente y compresión muy bajos, lo que estaría indicando una respuesta poco precisa (función lineal:  $R^2 = 0,91$ ;  $a = 1,25 \pm 0,11$  y la pendiente  $s = 0,34 \pm 0,04$ ; función de potencia:  $R^2 = 0,84$ ;  $a = 1,57 \pm 0,10$  y el exponente  $b = 0,37 \pm$

0,07).

Cuando se amplía el ancho de banda del estímulo hasta 1/6 de octava (Figura 7.4B), puede observarse un aumento en la precisión de la respuesta (función lineal:  $R^2 = 0,93$ ;  $a = 1,18 \pm 0,13$  y la pendiente  $s = 0,43 \pm 0,05$ ; función de potencia:  $R^2 = 0,91$ ;  $a = 1,56 \pm 0,10$  y el exponente  $b = 0,46 \pm 0,05$ ), aunque la diferencia entre estas dos respuesta no es significativa.

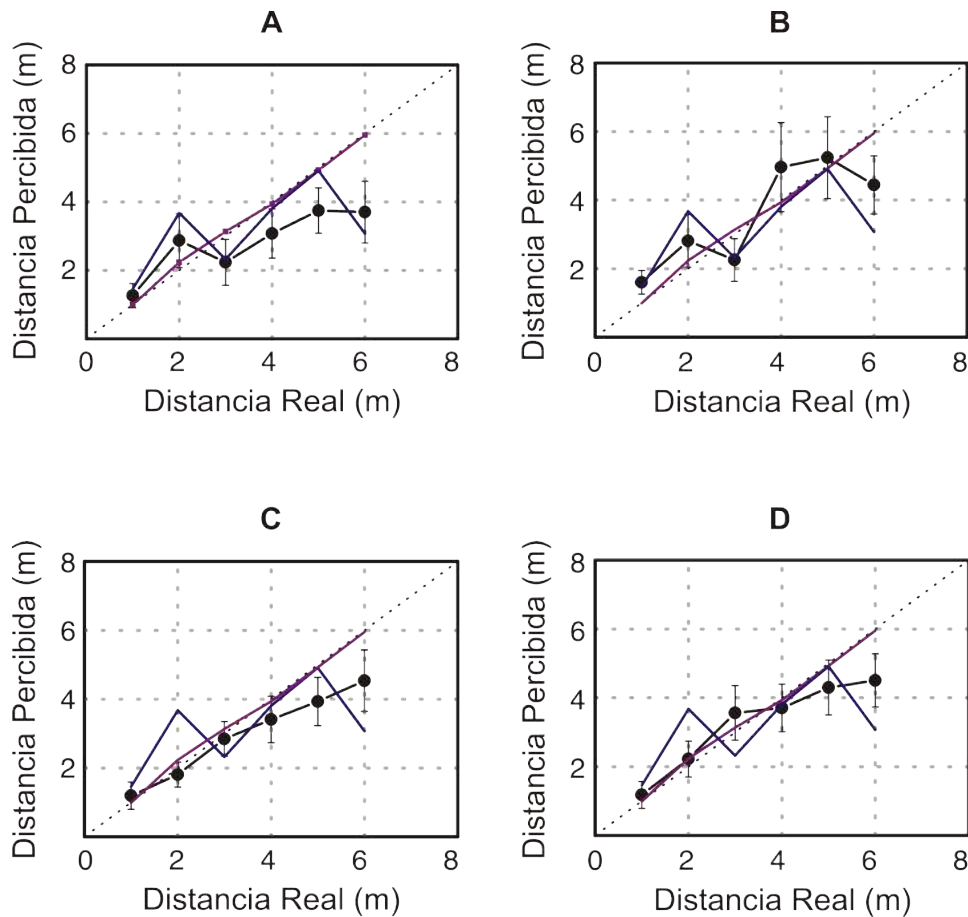
Al utilizar un ancho de banda de una octava (Figura 7.4), la respuesta no difiere significativamente a lo obtenido en la condición anterior, (función lineal:  $R^2 = 0,80$ ;  $a = 1,09 \pm 0,25$  y la pendiente  $s = 0,36 \pm 0,07$ ; función de potencia:  $R^2 = 0,90$ ;  $a = 1,36 \pm 0,12$  y el exponente  $b = 0,46 \pm 0,07$ ). Cuando el ancho de banda se aumenta hasta una octava y media, tanto el valor de la pendiente (s) como el del exponente (b) aumentan lo que representa un aumento en la precisión de la respuesta de los oyentes (función lineal:  $R^2 = 0,98$ ;  $a = 0,65 \pm 0,11$  y la pendiente  $s = 0,58 \pm 0,03$ ; función de potencia:  $R^2 = 0,99$ ;  $a = 1,19 \pm 0,04$  y el exponente  $b = 0,67 \pm 0,02$ ). Los resultados obtenidos muestran claramente que la precisión en los juicios de distancia mejora cuando el ancho de banda de la señal (en este caso centradas en 500 Hz) aumenta.

#### **7.12.2. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 1KHz**

En la Figura 7.5 se exponen las medias de los juicios subjetivos de distancia obtenidos como una función de la distancia física ( $n = 8$ ) utilizando ruido rosa filtrado con un filtro centrado en 1KHz con anchos de banda de: 1/12 de octava (A), 1/6 octava (B), una octava (C) y una octava y media (D). En todas las figuras se agregó, con fines comparativos, la respuesta obtenida para



Ruido Rosa en el Experimento 1 (línea purpura), y la respuesta para un tono puro de 1KHz obtenida en el experimento 2 (línea azul).



**Figura 7.5:** Ruido Rosa filtrado centrado en 1KHz con diferentes anchos de banda. (A)– 1/12 octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 1KHz con un ancho de banda de 1/12 de octava; (B)– 1/6 octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 1KHz con un ancho de banda de 1/6 de octava; (C)– Octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 1KHz con un ancho de banda de una octava; (D)– Octava y Media,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 1KHz con un ancho de banda de una octava y media. En todas las figuras se agregó para comparación la respuesta para ruido rosa (línea purpura) y la respuesta para un tono de 1KHz (línea azul).

De manera similar a lo observado en la condición anterior, puede verse claramente como la respuesta se vuelve más precisa al aumentar el ancho de banda

del estímulo. Con estímulos de Ruido Rosa centrados en 1KHz y ancho de banda de 1/12 de octava (Figura 7.5A), los ajustes de la función lineal y de potencia muestran una pendiente y un valor de compresión por encima de los 0.5, valores de respuesta que si bien no son totalmente precisos, muestran una mejora con respecto a los valores obtenidos con estímulos centrados en 500HZ (función lineal:  $R^2 = 0,86$ ;  $a = 0,84 \pm 0,29$  y la pendiente  $s = 0,54 \pm 0,09$ ; función de potencia:  $R^2 = 0,88$ ;  $a = 1,35 \pm 0,20$  y el exponente  $b = 0,59 \pm 0,10$ ). Al aumentar el ancho de banda del estímulo a 1/6 de octava (Figura 7.5B), se observa una pequeña mejora en la respuesta, distinguible por el valor de la pendiente de la curva no así en el exponente de la función de potencias, aunque estos valores no son significativamente diferentes (función lineal:  $R^2 = 0,77$ ;  $a = 0,97 \pm 0,42$  y la pendiente  $s = 0,64 \pm 0,15$ ; función de potencia:  $R^2 = 0,76$ ;  $a = 1,56 \pm 0,30$  y el exponente  $b = 0,61 \pm 0,14$ ). La precisión en la respuesta mejora cuando el ancho de banda aumenta hasta una octava (Fig. 5-C). Esto se ve reflejado en un aumento en el valor de la pendiente ( $s$ ) y del exponente  $b$ , lo que refleja una disminución en la compresión de la curva, de la función de potencia (función lineal:  $R^2 = 0,98$ ;  $a = 0,49 \pm 0,41$  y la pendiente  $s = 0,70 \pm 0,14$ ; función de potencia:  $R^2 = 0,98$ ;  $a = 1,12 \pm 0,07$  y el exponente  $b = 0,78 \pm 0,04$ ). Cuando el ancho de banda llega a una octava y media (Figura 7.5D), la respuesta es similar a la obtenida en la condición anterior (función lineal:  $R^2 = 0,92$ ;  $a = 0,64 \pm 0,44$  y la pendiente  $s = 0,72 \pm 0,14$ ; función de potencia:  $R^2 = 0,95$ ;  $a = 1,34 \pm 0,15$  y el exponente  $b = 0,71 \pm 0,07$ ).

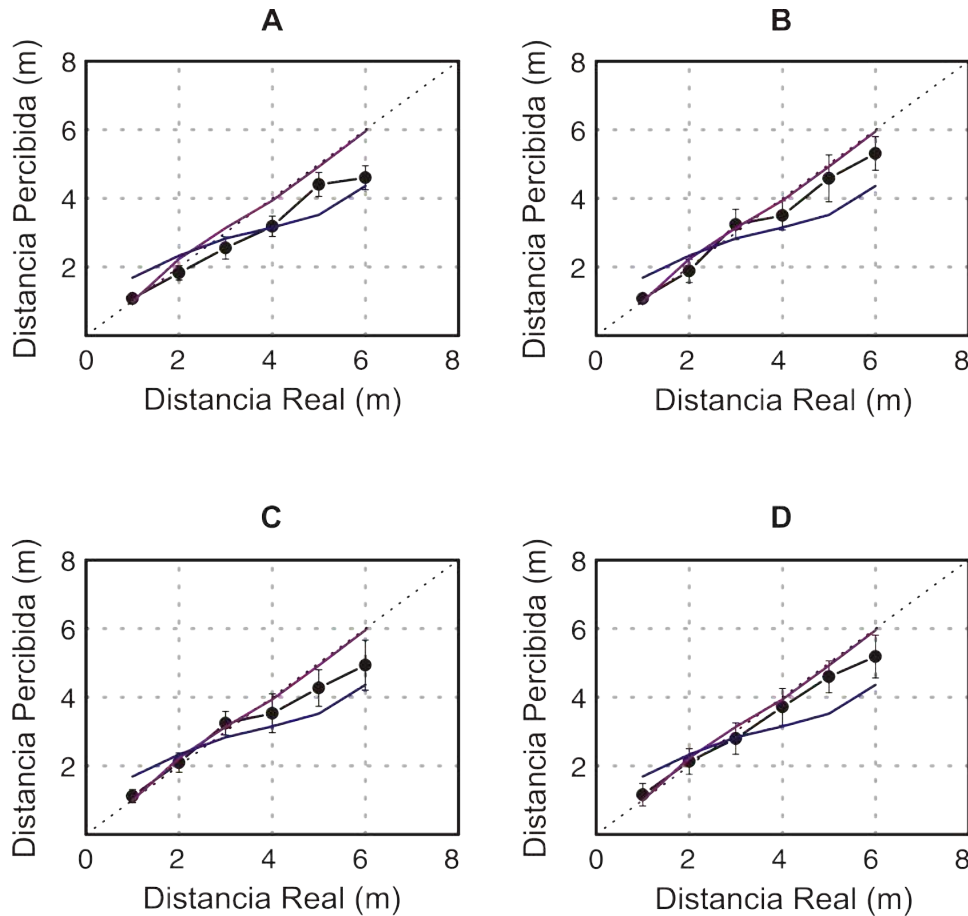
En este experimento se pudo observar como la respuesta de los oyentes es más precisa que en el experimento anterior cuando se utiliza estímulos de Ruido

Rosa centrados en 1KHz, y como esta precisión aumenta cuando el ancho de banda aumenta hasta por lo menos una octava.

### 7.12.3. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 2KHz

En la Figura 7.6 se puede observar las medias de los juicios subjetivos de distancia obtenidos como una función de la distancia física ( $n = 8$ ) utilizando como estímulos ruido rosa filtrado con un filtro centrado en 2KHz, con anchos de banda de: 1/12 de octava (A), 1/6 octava (B), una octava (C) y una octava y media (D). En todas las figuras se agregó, con fines comparativos, la respuesta para Ruido Rosa obtenida en el Experimento 1 (línea purpura), y la respuesta de los oyentes para un tono puro de 2KHz obtenida en el experimento 2 (línea azul).

A diferencia de la respuesta observada con los estímulos anteriores, se puede observar a simple vista una linealidad en la respuesta de los oyentes para estímulos de Ruido Rosa centrados en 2KHz. Cuando el ancho de banda fue de 1/12 de octava (Figura 7.6A), la respuesta de los oyentes ya superaba a la mejor respuesta obtenida con estímulos centrados por debajo de 1KHz (función lineal:  $R^2 = 0,98$ ;  $a = 0,33 \pm 0,14$  y la pendiente  $s = 0,74 \pm 0,05$ ; función de potencia:  $R^2 = 0,98$ ;  $a = 1,05 \pm 0,06$  y el exponente  $b = 0,83 \pm 0,04$ ). Cuando el ancho de banda aumenta a 1/6 de octava (Figura 7.6B), el resultado se acerca aun más a los valores obtenidos en el Experimento 1 (función lineal:  $R^2 = 0,98$ ;  $a = 0,22 \pm 0,15$  y la pendiente  $s = 0,85 \pm 0,07$ ; función de potencia:  $R^2 = 0,99$ ;  $a = 1,07 \pm 0,05$  y el exponente  $b = 0,88 \pm$



**Figura 7.6:** *Ruido Rosa filtrado centrado en 2KHz con diferentes anchos de banda. (A)– 1/12 octava  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 2KHz con un ancho de banda de 1/12 de octava; (B)– 1/6 octava  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 2KHz con un ancho de banda de 1/6 de octava; (C) – Octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 2KHz con un ancho de banda de una octava; (D) – Octava y Media,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 2KHz con un ancho de banda de una octava y media. En todas las figuras se agregó para comparación la respuesta para ruido rosa (línea púrpura) y la respuesta para un tono de 2KHz (línea azul).*

0,03).

La respuesta obtenida con anchos de banda de una octava (Figura 7.6C), (función lineal:  $R^2 = 0,96$ ;  $a = 0,38 \pm 0,24$  y la pendiente  $s = 0,81 \pm 0,09$ ; función de potencia:  $R^2 = 0,97$ ;  $a = 1,19 \pm 0,08$  y el exponente  $b = 0,81 \pm 0,05$ ) y una octava y media (Figura 7.6D), (función lineal:  $R^2 = 0,99$ ;  $a =$

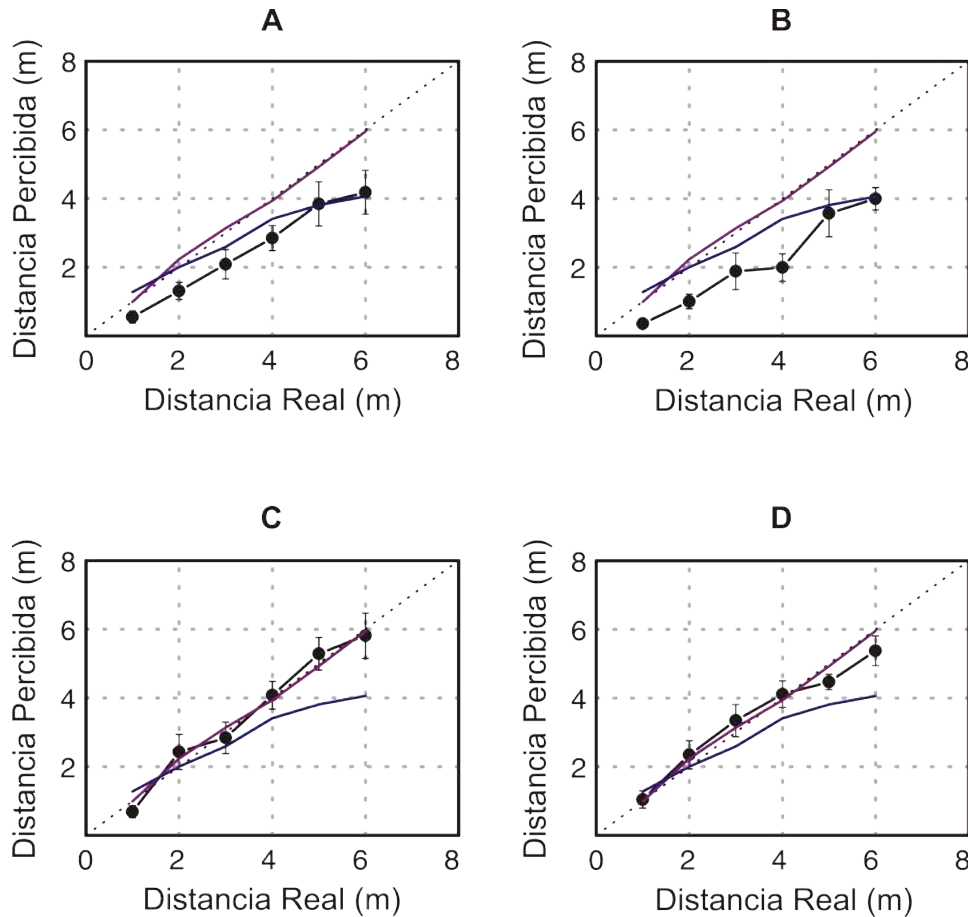
$0,38 \pm 0,34$  y la pendiente  $s = 0,82 \pm 0,10$ ; función de potencia:  $R^2 = 0,99$ ;  $a = 1,15 \pm 0,05$  y el exponente  $b = 0,84 \pm 0,02$ ) se mantienen parecidos a los logrados con 1/6 de octava. En este experimento se observa claramente que los sonidos cercanos a 2KHz necesitan un ancho de banda menor que sonidos centrados en frecuencias más graves (por debajo de 1KHz) para que un oyente pueda percibir de manera eficaz la distancia de la fuente que lo emite. Además estos resultados muestran que en este rango del espectro basta un ancho de banda de 1/6 de octava para obtener la mejor performance.

#### 7.12.4. Percepción de distancia y efecto de anchos de banda para estímulos centrados en 4KHz

En la Figura 7.7 se puede observar las medias de los juicios subjetivos de distancia obtenidos como una función de la distancia física ( $n = 8$ ) utilizando como estímulos ruido rosa filtrado con un filtro centrado en 4KHz, con anchos de banda de: 1/12 de octava (A), 1/6 octava (B), una octava (C) y una octava y media (D). En todas las figuras se agregó, con fines comparativos, la respuesta para Ruido Rosa obtenida en el Experimento 1 (línea purpura), y la respuesta de los oyentes para un tono puro de 4KHz obtenida en el experimento 2 (línea azul).

La respuesta a estos estímulos demuestra una diferencia importante con respecto al estímulo centrado en 2KHz, ya que hace falta un ancho de banda de una octava para lograr una respuesta cercana de las obtenidas en el Experimento 1. Cuando utilizamos un ancho de banda de 1/12 de octava (Figura 7.7A), los oyentes subestimaron todas las distancias testeadas. Sin embargo, tanto la pendiente como el exponente de la función de potencias fue elevado

(función lineal:  $R^2 = 0,99$ ;  $a = -0,20 \pm 0,22$  y la pendiente  $s = 0,76 \pm 0,08$ ; función de potencia:  $R^2 = 0,99$ ;  $a = 0,59 \pm 0,03$  y el exponente  $b = 1,12 \pm 0,04$ ).



**Figura 7.7:** *Ruido Rosa filtrado centrado en 4KHz con diferentes anchos de banda. (A) – 1/12 octava  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 4KHz con un ancho de banda de 1/12 de octava; (B) 1/6 octava  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 4KHz con un ancho de banda de 1/6 de octava; (C) – Octava,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 4KHz con un ancho de banda de una octava; (D) – Octava y Media,  $n=8$ : valor medio de los juicios de distancia como una función de la distancia real de la fuente sonora utilizando como estímulo ruido rosa centrado en 4KHz con un ancho de banda de una octava y media. En todas las figuras se agregó para comparación la respuesta para ruido rosa (línea purpura) y la respuesta para un tono de 4KHz (línea azul).*

Para un ancho de banda de 1/6 de octava (Figura 7.7B) se obtuvieron valores en los ajustes similares a los anteriores, sin que se observen diferencias

significativas (función lineal:  $R^2 = 0,98$ ;  $a = -0,35 \pm 0,11$  y la pendiente  $s = 0,70 \pm 0,05$ ; función de potencia:  $R^2 = 0,98$ ;  $a = 0,37 \pm 0,04$  y el exponente  $b = 1,32 \pm 0,07$ ). Cuando se utilizan anchos de banda más amplios, la respuesta comienza a semejarse a la obtenida en el Experimento 1. Para estímulos de Ruido Rosa centrados en 4KHz con un ancho de banda de una octava (Figura 7.7C), se logra el mejor resultado, ya que los valores de pendiente y compresión se acercan mucho a la unidad, mostrando una respuesta sumamente lineal (función lineal:  $R^2 = 0,98$ ;  $a = -0,36 \pm 0,22$  y la pendiente  $s = 1,09 \pm 0,08$ ; función de potencia:  $R^2 = 0,97$ ;  $a = 0,81 \pm 0,11$  y el exponente  $b = 1,14 \pm 0,09$ ). Cuando el ancho de banda llega a una octava y media, los valores obtenidos reflejan también juicios de distancia precisos, semejando a la respuesta con estímulos centrados en 2KHz (función lineal:  $R^2 = 0,97$ ;  $a = 0,38 \pm 0,28$  y la pendiente  $s = 0,84 \pm 0,07$ ; función de potencia:  $R^2 = 0,98$ ;  $a = 1,22 \pm 0,13$  y el exponente  $b = 0,82 \pm 0,07$ ).

### § 7.13 DISCUSIÓN

Los resultados obtenidos en los cuatro experimentos realizados muestran claramente que el contenido espectral del sonido afecta de manera sensible la PAD de oyentes humanos. Este fenómeno se aprecia con mayor contundencia cuando se compara la precisión de la respuesta obtenida ante un sonido espectralmente complejo como Ruido Blanco o Ruido Rosa con la obtenida con tonos puros. En el experimento 1 se puede observar una respuesta muy precisa en respuesta a estímulos que cubren todo el espectro de frecuencias, como lo son el Ruido Blanco (Figura 7.1A) y el Ruido Rosa (Figura 7.1B).

Este resultado concuerda con el experimento presentado en el capítulo anterior en donde se obtuvo una respuesta precisa al utilizar ruido blanco como estímulo en condiciones donde los oyentes conocían la sala donde se realizó el experimento (Capítulo 6 - Figura 6.6). Sin embargo, cuando utilizamos tonos puros, la respuesta fue notablemente menos precisa que la obtenida con estímulos de banda ancha. Como ya expusimos en la introducción de este capítulo, fenómenos similares ocurren en la percepción del ángulo de incidencia tanto en acimut como en elevación.

Este experimento nos permite observar además que la precisión para percibir la distancia de una fuente que emite tonos puros depende de la frecuencia del tono emitido. Por ejemplo, para tonos ubicados en la zona baja-media del espectro (500Hz – Figura 6.3A y 1KHz – Figura 6.3B) la respuesta fue sumamente imprecisa, lo cual se ve reflejado en la baja correlación obtenida tanto en el ajuste lineal como en la función de potencias. Cuando utilizamos tonos puros ubicados en la zona media-aguda del espectro (2KHz – Figura 6.3C y 4KHz – Figura 6.3D) observamos un aumento en la precisión de la respuesta lo cual se ve reflejado como un aumento en la correlación de las funciones lineal y de potencias. En estos casos la respuesta fue muy precisa en los primeros dos metros. Sin embargo, la distancia a la fuente fue subestimada en distancias mayores a dos metros. Este resultado indica que existe información relevante para percibir la distancia de una fuente en la zona media aguda del espectro ya que la respuesta obtenida con tonos puros ubicados en ese rango fue notablemente mas precisa que la obtenida en respuesta a tonos de 500 y 1000 Hz. Sin embargo, en ningún caso la respuesta fue tan precisa como la obtenida con sonidos de banda ancha indicándonos que no solamente influye



la frecuencia del estímulo sino su ancho de banda.

Los resultados obtenidos en el experimento 3 y 4 muestran claramente la importancia tanto de la frecuencia como del ancho de banda de los estímulos en la PAD. En el experimento 3 por ejemplo, se puede observar claramente como la respuesta depende de la zona del espectro de frecuencias donde se ubica la energía del estímulo. De manera similar a lo obtenido en el experimento 2 la precisión en la respuesta fue mayor para estímulos con información en el rango medio-agudo del espectro. Así, cuando se tiene información hasta 1KHz, se observa que la respuesta es precisa solo en los primeros metros testeados ya que cuando la fuente se encuentra a distancias mayores se produce una subestimación en la distancia de la fuente. Por el contrario, cuando se presentaron estímulos con información espectral por encima de 1KHz, la respuesta obtenida fue bastante precisa. Este resultado es comparable al obtenido en el experimento 2, ya que tanto para tonos puros como para estímulos con anchos de banda amplios la precisión en la respuesta mejora cuando se tiene información en la zona media-aguda del espectro. Esta relación queda claramente en evidencia en los resultados obtenidos en el experimento 4. Allí se observa claramente que la precisión en los juicios de distancia depende tanto del ancho de banda espectral del estímulo como del lugar en el espectro que este ocupa. Por ejemplo, para estímulos de Ruido Rosa centrados en 500Hz, con un ancho de banda espectral de  $1/12$  y  $1/6$  de octava ( Figura 6.5A y Figura 6.5B), se obtuvo una respuesta similar a la obtenida con un tono puro de 500Hz. Sin embargo, cuando el ancho de banda fue de una octava o una octava y media ( Figura 6.5C y Figura 6.5D) los oyentes detectaron la ubicación de la fuente de manera precisa en los primeros dos metros (las curvas fueron similares a las

obtenidas en el experimento anterior con un pasa-bajos de 20 a 1000 Hz). Es interesante lo que pasa en estos casos ya que la presencia de información en la zona grave del espectro, les sirvió a los oyentes para percibir correctamente la distancia de la fuente en 1 y 2 metros. Este resultado concuerda con lo reportado en el trabajo publicado por Kopko y Shinn-Cunningham, 2011 en donde se observa que un aumento de energía en la zona grave del espectro mejora la PAD de fuentes ubicadas en el campo cercano (hasta 1,7 metros). Este resultado avalaría la hipótesis de Békésy (1938) que postula que los cambios en la energía en la zona grave del espectro (efecto de proximidad) es un indicio importante de PAD. Sin embargo, como muestran los resultados obtenidos en este experimento, esta información no es suficiente para obtener respuestas precisas cuando la fuente se encuentra a distancias mayores a dos metros.

Cuando utilizamos estímulos de ruido rosa filtrado con filtros centrados en 1KHz de 1/12 y 1/6 de octava (Figura 6.6A y Figura 6.6B) observamos un comportamiento similar al que acabamos de describir ya que con anchos de banda pequeños la respuesta fue similar a la obtenida con un tono puro de 1KHz. Sin embargo, con anchos de banda de una octava y una octava y media (Figura 6.6C y Figura 6.6D) se obtuvieron respuestas mucho más precisas que las obtenidas con un tono puro de 1KHz o con estímulos centrados a 500 Hz de idéntico ancho de banda. De hecho, en estos casos, la respuesta fue similar a la obtenida en el experimento 1 con estímulos que contenían información en todo el espectro de frecuencias. Esto parece indicar que al aumentar el ancho de banda del estímulo centrado en 1KHz (por encima de una octava), se agregó información espectral (cercana a 2 kHz) importante para percibir correctamente la distancia de la fuente.

Este comportamiento parece ser reforzado cuando se evalúa la respuesta de los participantes con respecto a estímulos de ruido rosa filtrados con un filtro centrado en 2KHz. Lo interesante en esta zona del espectro es que aun con anchos de banda angostos ( $1/12$  y  $1/6$  de octava – Figura 6.7A y Figura 6.7B), la respuesta obtenida es notablemente mas precisa que la obtenida con un tono puro de 2 KHz. De hecho, aunque los anchos de banda utilizados fueron pequeños, la precisión de los oyentes para percibir la distancia a la fuente fue cercana a la obtenida con sonidos de banda ancha. Este aumento en la precisión se acentúa cuando el ancho de banda aumenta (una octava y una octava y media – Figura 6.7C y Figura 6.7D) aunque no es significativa ya que la respuesta a estímulos de ancho de banda angosta ya eran bastante precisos.

Los resultados obtenidos cuando utilizamos estímulos centrados en 4KHz fueron muy interesantes. Cuando utilizamos estímulos de ancho de banda pequeños ( $1/12$  y un  $1/6$  octava) la distancia a la fuente fue subestimada en todas las distancias de prueba. Sin embargo, los oyentes detectaron claramente los cambios en la distancia de la fuente. De hecho, tanto el ajuste lineal como el de potencias muestran que la respuesta tiene una pendiente y una compresión similar a la obtenida con ruido blanco y ruido rosa. Nosotros creemos que este resultado puede estar relacionado las características de la sala donde se realizó el experimento ya que las paredes de esta sala están cubiertas con material absorbente el cual actúa principalmente sobre frecuencias agudas. De este modo los estímulos de banda angosta centrados en 4 KHz utilizados aquí fueron absorbidos casi en su totalidad por las paredes de la sala por lo que al oyente llegó principalmente la señal directa y muy poca información reverberante. Vimos en la introducción (Capítulo 3) abundante evidencia que

indica que los oyentes tienden a percibir distancias mayores en presencia de reverberación. De este modo, podríamos pensar que la falta de reverberación hizo que los oyentes subestimaran la distancia real de la fuente sonora. Cuando el ancho de banda fue de una octava y una octava y media la respuesta obtenida fue muy similar a las obtenidas con Ruido Rosa y Ruido Blanco. Nosotros creemos que esta mejoría se debe a que estos estímulos contienen información espectral proveniente de la zona media del espectro (recordemos que el estímulo está centrado en 4 KHz y que al aumentar el ancho de banda se ganan tanto frecuencias agudas como graves) las cuales fueron importantes para la correcta estimación de la distancia a la fuente.

Los resultados aquí expuestos muestran claramente que el contenido espectral de un sonido puede afectar sensiblemente la PAD. Por ejemplo pudimos ver que es mucho más fácil ubicar en profundidad una fuente que emite sonidos de banda ancha que una que emite tonos puros y que además el ancho de banda mínimo necesario para percibir correctamente la distancia a la fuente depende de la zona del espectro que ocupe el estímulo. Estos resultados tienen también una connotación biológica ya que rara vez en el mundo natural estamos expuestos a tonos puros y por el contrario estamos expuestos a sonidos de banda ancha como el crujir de hojas, el sonido del agua, el sonido de pasos, de una rama quebrándose etc. Además, un sonido de banda ancha es susceptible de ser afectado en un gran rango de frecuencias lo que seguramente genera información relevante para la localización de fuentes sonoras mientras que en un tono puro solo puede ser afectada su amplitud.

## REFERENCIAS BIBLIOGRÁFICAS

- BASSO, G.: *Percepción Auditiva*. Editorial de la Universidad Nacional de Quilmes, Quilmes, Argentina, 2006.
- BATTEAU, D.W.: «The Role of the Pinna in Human Localization». *Proc. R. Soc. Lond.*, 1967, **168**, pp. 158–180.
- BÉKÉSY, G.V.: «Über die Entstehung der Entfernungsempfindung beim Hören (On the origin of the sensation of distance in hearing)». *Akustische Zeitschrift*, 1938, **3**, pp. 21–31.
- BLAUERT, J.: *Spatial hearing. The psychophysics of human sound localization*. Massachusetts, USA, 1997.
- BLOOM, P.J.: «Determination of monaural sensitivity changes due to the pinna by use of minimum-audible-field measurements in the lateral vertical plane». *J. Acoust. Soc. Am.*, 1977, **61(3)**, pp. 820–828.
- BUTLER, R.A.: «The bandwidth effect on monaural and binaural localization». *Hearing Research*, 1986, **21**, pp. 67–73.
- BUTLER, R.A. y PLANERT, N.: «The influence of stimulus bandwidth on localization of sound in space». *Attention, Perception and Psychophysics*, 1976, **19(1)**, pp. 103–108.
- CARLSON, V.R.: «Instructions and perceptual constancy judgment». *En Stability and Constancy in Visual Perception: Mechanisms and Processes Ed. W Epstein*, 1977, pp. 217–254.
- COLEMAN, P.D.: «Failure to localize the source distance of an unfamiliar sound». *J. Acoust. Soc. Am.*, 1962, **34**, pp. 345–346.
- : «Dual role of frequency spectrum in determination of auditory distance». *J. Acoust. Soc. Am.*, 1968, **44**, pp. 631–634.
- CONDAMINES, N.: «Le Reperage Directionnel Des Sons En Ecoute Mono et Biauriculaire». *5eme Cong. Int. D'Acoust. Liege, Belgium*, 1965.
- HEBRANK, J. y WRIGHT, D.: «Spectral cues used in the localization of sound sources on the median plane». *J. Acoust. Soc. Am.*, 1974, **56(6)**, pp. 1829–1834.
- HOFMAN, P.M. y VANOPSTAL, A.J.: «Spectro-temporal factors in two-dimensional human sound localization». *J. Acoust. Soc. Am.*, 1998, **103(5)**, pp. 2634–2648.
- INGARD, U.: «A Review of the Influence of Meteorological Conditions on Sound Propagation». *J. Acoust. Soc. Am.*, 1953, **25**, pp. 405–411.
- IWAYA Y., OTANI M., MAGARIYACHI T. y Y., SUZUKI: «Role of 4-8 kHz band component for wideband noise localization in median plane». *J. Acoust. Soc. Am.*, 2003, **131(4)**, pp. 3216–3216.
- JEFFRESS, L.A.: «Binaural signal detection - Vector theory (Vector correlation theory and neural mechanisms of binaural signal detection in human auditory system)». *Foundations of modern auditory theory*, 1972, **2**, pp. 351–368.
- JIN C., CARLILE S., CORDEROY A. y VANSCHAIK, A.: «Contrasting monaural and interaural spectral cues for human sound localization». *J. Acoust. Soc. Am.*, 2004, **115(6)**, pp. 3124–3141.
- KING, R.B. y OLDFIELD, S.R.: «The impact of signal bandwidth on auditory localization: Implications for the design of three-dimensional audio displays». *Human Factors*, 1997, **39**, pp. 287–295.

- KOPCO, N. y SHINN-CUNNINGHAM, B.G.: «Effect of stimulus spectrum on distance perception for nearby sources». *J. Acoust. Soc. Am.*, 2011, **130(3)**, pp. 1530–1541.
- KOPCO N., PUSZTA M. RACEK M., CELJUSKA D. y M., SARNOVSKY: «Effect of spectral content and learning on auditory distance perception». *En Proceedings of 2nd Slovak-Hungarian Joint Symposium on Applied Machine Intelligence*, 2012.
- LANGENDIJK, E.H.A y BRONKHORST, A.W.: «Contribution of spectral cues to human sound localization». *J. Acoust. Soc. Am.*, 2002, **112**, pp. 1583–1596.
- LICKLIDER J.C.R., WEBSTER J.C. y J.M., HEDLUM: «On the Frequency Limits of Binaural Beats». *J. Acoust. Soc. Am.*, 1950, **22**, pp. 468–473.
- LITTLE A.D., MERSHON D.H. y P.H., COX: «Spectral content as a cue to perceived auditory distance». *Perception*, 1992, **21**, pp. 405–416.
- MEIJER, P.B.L: «An Experimental System for Auditory Image Representations». *IEEE Trans. on Biomedical Engineering*, 1992, **39(2)**, pp. 112–121.
- MIDDLEBROOKS, J.C.: «Narrow-band sound localization related to external ear acoustics». *J. Acoust. Soc. Am.*, 1992, **92**, pp. 2607–2624.
- MIDDLEBROOKS, J.C. y GREEN, D.M.: «Sound Localization by Human Listeners». *Annual Review of Psychology*, 1991, **42**, pp. 135–159.
- MILLS, A.W.: «On the Minimum Audible Angle». *J. Acoust. Soc. Am.*, 1958, **30(4)**, pp. 237–246.
- : «Lateralization of high-frequency tones». *J. Acoust. Soc. Am.*, 1960, **32(1)**, pp. 132–134.
- : «Auditory localization». *En Tobias JV (Ed.) Foundations of Modern Auditory Theory*, 1972, **2**, pp. 303–348.
- MOORE, B.C.J.: *An introduction to the psychology of hearing*. Esmerald, Bingley, 2012.
- NYBORG, W. y MINTZER, D.: «Review of sound propagation in the lower atmosphere». *U.S. Air Force WADA Tech. Rept.*, 1955, pp. 54–602.
- PRATT, C.C.: «The spatial character of high and low tones». *Journal of Experimental Psychology*, 1930, **13**, pp. 278–285.
- RAYLEIGH, L.: «On our perception of sound direction». *Philosophical Magazine Series 6*, 1907, **13(74)**, pp. 214–232.
- ROFLER, S.K. y BUTLER, R.A.: «Factors That Influence the Localization of Sound in the Vertical Plane». *J. Acoust. Soc. Am.*, 1968a, **43(6)**, pp. 1255–1259.
- : «Localization of tonal stimuli in the vertical plane». *J. Acoust. Soc. Am.*, 1968b, **43**, pp. 1260–1266.
- SLATTERY, W.H.I. y MIDDLEBROOKS, J.C: «Monaural sound localization: Acute versus chronic impairment». *Hear. Res.*, 1994, **75**, pp. 38–46.
- SODNIK J., BOBOJEVIC G., SUSNIK R. y S., TOMAZIC: «The directional resolution of virtual sound sources in humans». *Electrotechnical Review*, 2004, **71(3)**, pp. 121–127.
- SODNIK J., STULAR M., SUSNIK R. y S., TOMAZIC: «Spatial sound resolution of an interpolated HRIR library». *Applied Acoustics*, 2005, **66(11)**, pp. 1219–1234.
- STERN, R.M. y TRAHOTIS, C.: «Models of binaural interaction». *En Handbook of Perception and Cognition, Hearing, B. C. J. Moore., ed., Academic Press*, 1995, **6**, pp. 347–386.

- STEVENS, S.S. y NEWMAN, E.B.: «The localization of actual sources of sound». *The American Journal of Psychology*, 1936, **48(2)**, pp. 297–306.
- THURLOW, W.R. y RUNGE, P.S.: «Effect of Induced Head Movement on the Localization of the Direction of Sounds». *J. Acoust. Soc. Am.*, 1967, **42**, pp. 480–488.
- TRIMBLE, O.C.: «Localization of sound in the anterior, anterior, posterior and vertical dimensions of auditory space». *British Journal of Psychology*, 1934, **24(3)**, pp. 320–334.
- WATKINS, A.: «Psychoacoustical aspects of synthesized vertical local cues». *J. Acoust. Soc. Am.*, 1978, **63(4)**, pp. 1152–1165.
- ZIBERNA, G. y ZAZULA, D.: «Computer generation of 3D sound in virtual spaces». *Electrotechnical Review*, 2003, **70(3)**, pp. 96–102.
- ZWISLOCKI, J. y FELDMAN, R.D.: «Just noticeable differences in dichotic phase». *J. Acoust. Soc. Am.*, 1956, **28**, pp. 860–864.





## - Capítulo 8 -

---

---

### Desarrollo de aplicaciones informáticas para la espacialización del sonido

---

#### § 8.1 LA ESPACIALIDAD EN LA MÚSICA ELECTROACÚSTICA Y EL DISEÑO DE PROGRAMAS INFORMÁTICOS PARA SU TRATAMIENTO

Como vimos en el desarrollo de esta tesis, en la música electroacústica, la calidad espacial de los sonidos juega un rol relevante dentro del marco estructural de las composiciones. A causa de esto, el análisis de la influencia estética de la calidad espacial de los sonidos y su relación con otras organizaciones sonoras es una parte esencial de la teoría de este estilo musical (ver Capítulo 4, enfoques analíticos de la espacialidad en la música electroacústica). Una trayectoria sonora puede determinar claramente el fin de un motivo musical, o la articulación formal de una composición musical, solo cambiando su dirección de forma repentina o parando su movimiento abruptamente y desarrollando una sonoridad particular desde una posición estable (para una revisión de diferentes formas de movimiento ver Stockhausen, 1952-60). Debido a esto surge en el compositor la necesidad de tener un control cada vez más fino sobre la calidad espacial del sonido a través de herramientas tecnológicas que

permitan la coordinación temporal precisa entre múltiples señales de audio y sus trayectorias y posiciones en un espacio virtual.

La espacialización del sonido puede desarrollarse a partir de muchas tecnologías y áreas científicas mutuamente relacionadas. Dentro de estas áreas puede mencionarse: la acústica (particularmente acústica de salas), psicoacústica (específicamente escucha espacial), procesamiento digital de audio e ingeniería de audio.

Particularmente, y junto con el desarrollo de la música por computadoras, mucho de los resultados de las investigaciones científicas en este campo fueron aplicados en la creación de aplicaciones de programas de computadora con el propósito de proveer herramientas concretas para producción sonora y musical a los compositores e ingenieros de audio. Los trabajos de Chowning (1971) y Moore (1983, 1989 y 1990) pueden ser considerados como una referencia inicial y valiosa las cuales fueron seguidos de desarrollos significativos. Entre ellos pueden ser mencionados algunos desarrollos de Kendall et al (1989), Furse (1995, 1999), López-Lezcano (1992), Karpen (1998), Pulkki (2001) y Varga (2000).

En el presente, existe una cantidad creciente de programas de espacialización, la mayoría de ellos como ambientes de alto nivel alcanzados por la combinación de unidades de software de propósito general (como Csound, SuperCollider o Pure Data, para solo mencionar algunos). Dentro de estos entornos se pueden mencionar AUDIOSCAPE (Wozniowski et al 2007), y Bin-Ambi (Musil 2007). Repasando, el tratamiento espacial del sonido involucra tres aspectos básicos:

1-Entorno: el espacio físico dentro del cual las fuentes sonoras están localiza-

das.

2-Localización – Movimiento: los puntos del espacio físico donde las fuentes sonoras están localizadas, o las trayectorias que estas siguen en el espacio donde se están moviendo.

3-Directividad: el patrón de extensión de la energía acústica de las fuentes sonoras y de acuerdo a su naturaleza y orientación con respecto al punto de escucha (ver por ejemplo, Otondo et al 2002).

Los programas anteriormente nombrados, fueron diseñados para desarrollar los puntos 1 y 2 (entorno, localización y movimiento). Solo algunos de estos programas (Moore 1983, 1990; Furse 1999), sin embargo, pueden llevar a cabo la simulación de directividad de la fuente sonora.

Por otra parte, desde el punto de vista del diseño, se pueden distinguir tres secciones básicas entre los diferentes programas de espacialización:

1-Un ajuste general de los parámetros de configuración relacionados a la sala o ambiente virtual, como por ejemplo: tamaño, forma, posición del oyente, cantidad de eco, calidad de la reverberación densa, etc.

2-Generación de trayectorias sonoras y/o posición para las fuentes acústicas virtuales.

3-Bloques de procesamiento digital de las señales (DSP Block – Motor de Audio) dedicados al procesamiento de las señales de audio digital a ser espacializadas de acuerdo con los datos obtenidos de las secciones 1A y 2A y las técnicas de espacialización que se utilizaran (Ambisonics, Intensity Panning, VBAP, etc.).

A partir de estos conceptos se procederá a la presentación de tres programas informáticos que trabajan de forma diferente la espacialidad del sonido. Estos fueron desarrollados en el marco de esta tesis por Esteban Calcagno bajo la supervisión de Oscar Pablo Di Liscia, como parte del proyecto de investigación “Desarrollo de software para la espacialización de sonido y música” (Programa de investigación Teatro Acústico, Universidad Nacional de Quilmes).

## § 8.2 GTKONVOLVE: UNA APLICACIÓN PARA REALIZAR CONVOLUCIÓN MÚLTIPLE

Como vimos en el Capítulo 5 (Técnicas y sistemas de espacialización) una de las formas en que el compositor electroacústico puede utilizar el espacio dentro de sus composiciones es simulando las características acústicas de un ambiente determinado, en general mediante las características de reverberación que posee el espacio elegido.

Como también vimos en ese capítulo, una de las formas de capturar las características acústicas de los ambientes es mediante la medición de su respuesta impulso (para un método eficaz para obtener la respuesta impulso ver Farina, 2000). La respuesta impulso lleva impregnada las reflexiones en las paredes y superficies del ambiente, por lo que teniendo esta medición se tiene una huella acústica del espacio. Una vez que se obtiene la respuesta impulso esta puede ser utilizada para traspasar sus características temporales y espectrales a otros sonidos, que pueden tener o no tener una reverberación propia (por ejemplo, los estímulos grabados en una cámara anecoica). Este procedimiento se realiza mediante la técnica de convolución, la cual revisaremos más adelante en este

capítulo.

En este capítulo se revisará Gtkonvolve, un programa informático, el cual funciona tanto en sistemas operativos GNU/Linux como MSWindows, con una interfaz gráfica de usuario o GUI (del inglés, Graphics User Interface), desarrollado mediante el lenguaje de programación C. Este programa sirve para configurar una matriz de convolución mediante el cargado de un archivo de audio, al cual queremos impregnar las características de reverberación de un ambiente y la respuesta impulso de dicho ambiente. Lo interesante, es que podemos cargar archivos de sonido multicanal y archivos de respuestas a impulso multicanal y combinar cada uno de los canales para crear un archivo de audio nuevo, que contenga la combinación de canales elegida. Más adelante veremos un ejemplo de la utilidad que puede tener estas posibilidades que permite Gtkonvolve.

A continuación se explicará de qué se trata la técnica de convolución, como se utiliza dentro del ámbito de la informática y específicamente en el procesamiento de señales de audio. Luego de esta explicación se procederá a realizar un análisis de la herramienta Gtkonvolve, sus características básicas de funcionamiento, y algunos ejemplos aplicados.

### 8.2.1. Convolución y procesamiento digital de señales de audio

El efecto que se intenta obtener al realizar el proceso de convolución entre dos señales de audio, es que el espectro de la señal resultante sea equivalente a la multiplicación de los espectros de estas señales. O sea, se intenta filtrar una señal con otra para traspasar las características espectrales de una a la otra (Di Liscia, 2005).

La convolución no es ni más ni menos que un proceso matemático, y dentro del audio digital es utilizada frecuentemente para simular el efecto de reverberación. A este efecto se lo denomina reverberación por convolución.

Si se requiere utilizar el proceso de convolución para impregnar las características acústicas de un ambiente a un sonido determinado, se debe utilizar la respuesta a impulso de la misma.

Como indica Cetta (2010), el efecto es sumamente preciso ya que la convolución aplica el comportamiento de la respuesta a impulso a cada una de las muestras de la señal de entrada. En definitiva, la operación de convolución establece una copia del espectro de una de las formas de onda (en este caso la respuesta a impulso), centrada sobre cada uno de los componentes de frecuencia del espectro de la otra onda (Di Liscia, 2005).

Dentro del procesamiento de audio digital el proceso de convolución se puede aplicar de diversas maneras. La forma más directa es la denominada convolución lineal, la cual presenta algunos inconvenientes relacionados con los tiempos de procesamiento. Una opción más eficiente es la denominada convolución rápida.

### 8.2.2. Convolución lineal

El procedimiento de convolución lineal, o aperiódica, consta de tomar una copia de una de las señales y multiplicar cada una de sus muestras por la primera muestra de la otra señal. A este primer resultado se le suma, una muestra más tarde, el mismo procedimiento pero realizado a la segunda muestra de la segunda señal. Este proceso se realiza hasta que se acaban las muestras de la segunda señal (Di Liscia, 2005). De forma obvia el resultado final será una

señal que contiene una cantidad de muestras igual a la suma de las muestras de cada una de las señales convolucionadas, menos una muestra.

Este procedimiento puede resumirse en la siguiente ecuación obtenida de Cetta (2010):

$$r(n) = f(n) * s(n) = \sum_{m=0}^n f(m)s(n - m)$$

Donde  $r(n)$  es la señal resultante y  $f(n)$  y  $s(n)$  son las señales a las que se aplica el proceso de convolución.

Este procedimiento, a pesar de ser el más directo, presenta como inconveniente principal el hecho de que se requieren realizar demasiadas operaciones, acusando un tiempo de procesamiento bastante importante y además que es sumamente dependiente de la cantidad de muestras que presentan las señales a las que se realiza el proceso.

En general, dentro del procesamiento digital de señales de audio se utiliza otro tipo de procedimiento al cual se denomina convolución rápida.

### 8.2.3. Convolución Rápida

La convolución rápida parte de la noción de que si multiplicamos los espectros de dos señales se logra el mismo resultado que el obtenido con la convolución lineal. Esto conlleva el beneficio de que en el caso de multiplicar los espectros, se reduce el tiempo de procesamiento, obteniendo un procedimiento más eficiente.

Como explica Cetta (2010), es posible obtener el espectro de una señal (dominio de frecuencias), la cual está representada en el dominio del tiempo,

mediante la Transformada Discreta de Fourier o DFT (del inglés, Discrete Fourier Transform). También es posible realizar este procedimiento, de una forma aun más veloz, mediante la aplicación de un algoritmo de la Transformada Rápida de Fourier o FFT (del inglés, Fast Fourier Transform). Este último método solo puede lograrse cuando el número de muestras es potencia de 2. De no suceder esto, se puede agregar a la señal tantas muestras con valor “0” como sea necesario. Como veremos más adelante, Gtikonvolve utiliza un algoritmo de FFT para realizar convolución rápida.

Una vez que se obtienen los espectros de ambas señales, se procede a realizar la multiplicación compleja entre ambos espectros y luego aplicar, al resultado de esta multiplicación de espectros, la Transformada Rápida Inversa de Fourier o IFFT (del inglés, Inverse Fast Fourier Transform) para volver al dominio del tiempo, y obtener así el mismo resultado que con el método de convolución lineal, pero de una forma más efectiva.

En Tommasini et al (2010) se puede obtener una comparación de performance entre el método de convolución lineal (o discreta) y el de convolución rápida mediante FFT. Allí se puede observar como el procesamiento puede ser hasta un 80 % más rápido utilizando la multiplicación compleja de los espectros de las señales.

Igualmente, Cetta (2010) aclara que si bien el costo del procesamiento se reduce considerablemente con este método, no es posible su aplicación en tiempo real debido a que se debería conocer de antemano la señal de entrada en su totalidad para realizar la operación en un solo paso. Para subsanar este problema existe un método denominada superposición y suma, mejor conocido como Overlap/Add.



#### 8.2.4. Método Overlap/add

Este método mejora la performance de los anteriores mediante la consideración de la señal por bloques. De esta manera se puede multiplicar directamente el espectro de cada bloque de la señal de entrada contra el espectro de la respuesta a impulso, y de esta manera acelerar el procesamiento. Este método provee una manera eficiente de calcular la convolución de una señal muy extensa (Tommasini et al, 2010). El nombre de Overlap/Add proviene del hecho de que cada bloque de salida es solapado con el siguiente y luego sumado (Oppenheim y Schaffer, 1975; Hayes, 1999).

Un ejemplo desarrollado por Cetta (2010) explica el desarrollo de esta técnica de una forma muy sencilla:

Si tenemos una respuesta a impulso con 128 muestras, esta se procesa con un bloque de la señal de entrada que también tenga 128 muestras, mediante convolución rápida, o sea la multiplicación compleja de sus espectros, obteniendo una señal de salida de 255 muestras (el total de muestras de la suma de las muestras de cada señal menos una muestra). Las primeras 128 muestras de esta señal se derivan al archivo de salida, mientras que el resto se almacenan para superponerse en la siguiente vuelta y luego derivarse al archivo de salida.

Este método de Overlap/Add es el que se utiliza en el funcionamiento de Gtconvolve, aunque no está aplicado al procesamiento en tiempo real.

#### 8.2.5. Gtconvolve

Gtconvolve fue desarrollado a partir del programa Convolve y fue escrito en su totalidad en el lenguaje C. El programa Convolve fue desarrollado por

Pablo Di Liscia para realizar convolución entre un archivo de audio y una respuesta a impulso. Dicho programa podía ser ejecutado desde una terminal de consola, tanto en sistemas operativos GNU/Linux como en MSWindows, y utilizaba la librería sndlib (Bill Schottstaedt, <https://ccrma.stanford.edu/software/snd/snd/sndlib.html>) para leer y procesar archivos de audio. Como su documentación lo indica, el bucle principal de Convolve está basado a su vez, en el programa Convolvel de Mark Dolson (CARLs, kit de programas para DSP, ver también CCRMA:[https://ccrma.stanford.edu/~jos/sasp/Example\\_Overlap\\_Add\\_Convolution.html](https://ccrma.stanford.edu/~jos/sasp/Example_Overlap_Add_Convolution.html)).

Al igual que este programa, Convolve utiliza una adaptación de las rutinas de FFT e IFFT de las librerías IEEE. A partir del estudio del programa Convolve, se trabajó en una interfaz gráfica que permitiera combinar un archivo de audio con otro archivo que contuviera una colección de respuestas a impulso. De esta manera se podría obtener un nuevo archivo, de tal manera que cada uno de los canales contenga el archivo de audio original procesado con diferentes respuestas a impulso. La génesis de este programa surgió de la idea de grabar la respuesta a impulso de una sala en diferentes posiciones de la misma, por ejemplo en 8 posiciones diferentes, y guardar cada una de estas respuestas a impulso en un canal diferente de un mismo archivo. Una vez obtenida la grabación de estas respuestas, mediante Gtkonvolve se puede convolucionar un archivo de audio con cada una de ellas y obtener, siguiendo con el ejemplo, un nuevo archivo con 8 canales, el cual contendría cada una de las convoluciones. Al reproducir estos 8 canales (en un sistema octofónico, por ejemplo) respetando las distancias y la localización a las que se grabaron las respuestas a impulso, se podría tener una reproducción muy parecida al

evento inicial. Así, Gtkonvolve se convierte en una herramienta para emular la respuesta acústica de un espacio en múltiples ubicaciones. Por supuesto que esta es la idea base del programa, lo que no inhabilita utilizarlo para otros tipos de procesamiento entre dos archivos de audio.

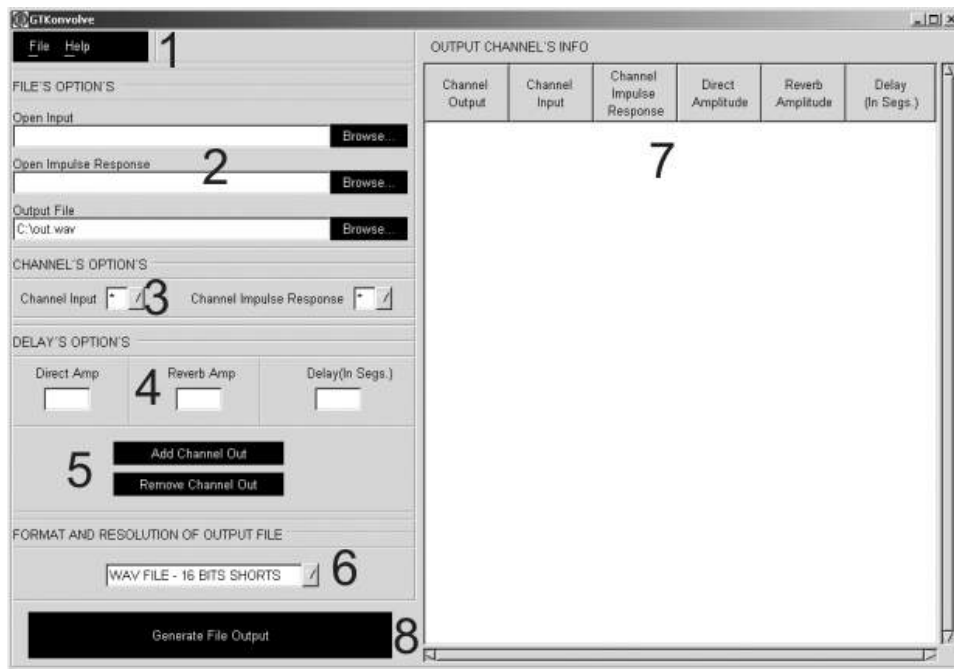
### 8.2.6. Desarrollo de la interfaz

La interfaz gráfica de usuario de Gtkonvolve fue diseñada y programada en lenguaje C, utilizando la librería de código abierto para el desarrollo de interfaces gráficas GTK (gtk-2.0, <http://www.gtk.org/>, para sistemas GNU/Linux y winGTK, <http://wingtk.sourceforge.net/>, para sistemas MSWindows). Una imagen del panel principal de Gtkonvolve puede apreciarse mediante la Figura 8.1.

### 8.2.7. Lógica de funcionamiento

Como se observa en la figura 2, Gtkonvolve se divide en 8 secciones. En la sección 1, se puede observar el menú principal. Este menú, mediante la pestaña File, permite crear una nueva sesión de trabajo, guardar los datos de la sesión actual en un archivo de texto, abrir una sesión pre-guardada o salir de la aplicación. Mediante la pestaña Help, se puede tener información sobre la aplicación y acceso al archivo de ayuda.

La sección 2 muestra donde se configura las opciones de los archivos de audio y respuesta a impulso que vamos a cargar, y en que ubicación será guardado el archivo de salida que generemos. En la sección 3, tenemos las opciones de



**Figura 8.1:** Frontend de Gtkonvolve. 1) Menú principal; 2) ingreso de la ruta del archivo de audio, respuesta a impulso y archivo de salida; 3) elección del canal del archivo de audio y el canal del archivo de las respuestas a impulso a combinar; 4) Configuración de la amplitud del sonido directo, de la reverberación y si se quiere agregar un delay entre ambos; 5) agregar o quitar combinaciones; 6) elección del formato y resolución del archivo de salida; 7) Información de las combinaciones de las que estará compuesto el archivo de salida; 8) Generación del archivo de salida

canales, aquí se irán configurando la combinación de canales entre el archivo de audio y la respuesta a impulso. La sección 4 es donde se configurará la amplitud del sonido directo, la del sonido reverberante y el tiempo de retraso de la señal procesada. En la sección 5 están las opciones de agregar o quitar un canal del archivo de salida, los cuales estarán configurados según las opciones que se dispusieron en 3 y 4. En la sección 6 podemos determinar cuál será el formato y la resolución que tendrá el archivo de salida. Se puede elegir entre WAV de 16 bits, WAV de 32 bits, AIF de 16 bits o SND de 16 bits. La sección 7 corresponde a una matriz visual donde podremos observar cómo se va configurando cada uno de los canales del archivo de salida, y además, como veremos en un ejemplo más adelante, se puede realizar retoques sobre lo previamente seleccionado. De esta manera se puede ir observando y arreglando

cualquier configuración que se haya elegido. Por último la sección 8 muestra el botón que se utiliza para generar el archivo de salida, una vez que se haya realizado la configuración correspondiente.

Se procederá a comentar un ejemplo básico de procesamiento utilizando Gtkonvolve.

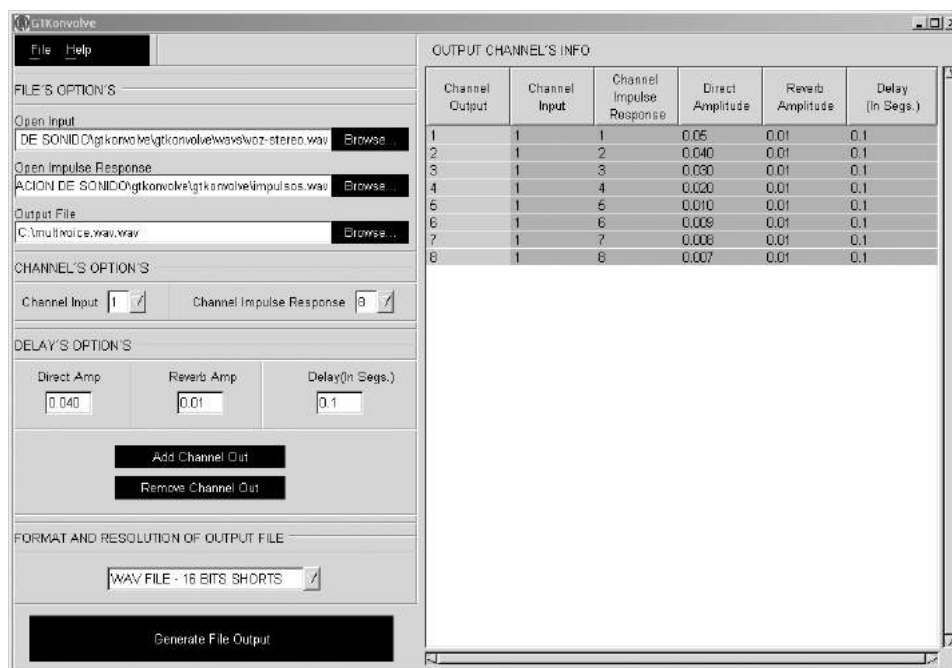
### 8.2.8. Ejemplo básico de funcionamiento

Como vimos en el Capítulo 4 (Enfoques analíticos de la espacialidad en la música electroacústica), Kendall (2010) desarrolló una metodología para analizar la estructura espacial de una obra electroacústica. Una de las facetas interesantes que propone Kendall es lo que el mismo denomina “juegos con el agrupamiento perceptivo”. Mediante la definición de Fuentes Conceptuales e Imágenes Fuente, se pueden dar diferentes configuraciones que pueden producir un desfase perceptual entre lo que entendemos y lo que escuchamos. Uno de estos juegos es cuando se tiene una sola Fuente Conceptual (como por ejemplo una voz humana, sumamente reconocible) y múltiples Imágenes Fuente, o sea múltiples fuentes sonora localizadas en el espacio que contienen la misma información sonora. En definitiva se está hablando de una voz que se escucha en diferentes espacios o localizaciones al mismo tiempo.

Este ejemplo puede generarse de manera muy efectiva con Gtkonvolve.

Si por ejemplo tenemos grabadas 8 respuestas a impulso de 8 posiciones diferentes de un ambiente, podemos tomar una voz determinada y procesarla con Gtkonvolve de tal manera que la voz sea convolucionada contra cada respuesta a impulso. De esta manera tendríamos un archivo de salida, que luego puede ser reproducido por un sistema octofónico produciendo este juego perceptual.

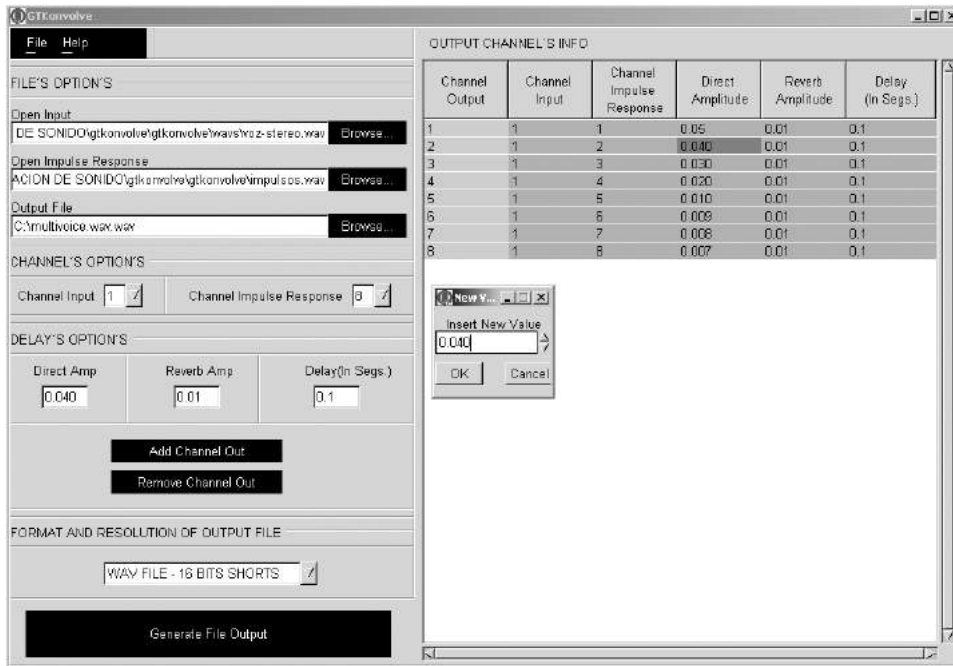
Si además se cambia la relación entre el sonido directo y el reverberante y el retraso de cada canal, podríamos simular diferentes distancias en cada uno de los canales (recordemos que la relación entre el sonido directo y el reverberante es una de las pistas más importantes en la percepción auditiva de distancia). La Figura 8.2 muestra como quedarían conformados los diferentes canales de la señal de salida en este ejemplo en particular:



**Figura 8.2:** Ejemplo de Gtikonvolve para generar el juego perceptivo indicado por Kendall (2010): Una fuente Conceptual Vs. múltiples Imágenes Fuente

De querer cambiar algún parámetro de los canales del archivo de salida, solamente se debe hacer doble click en el casillero al cual se desea cambiar el valor, como se observa en la Figura 8.3.

Además, una vez que Gtikonvolve procesa los datos que ingresamos (al oprimir el botón “Generate File Output”), este devuelve un reporte que indica como fue el resultado canal por canal. Este reporte es importante debido a que nos indicará si algún canal se fue del rango de amplitud (Figura 8.4). Por último, puede obtenerse una copia del archivo de salida del ejemplo

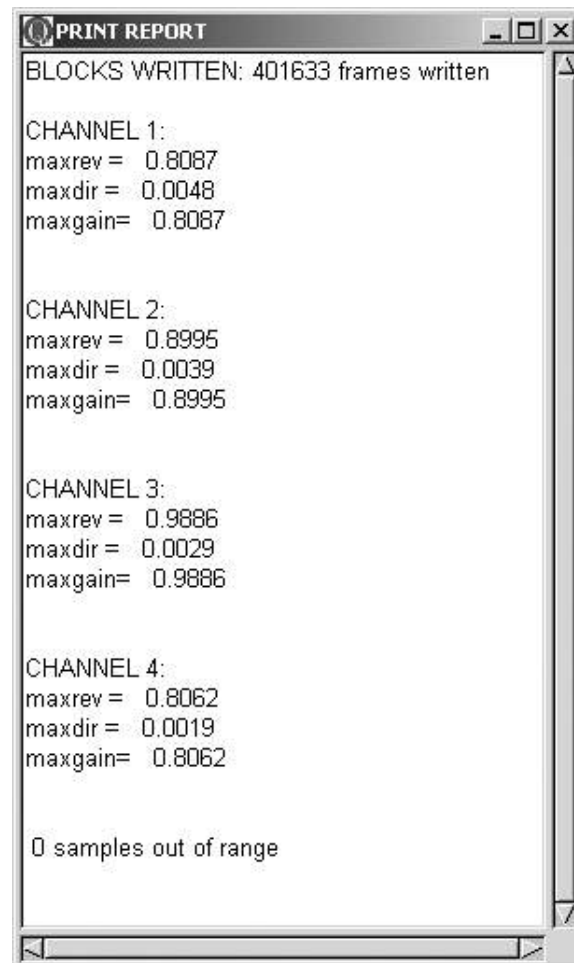


**Figura 8.3:** Ejemplo de cómo cambiar el valor de los casilleros previamente configurados.

aquí planteado en <http://www.estebancalcagno.com.ar/programas.html>. Por supuesto este archivo debe abrirse con un programa que soporte archivos de audio multicanal (Este ejemplo fue abierto sin inconvenientes por Audacity - <http://audacity.sourceforge.net>). Una copia de la última versión de GtKonvolve puede obtenerse en <http://www.estebancalcagno.com.ar/programas.html>.

### § 8.3 3DEV: UNA HERRAMIENTA PARA EL DISEÑO DE TRAYECTORIAS DE MÚLTIPLES SONIDOS DIRECCIONALES EN UN ESPACIO EN 3 DIMENSIONES

3DEV es un programa de código abierto desarrollado en el marco de esta tesis junto a Oscar Pablo Di Liscia para sistemas GNU/Linux y MS Windows, cuya utilidad es la creación, transformación y coordinación temporal de múltiples trayectorias de fuentes sonoras, dentro de un espacio de tres dimensiones.



**Figura 8.4:** Ejemplo de impresión de reporte generado por Gtconvolve una vez que se procesaron los datos de entrada.

3DEV fue concebido como una herramienta general a ser usada en composiciones de música electroacústica, y los datos generados por este pueden ser transmitidos de una forma simple y efectiva hacia varios programas de espacialización, entre ellos Csound y Pure Data.

El trabajo con 3DEV se enfoca en la generación de trayectorias de sonido y se centra en el diseño de una interface gráfica para la formación de múltiples trayectorias tridimensionales, independientemente del motor de espacialización de audio que se utilice. Esta separación permite al compositor seleccionar la técnica y programa de audio más apropiado para procesar las señales de audio, y luego, utilizar un solo conjunto de datos en los diferentes motores de



espacialización que se utilizan.

El desarrollo de 3DEV surgió a partir de poder comenzar a pensar una herramienta que permita el diseño a priori de las diferentes trayectorias que diferentes sonidos pueden tener en una composición electroacústica, dentro de un ámbito de tres dimensiones, y la posibilidad de observar y corregir estas trayectorias a partir de la evolución temporal que estas produzcan. A continuación se desarrollaran las características principales que presenta este programa.

### 8.3.1. Características básicas de 3DEV

3DEV fue concebido para brindar asistencia a los compositores de música electroacústica en la creación, transformación y coordinación temporal de múltiples trayectorias de fuentes sonoras direccionales en un espacio tridimensional. Los binarios para sistemas GNU/Linux y MSWindows, así como también el código fuente y documentación pueden ser descargados en: <http://www.estebanalcagno.com.ar/programas.html>. 3DEV permite generar un número arbitrario de trayectorias y trabajar sobre ellas utilizando algunas herramientas de edición. Las trayectorias son delimitadas mediante puntos (nodos) que pueden ser visualizados individual y globalmente. Una de las características más valiosa de 3DEV es su capacidad de establecer una relación temporal precisa entre cualquiera de las trayectorias generadas y las señales de audio que están asociadas a estas. Adicionalmente, también es posible crear y modificar un vector de rotación de cada fuente sonora virtual, el cual puede también ser coordinado con la evolución temporal de la señal de audio. Si el motor de audio que se decide utilizar permite el trabajo con

la directividad de las fuentes sonoras, entonces es posible transmitir de que manera la fuente virtual cambia eventualmente de orientación y simular su radiación direccionada. El control del tiempo entre los diferentes nodos espaciales (indicado como un porcentaje del total de la duración de la señal que se asocia a la trayectoria) es hasta el momento la única forma de manejar la velocidad del movimiento.

### 8.3.2. Descripción del programa 3DEV

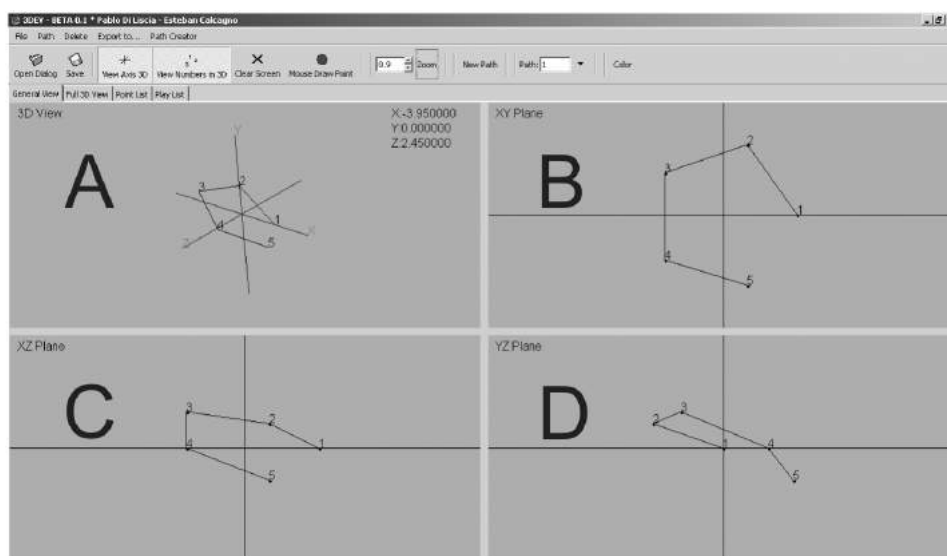
3DEV utiliza la librería GTK+, escrita en lenguaje C (<http://www.gtk.org/>), como herramienta de diseño de su interface gráfica, como también la librería Pthread (<https://computing.llnl.gov/tutorials/pthreads/>) y la librería escrita en lenguaje C para el manejo de entrada y salida de audio, Sndlib (Schottstaedt, <https://ccrma.stanford.edu/software/snd/snd/sndlib.html>). Además, utiliza el procesamiento de código de Python (versión 2.6: <http://www.python.org/>) embebido en el código C, para el envío de información, vía OSC (Open Sound Control, <http://opensoundcontrol.org/>), a otros programas. Hasta el momento, 3DEV se comunica mediante OSC con programas desarrollados bajo la plataforma Pure Data para mostrar el movimiento e interacción de las trayectorias, pero esta característica se encuentra en continuo desarrollo. 3DEV se divide en tres diferentes áreas de trabajo: la ventana de Trayectorias Espaciales (Spatial Trajectories), la ventana de Archivos de Audio (Audio Files) y la ventana de conexión al motor de audio. Las dos primeras son las que llevan la mayor parte del trabajo hasta el momento, y esta última esta aun en desarrollo, y por ahora se está usando como salida Csound y envío de datos hacia Pure Data. Sin embargo, como

mencionamos, la independencia entre el diseño del movimiento y el procesamiento de audio permitirá el uso de otros programas para la espacialización del sonido.

### 8.3.3. Ventana de trayectorias espaciales

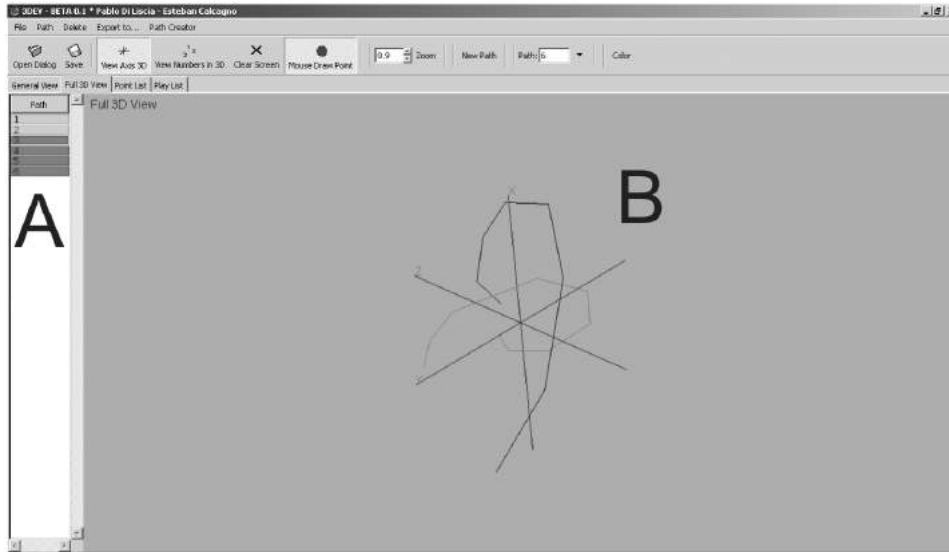
La interface gráfica para las trayectorias, cuyas funciones básicas para el manejo gráfico en tres dimensiones fueron adaptadas de Harlow (1999), tiene cuatro pestañas que se describen a continuación:

1- La vista principal ( – Figura 8.5), está dividida en cuatro secciones de igual tamaño: A) muestra una gráfica en tres dimensiones de las trayectorias que han sido diseñadas hasta el momento (con opciones de rotación y zoom 3D); B) indica la proyección de la trayectoria en el plano XY; C) indica la proyección de la trayectoria en el plano XZ; D) indica la proyección de la trayectoria en el plano YZ. Estas tres permiten el diseño de las trayectorias mediante la utilización del mouse.



**Figura 8.5:** General View. Pantalla principal de 3DEV, permite diseñar y construir las trayectorias mediante el agregado de puntos en cualquiera de los tres planos. Estos puntos pueden ser relocalizados mediante el mouse.

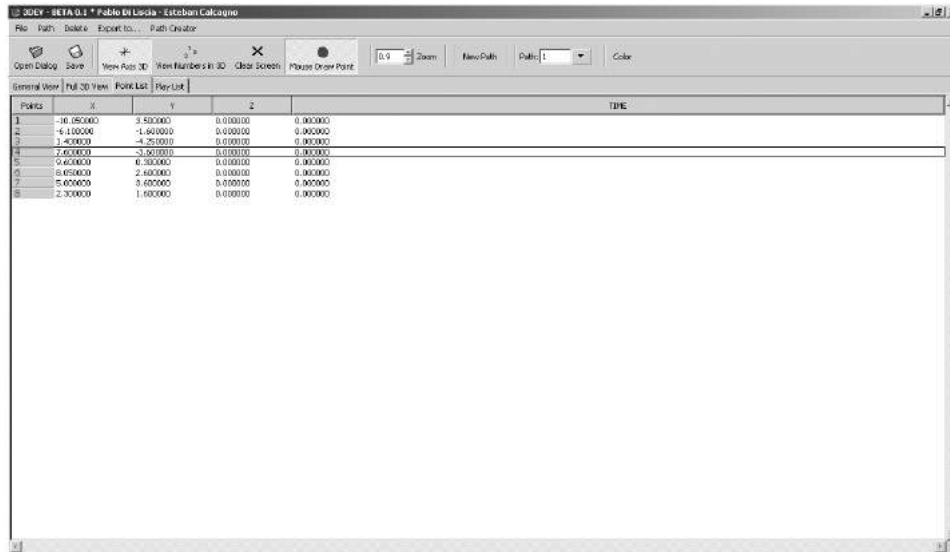
2- La vista 3D completa (Full 3D View – Figura 8.6), donde se puede ver: A) un menú con todas las trayectorias y B) una gráfica en tres dimensiones de todas las trayectorias creadas interactuando entre sí. En esta vista pueden seleccionarse las trayectorias que se desea visualizar y así poder observar interacciones entre trayectorias diferenciadas.



**Figura 8.6:** Full 3D View. En esta pantalla se pueden visualizar todas las trayectorias diseñadas interactuando entre sí, además de una lista de las trayectorias las cuales se pueden esconder o mostrar.

3- El menú de las listas de puntos (Point List – Figura 8.7), el cual muestra una lista de las coordenadas cartesianas de cada uno de los puntos que contiene la trayectoria elegida. Este menú permite la edición manual de las coordenadas de cada punto.

4- El menú de lista de reproducción (Play List) muestra una lista de todas las trayectorias, las cuales se pueden asociar con archivos de audio. Mediante este menú es posible acceder a la ventana de archivos de audio (Audio File Window), que se describe en la siguiente sección.



**Figura 8.7:** *Point List.* Lista de puntos de cada una de las trayectorias, la cual permite corregir manualmente las coordenadas cartesianas de los puntos. Esto permite un mayor control sobre el diseño.

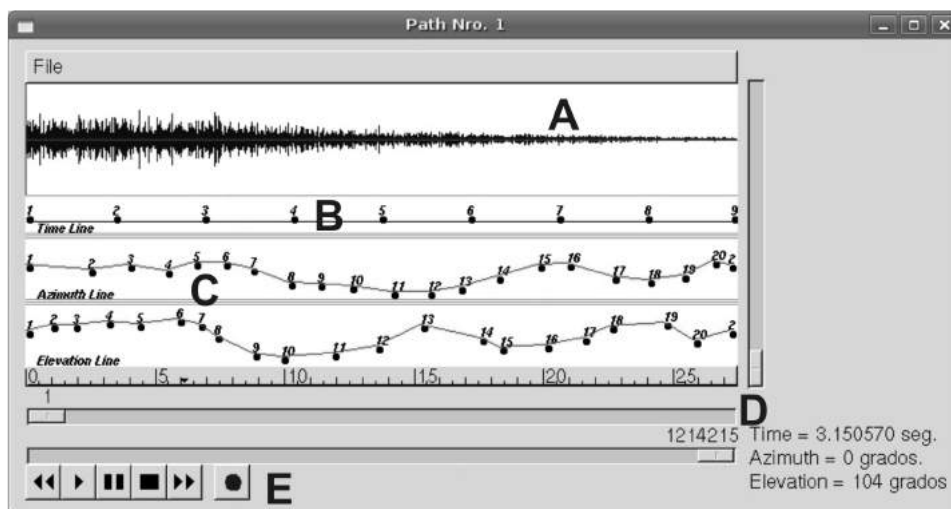
#### 8.3.4. Ventana de archivos de audio

La Figura 8.8, muestra el visualizador y reproductor de archivos de audio. Empezando desde arriba hacia abajo, esta ventana está compuesta por:

- A- Una imagen de la forma de onda del archivo de audio seleccionado.
- B- Una línea horizontal que relaciona la duración del archivo de audio con cada nodo de la trayectoria diseñada. La posición horizontal de cada uno de estos nodos puede ser modificada para cambiar la duración de cada segmento de las trayectorias. De esta manera, se puede lograr la sincronización exacta de tiempo entre el movimiento y el audio relacionado a esa trayectoria.
- C- Un sistema de dos envolvente editables indican la orientación de la fuente sonora (ángulos de acimut y elevación). Mediante este sistema, la orientación de la fuente sonora puede ser cambiada dinámicamente y también sincronizada con efectividad con la señal de audio.
- D- Una sección con una referencia temporal de las señales de audio selecciona-

da (en segundos y en samples), que además tiene opciones de zoom y controles de reproducción.

E- Los controles de Play, Stop y Pause del reproductor de archivos de audio, mediante los cuales se puede realizar una escucha de los mismos. Hasta el momento esta opción es funcional solo en sistemas GNU/Linux, pero se encuentra en desarrollo para funcionar bajo sistemas MSWindows.



**Figura 8.8:** *Audio File Window. Permite la visualización de los archivos relacionados con cada trayectoria, además de la coordinación temporal de los puntos, rotación en acimut y elevación. En algunas versiones se permite la reproducción del archivo de audio.*

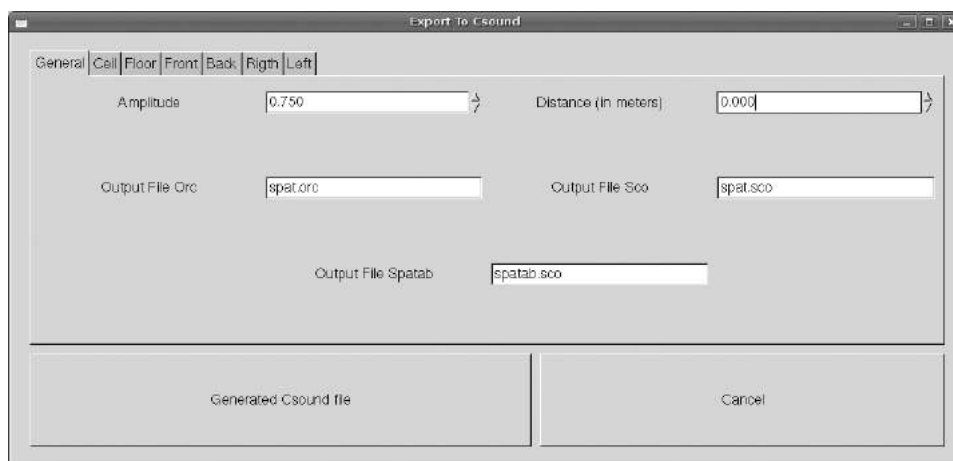
### 8.3.5. Conexión a los motores de audio y espacialización

Hasta el presente 3DEV puede enviar datos de las trayectorias diseñadas hacia dos programas que realicen el procesamiento espacial de los archivos seleccionados: Csound y Pure Data.

### 8.3.6. Csound

3DEV utiliza el programa Csound (Barry Vercoe) como motor de espacialización. Para esto, 3DEV genera un archivo de texto con el formato de una

partitura Csound. Este archivo de texto es un tipo de script que puede ser utilizado por Csound para sintetizar una señal espacializada basada en una orquesta (prediseñada por el Dr. Oscar Pablo Di Liscia), el cual constituye el modulo básico de DSP. En este caso, la orquesta de Csound usa la unidad generadora spat3d (Varga, 2000) que implementa técnicas de espacialización Ambisonic (ver Malham, 2010). La orquesta diseñada provee espacialización en tres dimensiones tanto para la señal directa como para las reflexiones tempranas dentro de una sala virtual, cuyas características deben ser configuradas por el usuario. La Figura 8.9 muestra los datos de entrada necesarios para especificar las características de nuestra sala virtual.



**Figura 8.9:** *Export To Csound.* Este modulo permite la configuración para la creación de una partitura de Csound mediante la información de las trayectorias diseñadas.

La conexión de las trayectorias de sonido es muy fácil de lograr por medio de tripletes de tablas (vectores numéricos legible por Csound) que contienen cada una de las coordenadas cartesianas (X, Y, Z). Cada trayectoria de sonido se representa en la partitura de Csound como una nota con su tiempo correspondiente de inicio, la duración de la escala de amplitud, y las tablas que indican su movimiento. De esta manera, es posible espacializar y mezclar múltiples fuentes de sonido virtuales que se mueven de diferentes maneras.





### 8.3.8. Consideraciones finales

3DEV está todavía en desarrollo pero ya constituye una herramienta muy útil para los compositores de música electroacústica. El siguiente paso es mejorar el conexionado con otros programas de espacialización, como Pure Data. Además, otras posibilidades de desarrollo sería incluir dibujos y varias rutinas de transformación de las trayectorias 3D, así como el almacenamiento de los datos en diversos formatos para su uso con los programas de espacialización estándar. Pudiendo almacenar los datos generados y almacenados con precisión, sería un asunto sencillo convertirlos en diferentes sistemas de coordenadas, así como adaptarlo a los requisitos de sintaxis y tipos de datos utilizados por otros programas de espacialización.

Por último, como la visualización de muchas trayectorias espaciales complejas que funcionen juntas puede ser difícil de lograr, una sección de animación en tiempo real sincronizado con sus archivos de audio asociados es también una característica a desarrollarse en el futuro.

## § 8.4 GUISPA: CONTROL DE ESPACIALIZACIÓN DE FUENTES SONORAS EN TIEMPO REAL

GUISPA, fue diseñado y desarrollado para la obra “120 Máquina Lírica” del colectivo Buenos Aires Sonora (BAS). Esta aplicación fue íntegramente desarrollada dentro del programa Pure Data, el cual permite el procesamiento de audio en tiempo real, además de proveer una amplia colección de herramientas para la manipulación de fuentes sonoras.

GUISPA ha sido diseñado como un entorno con características de manejo interesantes: permite el control de las trayectorias de hasta 16 fuentes sonoras en tiempo real, también elegir entre varias técnicas de espacialización sobre un arreglo de 4 parlantes, y además contiene un flujo de señales de entrada y salida similar a la de una consola de sonido, lo que permite una rápida adaptación al sistema para personas que están acostumbradas a este hardware.

Este capítulo explicará las fases de desarrollo de esta aplicación, como se comportó en performance y las ventajas y desventajas de su uso en la obra además de las posibles mejoras que podrían aplicarse para un funcionamiento más efectivo.

#### 8.4.1. Características y desarrollo

En el Capítulo 9, Sección 7, de esta tesis se realizará un análisis sobre la producción de la obra “120 Máquina Lírica”. Este análisis presenta el desarrollo de las diferentes escenas de la obra poniendo el énfasis en los sistemas utilizados y la composición de la misma. Para la última escena, planteada en la Sala TACET del Teatro Argentino de La Plata, fue necesario idear un sistema de reproducción multicanal hacia donde se podían dirigir 16 fuentes sonoras compuestas por voces cantadas, instrumentos acústicos (ensamble de metales, xilófono y percusión), grabaciones de audio y el procesamiento en tiempo real realizado a cada una de estas fuentes, las cuales fueron tomadas con distintos tipos de micrófonos e ingresados a una consola, que luego enviaba 16 señales hacia la placa de sonido de una computadora que alojaba el programa GUISPA. Según el planteo de la obra, estas señales debían poder ser localizadas en el espacio del escenario central de la sala (un espacio cuadrado de aproxi-

madamente 20 metros de lado), por lo que se optó por plantear un sistema de 4 parlantes que rodeara a la audiencia. El desarrollo mismo de la escena planteaba la localización y movimiento de cada una de estas fuentes en este espacio debido al desarrollo de juegos perceptivos entre información visual y sonora que necesitaban de estas características de posicionamiento de fuentes. A partir de estas ideas, se comenzó a desarrollar GUISPA.

El primer paso fue elegir la plataforma donde se desarrollaría el programa de espacialización. Como vimos en los capítulos anteriores, en general para los programas GTKonvolve y 3DEV se eligió realizar los desarrollos desde cero mediante la programación a bajo nivel en lenguaje C. Esto permitió el diseño de herramientas propias y adaptadas a cada problema. En este caso se eligió realizar la programación del mismo dentro del software de procesamiento de audio en tiempo real Pure Data, por dos razones principales: en primera instancia, Pure Data ofrece tiempos de desarrollo muy rápidos ya que posee una gran cantidad de librerías y objetos que resuelven la mayoría de los problemas que aparecen a la hora de reproducir, procesar y espacializar audio en tiempo real. Por otro lado un trabajo previo de desarrollo de programas bajo esta plataforma realizado por integrantes del Grupo Buenos Aires Sonora (entre ellos quien escribe) generó una base sustentable para el desarrollo de un programa con las características buscadas en esta ocasión.

Una vez elegida la plataforma, se diseñó la pantalla principal, la cual debía contener una Interfaz Gráfica de Usuario o GUI <sup>1</sup> que permitiera manejar las diferentes fuentes a espacializar mediante el uso de un mouse y una consola MIDI. Además debía tener la posibilidad de agrupar varias fuentes para rea-

---

<sup>1</sup>Del inglés, Graphics User Interface

lizar movimientos coordinados.

De esta manera GUISPA se dividió en tres partes fundamentales: la interfaz de usuario principal, la cual permitía el movimiento manual y automatizado de las trayectorias de las fuentes sonoras, la interfaz de usuaria secundaria, que permitía el direccionamiento de señales de entrada y salida y los objetos de Pure Data que se encargan de la espacialización. A continuación se realiza una reseña de cada parte de GUISPA.

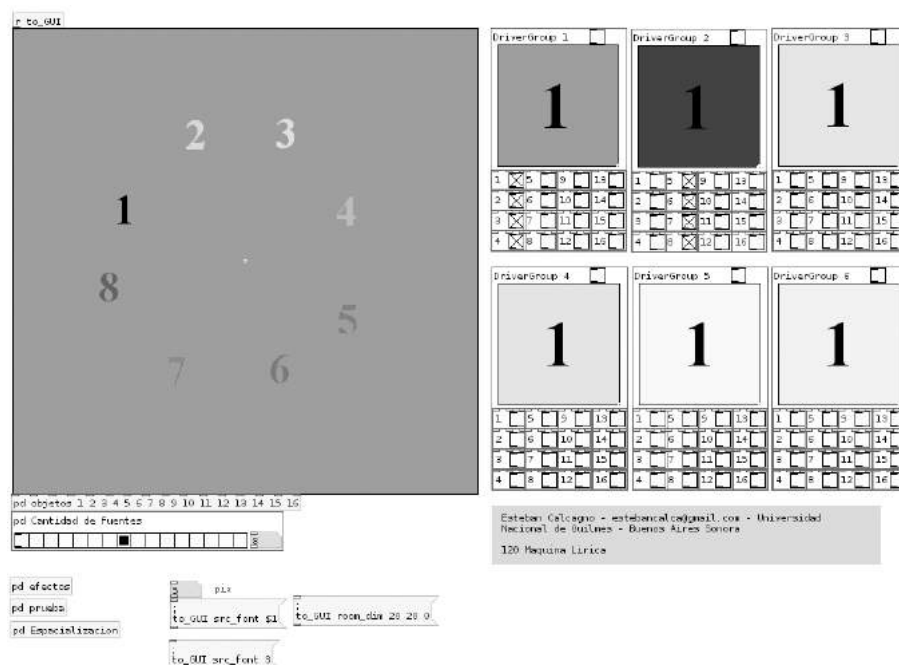
#### 8.4.2. Interfaz Gráfica de Usuario Principal

Para el desarrollo de la GUI que mostraba cada una de las fuentes a espacializar y el sistema de agrupamiento de fuentes se utilizó la librería `iemgui` (la cual viene por defecto en el paquete `Pd-extended`), mediante la cual se puede manejar una cierta cantidad de objetos gráficos que facilitan el diseño visual. Entre estos objetos se destaca `room_sim_3d`, el cual permite el posicionamiento en tres dimensiones de elementos dentro de un simulador de salas, al cual se le puede indicar la cantidad de elementos que contiene, la posición de un oyente virtual y las dimensiones de la sala.

En la Figura 8.11 se puede observar la pantalla principal de GUISPA.

Mediante la utilización del mouse de la PC se puede mover de forma independiente cada una de las fuentes simbolizadas por un número que aparecen a la derecha de la GUI. Lo interesante de `room_sim_3d` es que si bien se puede mover solamente de a una sola fuente por vez mediante el mouse, es posible realizar el movimiento de varias fuentes mediante el envío de mensajes al objeto. Estos mensajes podían ser enviados por los conductores de grupo

(Drivergroup) los cuales a su vez podían ser manejados de forma manual o mediante las ordenes enviadas por una consola MIDI <sup>2</sup>.



**Figura 8.11:** Pantalla principal de GUISPA. A la izquierda se puede apreciar el simulador de sala con 8 objetos rodeando un sweet spot; a la derecha los 6 módulos que permiten agrupar fuentes y moverlas de forma coordinada.

De esta manera, se pudo realizar el movimiento de múltiples fuentes en este espacio mediante el diseño de un conductor de grupos. En total, GUISPA podía manejar de forma coordinada 6 grupos de fuentes, los cuales fueron suficientes para asistir la espacialización de las fuentes de la obra planteada. La Figura 8.12 muestra el detalle de uno de los conductores de grupo que se podían utilizar. Como se observa, se puede elegir cualquiera de los 16 canales disponibles y manejarlos directamente con el objeto “1”, por lo que se puede desplazar mediante el mouse o la consola MIDI los elementos seleccionados de forma coordinada. Este módulo presentó además la posibilidad de automatizar el movimiento para que las fuentes se muevan de forma coordinada en círculos,

<sup>2</sup>En la obra “120 Máquina Lírica” se utilizó una consola MIDI usb Behringer Bcontrol Fader BCF2000 sin ningún inconveniente

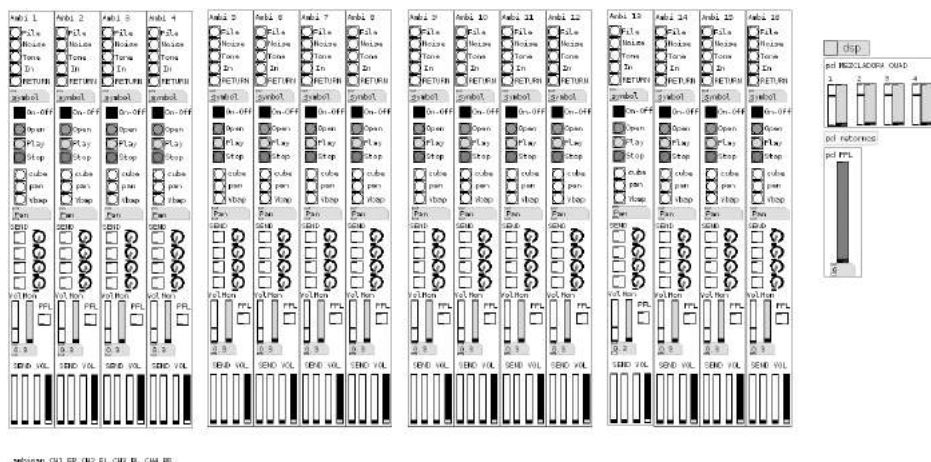
de mayor o menor tamaño.



**Figura 8.12:** *DriverGroup*: uno de los módulos manejadores de trayectorias de grupos de fuentes sonoras. Como se observa puede seleccionarse cualquiera de las 16 fuentes para que estas sean manejadas como un grupo y moverse en el espacio de forma coordinada.

### 8.4.3. Interfaz Gráfica de Usuario Secundaria

Dentro de GUISPA la Interfaz Gráfica de Usuario Secundaria permite el direccionamiento de las señales de entrada y salida mediante una matriz de canales que simula una consola de audio. Esta matriz se compone de 16 módulos que permite realizar selecciones de las cuales depende el comportamiento de las señales. La Figura 8.13 muestra la vista completa de la interfaz secundaria.



**Figura 8.13:** *Interfaz de direccionamiento de señales de audio de GUISPA.*

Cada uno de los 16 módulos de flujo corresponde a una de las fuentes que se observan en la GUI principal, y sus números están relacionados: el módulo 1 corresponde a la fuente 1, el módulo 2 a la fuente 2 y así sucesivamente. La configuración de cada uno de estos módulos es la siguiente:

- a) Un sector donde se indica de que manera entrará el audio al programa (Figura 8.14).
- b) Un sector que indica que tipo de técnica se utiliza para espacializar la fuente (Figura 8.15).
- c) Un sector de re direccionamiento y monitoreo de la señal (Figura 8.16).

La Figura 8.14 muestra entonces el sector donde se indica de donde proviene la señal de entrada al canal.

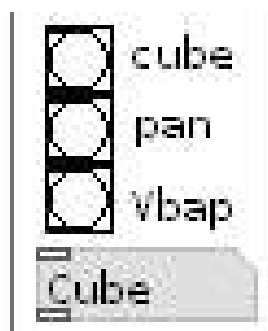


**Figura 8.14:** Sector de selección de señal de entrada.

Como se puede observar, se puede elegir como señal de entrada tanto un archivo de audio localizado en la computadora (File), para espacializar sonido

pregrabado, como Ruido Blanco (Noise)<sup>3</sup> o un tono puro de 440HZ (Tone)<sup>4</sup>. Luego se puede elegir una entrada de audio de la placa de sonido (In), donde cada módulo está direccionado a la entrada correspondiente (módulo 1, entrada 1), la cual es muy útil para espacializar sonidos que se están ejecutando en vivo (voces, instrumentos, etc), y por último se puede optar por una señal que provenga de otro de los canales (Return), muy útil para ubicar, duplicar o mezclar señales en una posición. Más abajo se observan los controles de carga de archivo de audio (Open), reproducción y detención (Play-Stop). Por último se puede elegir silenciar o reproducir el canal (On-Off).

La Figura 8.15 muestra el sector del módulo que permite elegir la técnica utilizada para espacializar el sonido que ingresa al canal.



**Figura 8.15:** Sector de elección de técnica de espacialización.

En este caso se puede elegir entre tres técnicas: Ambisonics, para arreglos 8 parlantes (Cube), Ambisonics para arreglos de 4 parlantes (Pan), esta opción se utilizó en “120 Máquina Lírica”, y VBAP. Esta última no es acorde a una arreglo de 4 parlantes, pero podría utilizarse si se dispusiera del arreglo correspondiente (ver Capítulo 5, técnicas y sistemas de espacialización).

La Figura 8.16 muestra el sector de re direccionamiento y monitoreo de la

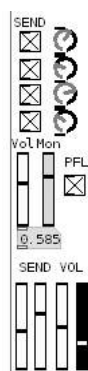
---

<sup>3</sup>Estímulos muy útil para realizar calibraciones y probar el sistema espectralmente, ya que como vimos en varios capítulos de esta tesis, los sonidos con contenido espectral complejo ayudan a la localización adecuada de las fuentes sonoras, tanto en acimut como en distancia

<sup>4</sup>Para medir niveles de intensidad y ganancia del sistema de reproducción



señal. Aquí se puede observar cuatro opciones de envíos auxiliares, los cuales estaban pensados para poder contar con mezclas de monitoreo de cada canal (SEND). Así mismo, en este sector se puede establecer la intensidad de salida del canal (Vol), y se tiene un monitoreo de la señal mediante un vúmetro (Mon). Por último está disponible la opción de PFL (Post Fader Level) y de monitoreo de los niveles de amplitud de los monitoreos (SEND VOL).

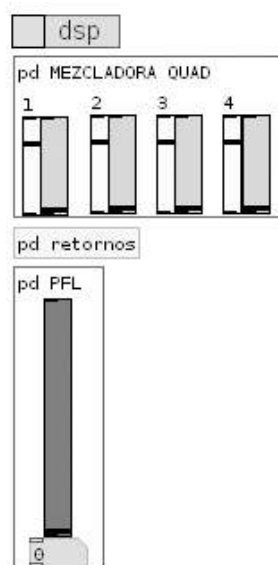


**Figura 8.16:** Sector de monitoreo y re direccionamiento de la señal

Además de los módulos que manejan las señales de entrada a espacializar, esta interfaz secundaria contiene un módulo extra que se encarga de calibrar la intensidad para la salida de cada canal de la cuadrafonía, que además permite monitorear la salida de cada canal, un selector de inicio o parada del procesamiento, y el monitor de PFL. La Figura 8.17 muestra este módulo.

#### 8.4.4. Espacialización de las fuentes sonoras

Para realizar el procesamiento de espacialización de las señales que entraban a GUISPA se utilizó objetos de las CICM Tools, ambicube, ambipan y vbapan (<http://cicm.mshparisnord.org>). Ambicube realiza ubicaciones y trayectorias espaciales de fuentes sonoras en Ambisonics de 1er orden, para dispositivos de parlantes que conforman un cubo. Ambipan, hace lo mismo para una cuadrafonía, y VBAPAN, aplica la técnica VBAP pero mediante



**Figura 8.17:** Representación esquemática del acimut, elevación y distancia.

ternas de parlantes.

Los tres objetos reciben como argumentos las coordenadas cartesianas en dos dimensiones de cada uno de los objetos a espacializar. Cada uno de estos módulos tiene integrado estos tres objetos los cuales se activan al seleccionarlos en el selector, para que no estén procesando los tres a la vez y consuman recursos innecesariamente.

#### 8.4.5. Performance en vivo

La performance en tiempo real de GUISPA ha sido satisfactoria pero no ideal. En general el sector de espacialización fue bastante bueno y no presentó inconvenientes a la hora de procesar 16 fuentes sonoras al mismo tiempo. Sin embargo los problemas más evidentes de procesamiento en tiempo real aparecieron en la interfaz gráfica de usuario principal, específicamente cuando se estaba al tope de procesamiento (16 fuentes sonoras al mismo tiempo) se percibía un retraso importante en la respuesta cuando se intentaba mover una fuente mediante el mouse sobre el simulador de sala. Sin embargo lo mismo no

ocurría cuando se movían las fuentes desde el manejador de grupos o mediante la consola MIDI.

#### 8.4.6. Mejoras a futuro

La herramienta GUISPA ha servido sin inconvenientes para organizar y controlar eficazmente la localización y el movimiento de fuentes sonoras en tiempo real para la obra “120’ Máquina Lírica”. Pueden sin embargo, plantearse mejoras a futuro que tienen que ver exclusivamente con la optimización y la ampliación de las características que este programa posee.

La primera mejora a plantear sería el diseño y desarrollo de la interfaz de control primaria mediante otra plataforma, por ejemplo Processing (<http://processing.org>), lo que bajaría los niveles de consumo de Pure Data, quien solamente estaría dedicado a procesar audio, dejando margen para el manejo de una mayor cantidad de fuentes sonoras. (esto tendría que correr en otra computadora, dado que de lo contrario, también consumiría una porción importante de capacidad de procesamiento de la CPU). Otro agregado interesante sería la posibilidad de contar con módulos de procesamiento alternativos como filtros, granuladores, reverberación, etc. Pure Data cuenta con todas las herramientas necesarias para desarrollar estas herramientas, y que dotarían a GUISPA de una funcionalidad más profunda.

Por último, sería interesante permitir que GUISPA pueda recibir datos de trayectorias espaciales de otros programas. Una posible e interesante unión podría darse con 3DEV (Capítulo 8, Sección 3), lo cual permitiría el diseño y ejecución de trayectorias espaciales en tiempo real generando una suite de espacialización sumamente interesante.

---

Puede obtenerse una copia de GUISPA, junto a todas las librerías necesarias para su funcionamiento en <http://estebanalcagno.com.ar/programas.html>.

## REFERENCIAS BIBLIOGRÁFICAS

- CETTA, P.: *Un modelo para la simulación del espacio en música*. Editorial de la Universidad Católica Argentina, Buenos Aires, Argentina, 2007.
- CHOWNING, J.: «The simulation of Moving Sound Sources». *Journal of the Audio Engineering Society*, 1971, **19**, pp. 48–52.
- DI LISCIA, O.P.: «Los modos de vínculo de la concepción espacial del sonido con la poiesis de la música electroacústica». *Tesis Doctoral, Universidad Nacional de Rosario, Facultad de Humanidades y Arte.*, 2005.
- FARINA, A.: «Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique». *AES convention 108*, 2000.
- FURSE, R.: «Csound Phase Vocoder and Ambisonic/Stereo Spatialisation Utilities». *MN Audio Library, disponible en <http://www.muse.demon.co.uk/csound.html>*, 1995.
- : «VSpace Introduction». *MN Audio Library, disponible en <http://www.muse.demon.co.uk/csound.html>*, 1999.
- HARLOW, E.: *Desarrollo de Aplicaciones Linux con GTK+ y GDK*. Prentice Hall, Madrid, 1999.
- HAYES1999: *Schaum's outline of theory and problems of digital signal processing*. McGraw-Hill Professional, New York, 1999.
- KARPEN, R.: «Space and Locsig Ugs». *En: The Csound Manual, disponible en: <http://www.csounds.com/manual/html/locsig.html>, <http://www.csounds.com/manual/html/space.html>.*, 1998.
- KENDALL, G.: «La interpretación de la espacialización electroacústica: atributos espaciales y esquemas auditivos». *En Música y espacio: ciencia, tecnología y estética. Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 241–259.
- KENDALL G., MARTENS W.L. y S.L., DECKER: «Spatial reverberation: discussion and demonstration». *Current directions in computer music research. MIT Press*, 1990, pp. 65–87.
- LOPEZLEZCANO, F.: «Quad sound playback hardware and software». *C.C.R.M.A., Stanford University.*, 1992.
- MALHAM, D.: «El espacio acústico tridimensional y su simulación por medio de Ambisonics». *En Música y espacio: ciencia, tecnología y estética, Editorial de la Universidad Nacional de Quilmes*, 2010, pp. 161–202.
- MOORE, F.R.: «A General Model for Spatial Processing of Sounds». *Computer Music Journal*, 1983, **7(3)**, pp. 6–15.
- : «Spatialisation of sounds over loudspeakers». *Current directions in computer music research. MIT Press*, 1989, pp. 89–103.
- : *Elements of Computer Music*. Prentice Hall., New Jersey, 1990.
- MUSIL, T.: «Binaural-Ambisonic 4. Ordnung 3D-Raumsimulationsmodell mit ortsvarianten Quellen und Hörerin bzw. Hörer für PD». *IEM Report N°38/07 [http://iem.at/projekte/publications/iem\\_report/report38\\_07/report38\\_07](http://iem.at/projekte/publications/iem_report/report38_07/report38_07)*, 2007.
- OPPENHEIM, A.V. y SCHAFER, R.W.: *Digital signal processing*. Prentice-Hall, New Jersey, 1975.

- OTONDO F., CAUSSE R., RINDEL J. y DELACUADRA P.: «Directivity of musical instruments in a real performance situation». *Proceedings of the International Symposium on Musical Acoustics (ISMA), Mexico City*, 2002, pp. 312–318.
- PULKKI, V.: *Spatial sound generation and perception by amplitude panning techniques*. Technological University of Helsinki, Helsinki, 2001.
- STOCKHAUSEN, K.: «Kontakte». *Texto publicado en el CD N°3, Electronische Musik, Stockhausen Verlag, Alemania.*, 1952-1960.
- TOMMASINI F.C., ARANEDA M. y O.A., RAMOS: «Auralización en tiempo real: implementación del modelo del oyente». *2º Congreso Internacional de Acústica UNTREF*, 2010.
- VARGA, I.: «Spatial sound generation and perception by amplitude panning techniques». *En The Csound Manual, disponible en: <http://www.csounds.com/manual/html/spat3d.html>*, 2000.
- WOZNIEWSKI M., SETTEL Z. y S., COOPERSTOCK: «Spatial sound generation and perception by amplitude panning techniques». *Pure Data Convention, Quebec, Canada [http://www.audioscape.org/twiki/pub/Audioscape/AudioscapePublications/audioscape\\_pdconv07\\_final.pdf](http://www.audioscape.org/twiki/pub/Audioscape/AudioscapePublications/audioscape_pdconv07_final.pdf)*, 2007.



## - Capítulo 9 -

---

---

### Análisis de la espacialidad del sonido en obras de Música Electroacústica: el caso de Buenos Aires Sonora

---

#### § 9.1 INTRODUCCIÓN

En este capítulo se abordará el análisis de la utilización del espacio sonoro en obras de música electroacústica, específicamente tres obras realizadas por el colectivo Buenos Aires Sonora (BAS), el cual enmarca su producción dentro de lo que se denomina Arte Para Sitios Específicos (Labelle, 2006).

Para esto, en una primera instancia se desarrollará una introducción a las nociones de Arte Sonoro y Arte Para Sitios Específicos, generando un marco apropiado de entendimiento de este tipo de producción musical.

Luego se realizará el análisis de las tres obras, el cual se dividirá en las siguientes partes:

Para cada obra se realizará una descripción general del contexto en el que se desarrollaron, tanto histórico, como arquitectónico y socio-cultural, lo que indicará las características acústicas y simbólicas de cada lugar, que en sí configuran una primera determinación respecto del espacio elegido o propuesto para realizar las obras.



Luego se procederá a la descripción de las tecnologías electrónico-digitales utilizadas, las técnicas y dispositivos de espacialización (basándose en lo desarrollado en el Capítulo 5, técnicas y sistemas de espacialización), y el/los orgánicos instrumental/es y medios acústicos elegidos en cada caso.

Finalmente, se analizarán las obras desde el punto de vista musical y sonoro en base a la metodología propuesta por Kendall (2010), la cual fue presentada en el Capítulo 4, enfoques analíticos de la espacialidad en la música electroacústica.

## § 9.2 DISEÑO Y CONSTRUCCIÓN DEL ESPACIO SONORO PARA SITIOS ESPECÍFICOS

En la música para sitios específicos, el diseño y la construcción del espacio sonoro como elemento estructurante de una obra juega un rol preponderante. Bajo esta línea estética el colectivo argentino Buenos Aires Sonora ha realizado un conjunto de obras donde el espacio es tomado como base constructiva, tanto por la transformación que ejerce la acústica de un lugar sobre los sonidos que allí se desarrollan, como por la carga histórica y generadora de sentido que este produce en la obra. Este artículo intenta analizar las diferentes herramientas de construcción y diseño del espacio sonoro utilizado por Buenos Aires Sonora en tres obras que se diferencian tanto por el lugar en donde se desarrollan como por los materiales sonoros utilizados.

### § 9.3 ARTE SONORO Y ARTE PARA SITIOS ESPECÍFICOS

Dentro de las nuevas líneas artísticas (tanto visuales como sonoras) que iniciaron su desarrollo en la segunda mitad del siglo pasado, existe una corriente estética que se refiere a las obras en sitios específicos. Entendemos como creación en sitios específicos a aquellas que toman en cuenta en forma integral el espacio elegido para su presentación. Según Liut (2009a) esta integralidad implica no solo atender aspectos topológicos de un espacio sino también su carga simbólica y la relación que este lugar tiene con la comunidad que lo habita o transita. Sin duda, las características acústicas de un lugar específico serán parte importante de esta integración del espacio, más aun si se trata de una obra que utiliza el sonido como actor principal.

En esta línea de obras para sitios específicos aparece Buenos Aires Sonora (BAS)<sup>1</sup>, un grupo conformado por músicos argentinos que decidió trasladar su experiencia y saberes artísticos y científicos, de la sala de concierto a la calle. Desde 2003 BAS ha realizado intervenciones artísticas en diferentes espacios públicos de la Ciudad Autónoma de Buenos Aires tomando al sonido como principal herramienta.

BAS también desarrolla su trabajo en lo que se denomina arte sonoro, ocupando varias ramas dentro de esta clasificación. El arte sonoro nace de la plástica en la década de 1960, donde artistas visuales utilizan sonidos en sus obras, pero donde también hay participación de músicos experimentales. El arte sonoro incluye prácticas y formatos como la instalación sonora, la escultura sonora, la

---

<sup>1</sup>Si bien la formación de BAS ha variado desde sus comienzos, cuando se realizaron las obras aquí comentadas el grupo estaba conformado por: Martín Liut, Mariano Cura, Hernán Kerlleñevich, Pablo Chimenti y Esteban Calcagno

escritura sonora, la poesía sonora, las músicas habladas, el radio arte, el paisaje sonoro, el audio performance, los emplazamientos conceptuales, algunas música electroacústicas y otro tipo de interacciones tecnológicas en espacios reales y virtuales (López Cano, 2012).

Según Liut (2009a), el arte sonoro que aplica BAS en sitios específicos, comparte su vocación por una integración y un dialogo con los espacios a ser intervenidos con otras manifestaciones artísticas que se desarrollan en el espacio público, como la arquitectura, los monumentos y los graffitis y stencils. Entonces, las creaciones de Buenos Aires Sonora mantienen una comunicación con las características constructivas y estructurales de los espacios como la carga simbólica y las características acústicas que poseen, pero con la diferencia de que no espera ser monumento, sino arte efímero. Se utiliza el lugar por lo que es, por lo que fue y por lo que suena y en un momento de tiempo determinado.

Debido a las características arquitectónicas de cada lugar que es intervenido por BAS, sus cualidades sonoras son disímiles en tamaño, materiales, ruido de fondo y contexto.

Cada espacio tiene una arquitectura sonora característica (Blessner y Slater, 2006), la cual se convierte en uno de los materiales a trabajar, controlar, analizar y modelar. En muchos casos se trabaja directamente con las cualidades que estos espacios imprimen en los sonidos, pero otras veces se intenta transformar estas características. Con el advenimiento de las nuevas tecnologías es posible realizar varias tareas que en otros momentos de la historia hubiesen sido imposibles. Gracias al desarrollo de micrófonos, computadoras, placas de sonido, software y sensores, y al poder de trabajar con la información so-

nora en tiempo real (llámese a esto por ejemplo, poder capturar un sonido, transformarlo en el espectro de frecuencias, en intensidad o en el tiempo y reproducirlo un retraso casi imperceptible para el oído humano) es posible capturar las características acústicas de determinados materiales o espacios, y construir, espacializar o deconstruir estos espacios de formas simples o complejas. De esta manera el espacio sonoro se convierte en un bien que puede ser tratado como un elemento estructurante de la música, como lo son el ritmo, el timbre y la altura. En este trabajo en particular, se analizarán tres obras del colectivo Buenos Aires Sonora, donde el espacio es utilizado estructuralmente y como generador de sentido en performance, y donde el uso de tecnologías variadas de captura, procesamiento y reproducción del sonido permiten una apropiación del espacio y su posible construcción y deconstrucción.

#### § 9.4 TRES OBRAS, TRES FORMAS DE PENSAR EL ESPACIO MUSICAL

Buenos Aires Sonora ha presentado entre 2010 y 2011 tres obras en espacios que se diferencian tanto por sus características acústicas como por su contenido y significado. Cada una de estas obras presenta un diseño característico del espacio como elemento estructurante de la composición.

La primera de ellas, “Oí(r) el ruido”, estrenada en enero del 2010, trata de una intervención que el grupo realizó en el Pabellón Bicentenario que el gobierno de la Ciudad de Buenos Aires montó en la Plaza Jardines de Invierno (Av. Libertador y Fray Justo Santamaría de Oro) en conmemoración del Bicentenario de la República Argentina. El pabellón estaba construido con caños de hierro cuadrado y medía 45 metros de largo, 9 de ancho y 2,5 metros de

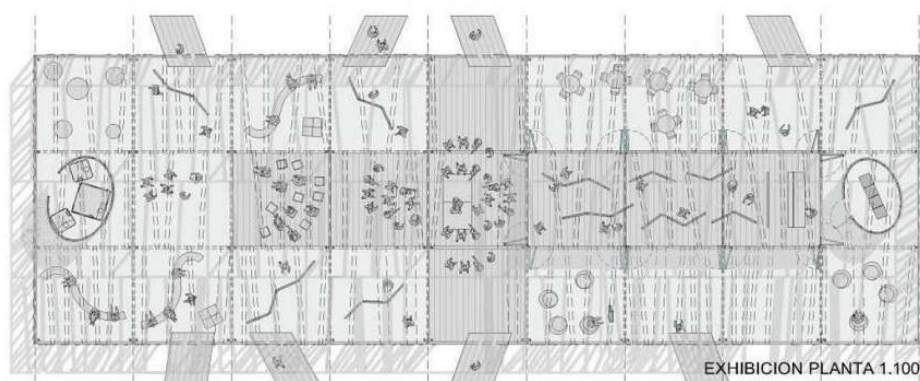
alto y constaba, además, con otros materiales como madera y telas plásticas. Todos los elementos de la estructura fueron capturados mediante microfones para obtener información sonora. La segunda obra a analizar, “Ocupación Invisible”, fue estrenada a mitad de 2010 en Buenos Aires y luego presentada en la ciudad de Montevideo (Uruguay) a finales del mismo año. Esta obra fue el primer trabajo interactivo de BAS, que constaba de una habitación vacía y oscura dotada con sensores de movimiento que captaban la circulación de los espectadores y reproducían audios determinados.

Una tercer obra, “120 Máquina Lírica”, estrenada en septiembre de 2010, se desarrolló en varios espacios del Teatro Argentino de la Plata, lo que generó que esta se desarrollara dinámicamente en diferentes ambientes acústicos reales y virtuales. En las siguientes páginas se intenta analizar cada una de estas composiciones desde el punto de vista espacial, o sea determinar cómo cada una de estas obras utiliza el espacio que interviene y como este espacio constituye uno de los aspectos fundamentales en su concepción.

## § 9.5 OÍ(R) EL RUIDO

### 9.5.1. Descripción general

Para los festejos del primer bicentenario de la República Argentina, la Ciudad de Buenos Aires decidió llamar a un concurso para diseñar el llamado “Pabellón Argentino Bicentenario 2010”, en el cual se desarrollarían los festejos para la ocasión. El diseño ganador, presentado por los arquitectos Frangella, Sardìn, del Puerto y Berson, estaba conformada por 27 cubos de 5mX5mX3m, construidos con perfil de caño de hierro hueco, que unidos conformaban una estructura de 45 metros de ancho por 9 de largo por 3 de alto (Figura 9.1).



**Figura 9.1:** *Planta Pabellón Argentino Bicentenario 2010 – Imagen: Catálogo Pabellón Bicentenario Proyecto Arquitectónico.*

Esta estructura contaba además con un piso de tablas de madera, telas plásticas dispersadas en su contorno (Figura 9.2) y dos gazebos de chapas de cinc, que se observan en la Figura 9.3.

No es menor describir la zona y el lugar exacto donde se emplazó el pabellón. La Ciudad de Buenos Aires decidió construir este homenaje a la patria en uno de los barrios de poder adquisitivo más alto de la Ciudad de Buenos Aires, Palermo, en Av. Libertador 3600, dentro del Parque Jardines de Invierno.



**Figura 9.2:** *Pabellón Bicentenario de la Ciudad de Buenos Aires. Imagen: Catalogo Pabellón Argentino Proyecto Arquitectónico.*



**Figura 9.3:** *gazebos de chapas de cinc en uno de los extremos del Pabellón Argentino del Bicentenario. – Imagen: <http://www.eternautasblog.com>*

El grupo Buenos Aires Sonora fue comisionado para presentar una obra que utilizará el Pabellón como escenario. Sin embargo BAS decidió utilizar al Pabellón no como un escenario sino como un instrumento en sí mismo. Debido a la cantidad y diversidad de materiales de construcción, de los cuales se podían obtener diferentes sonoridades, este objetivo parecía posible. El día jueves 28 de enero de 2010 el grupo Buenos Aires Sonora, junto a la Compañía de Danza Espacio Contemporáneo, de Diana Theocharidis, presentó su obra “Oí(r) el ruido”. Entonces, la obra se ocupa solamente del espacio físico en términos puramente materiales: el pabellón como un gran instrumento musical, y el

espacio en cuanto a la particularidad física de estar al aire libre, tener una de las avenidas más transitadas de buenos aires “pegada” al pabellón, y una “pared” de edificios enfrente. El Parque es un lugar neutro a la vera de la avenida, y no mucho más.

En “Oí(r) el ruido” se utilizó entonces el pabellón como un instrumento gigante, que debido a sus dimensiones representó un desafío interesante, tanto por el espacio a cubrir, como por la idea de “tocar la estructura” y que el público la percibiera como un instrumento único. Así comenzó un periodo de estudio para el armado de la obra que constó de la elección del material sonoro y la posibilidad de utilizar un sistema de captura y reproducción no convencional.

### 9.5.2. Orgánico sonoro

Tomando como punto de partida el conjunto de materiales de construcción del Pabellón, del cual se podía obtener diferentes tipos de sonoridades, se comenzó por la captura de sonidos con diferentes tipos de micrófonos y diferentes tipos de excitadores.

Mediante la utilización de baquetas de diversos materiales (varillas rosca-das, palillos de madera de distinto grosor y palillos con pelotas de goma (comúnmente denominadas boligomas), mazos de goma y madera, escobillas, resortes gigantes de metal y caños de hierro) y micrófonos de diferentes características (piezoeléctricos, dinámicos activos de bajo y guitarra y condensadores aéreos)<sup>2</sup> se obtuvieron una serie de texturas sonoras que dieron forma a la obra.

Estos materiales fueron organizados de tal manera que conformaron el orgáni-

---

<sup>2</sup>Para una referencia sobre las características de construcción y funcionamiento de los micrófonos utilizados ver Ballou (2009)



co sonoro de la obra. Los sonidos captados por los micrófonos eran enviados a dos consolas Yamaha LS9 (digitales), las cuales adecuaron las señales para su reproducción. Estas mismas señales fueron enviadas a una computadora que realizó procesos en tiempo real: reverberaciones, retrasos, granulaciones y también espacialización virtual de las fuentes sonoras (Subía Valdez, 2011) y toda esta información sonora fue direccionada a un complejo sistema de reproducción multicanal de 32 parlantes amplificados diseñado de tal manera que permitió la ubicación y movimiento de varias fuentes sonoras en el espacio de una forma dinámica y sencilla.

### 9.5.3. Sistemas de captura, reproducción y espacialización

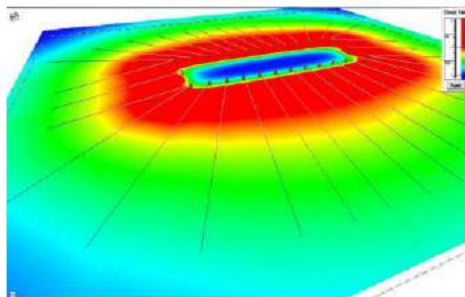
Al presentarse la obra al aire libre, el uso de las características acústicas y espaciales del ambiente se magnificaron. Cómo el sonido se desarrollaba en la distancia pasó a ocupar un rol muy importante en la composición misma de la obra. Como se explicó en el Capítulo 3, localización en distancia de fuentes sonoras, uno de las principales indicios de percepción de distancia de una fuente sonora es la amplitud, la cual comienza a decaer cuando la distancia se incrementa, y otro también muy importante es la pérdida de energía en el rango agudo del espectro. En ambos casos, las distancias extremas generan un efecto aun más marcado en estas dos pistas de percepción auditiva del sonido, por lo que pensar en estas características pasó a ser parte del problema. Además debíamos agregar que la obra debía competir con el ruido proveniente de la Av. Libertador, muy próxima a la instalación. Por eso se llegó a la conclusión de que se debía utilizar un sistema de reproducción multicanal, el cual además permitiera además aprovechar las características de distancias extremas que

poseía el pabellón. El hecho de tomar esta decisión devolvía el problema de siempre: ¿cómo lograr que la obra no suene en un sólo espacio, sino que pueda desarrollarse en varios planos espaciales superpuestos y que además suene como un instrumento único? La solución llegó mediante un diseño de sonido no convencional donde se utilizaron 32 parlantes auto amplificadores (además de 4 subwoofers –sub bajos- para amplificación de sonidos graves) que representaron 16 fuentes móviles diferentes, en formato de espejo (Figura 9.4).



**Figura 9.4:** configuración en espejo para el movimiento y ubicación de las fuentes sonoras de la obra “Oí(r) el ruido” – Imagen: José Rafael Subía Valdez

Es decir, se podían direccionar todas las señales sonoras provenientes de los micrófonos hacia la consola, procesarlas mediante una computadora, y elegir 16 posiciones diferentes esparcidas en los 45 metros del largo de la estructura, o mover los sonidos por entre esas posiciones simulando movimientos libres. Para esto se debió realizar un análisis previo de las posibilidades, diseñar el sistema de audio y determinar la posición de todos los micrófonos que intervenían en la puesta. Al final se llegó a la solución que muestra en la Figura 9.5, donde se muestra el diseño del patrón de radiación de los parlantes para que no haya huecos en la cobertura de todo el espacio, y así reforzar la idea de un instrumento único.



**Figura 9.5:** *Patrón de radiación del sistema de reproducción de “Oí(r) el ruido” – Diseño: Daniel Hernández.*

La estructura formal de la obra se apoyó este diseño sonoro espacial, ya que la misma contemplaba movimientos coreografiados de músicos y bailarines de tal manera que se aprovechó todo el espacio del pabellón y los diferentes materiales sonoros que en él se presentaban.

#### 9.5.4. Análisis del espacio como estructura de la obra

Esta obra tiene la característica de estar ejecutada en vivo sobre una estructura de metal, madera y telas de plástico (45 metros x 9 metros) y ubicada en un espacio público amplio. Además se trata de una obra músico-coreográfica, en donde el aparato tecnológico aporta el factor de unidad de todas las partes, ya que el sistema de captura y reproducción logra aunar todos los sonidos que se están produciendo. La idea primordial de esta obra fue tomar este gigantesco pabellón y convertirlo en un instrumento único. A pesar de la gran cantidad y variedad de material sonoro que se obtuvo y se reprodujo, es interesante observar que la ubicación de este pabellón en un lugar aun más amplio (una plaza en la Av. Libertador), cumplió con las expectativas.

Este análisis abarcará las 4 escenas desarrolladas de la obra intentando aplicar

la taxonomía desarrollada por Kendall (2010) y en algunos casos extendiendo su aplicación.

La obra presenta desde este punto de vista algunos rasgos interesantes que se observaran con el correr del análisis más precisamente las ideas de RECINTO y CONTENEDOR que maneja y las grandes distancias que tanto los ejecutantes como los espectadores recorren y perciben, lo que hace que la percepción misma de la obra cambie sensiblemente dependiendo de la ubicación en el espacio.

#### **9.5.5. Análisis de los diferentes tipos de FUENTE**

#### **9.5.6. Señales Fuente**

Las Señales Fuente utilizadas son de diversa índole. Se tienen 24 señales acústicas tomadas por 24 micrófonos de diversas características (Figura 6). Estos micrófonos estaban repartidos alrededor y dentro de toda la estructura planteada, y recogían las señales producidas en los caños, maderas, telas y gazebos. Una lista descriptiva de estas señales sería:

4 micrófonos piezoeléctricos en telas de lona.

10 micrófonos piezoeléctricos en parantes de hierro.

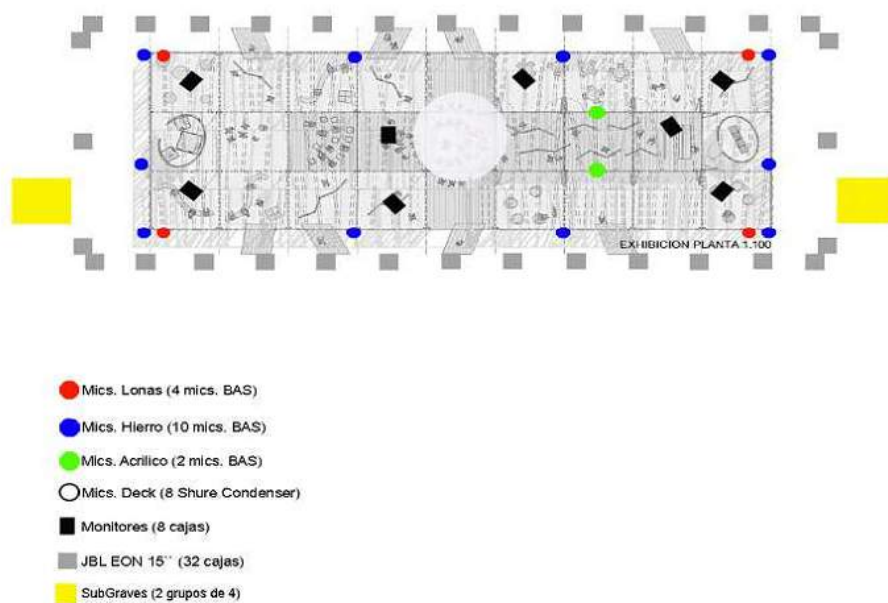
4 micrófonos condensador en deck de madera.

4 micrófonos activos de guitarra en deck de madera.

2 micrófonos piezoeléctricos en gazebos de chapa de cinc.

Otra Señal Fuente utilizada fueron texturas electroacústicas generadas a par-

tir de grabaciones previas sobre los materiales del pabellón y tratados con un granulador (técnica por la cual se divide a un sonido en secciones que luego serán procesadas en forma individual) y un Pitch Shifter (modificación de la altura de un sonido, intentando ajustarlo a una escala musical).



**Figura 9.6:** Planta final del diseño sonoro de “Oí(r) el ruido”. Diseño Daniel Hernández, Mariano Cura y Esteban Calcagno. Imagen: Esteban Calcagno.

### 9.5.7. Fuentes conceptuales

Si bien podría llegarse al acuerdo de que no todas las Señales Fuente utilizadas en esta obra se relacionan con Fuentes Conceptuales, debido a que los sonidos utilizados fueron de carácter concreto y quizás el espectador no pueda generar una vinculación al escucharlos, existe un dato no menos interesante al respecto que puede generar una salvedad. Los materiales constructivos presentes en el pabellón eran simples estructuras de caño de metal, telas plásticas, maderas y chapas de zinc. Los espectadores podían observar sin inconveniente estos

materiales y ver a los ejecutantes tocarlos, y entender que los sonidos que escuchaban eran obtenidos de esos materiales. En este nivel de comprensión los espectadores pueden conceptualizar los sonidos producidos, ya que aunque están reproducidos por parlantes observan y entienden la producción, generando un vínculo inmediato. Esta característica que se da en general en obras donde hay una performance en vivo, podría potenciar la taxonomía de análisis de Kendall, ya que se extendería más allá del sistema de reproducción y servir para analizar escenas más complejas.

Habiendo realizado esta salvedad, se pueden tratar como Fuente Conceptual a los diversos sonidos obtenidos mediante la ejecución de cada participante y que se resumen en la siguiente lista:

- Secuencias agudas y cortas generadas por el golpeteo de los palillos de madera sobre los caños de hierro de la estructura (que además poseían una reverberación propia).
- Secuencias más graves y continuas con una rugosidad característica, obtenidos mediante el uso de boligomas frotadas sobre el hierro.
- Sonidos de frecuencias medias generados al percutir mazos de goma y madera sobre las maderas del piso (deck).
- Secuencias sonoras rápidas y de gran riqueza tímbrica, cuando las telas eran frotadas con las diversas baquetas.

Una fuente conceptual distinta a las antes mencionadas fue el caso de los corales generados a partir de la ejecución vocal sobre los enormes gazebos de zinc. La voz cantada tiene un nivel de conceptualización muy alto, a pesar de

estar modulada por la reverberación del gazebo, y por lo tanto el vínculo es directo. A esta altura, dentro de la noción de Fuente Conceptual planteada por Kendall, se puede hablar quizás de diferentes niveles de vinculación con la fuente, ya que como es el caso de esta obra, se puede hacer una distinción entre una Fuente Conceptual directamente vincular, como las voces cantadas, con las Fuentes Conceptuales generadas a partir de la obtención de sonidos mediante materiales conocidos, donde la vinculación por parte del espectador se da por la información visual obtenida.

#### 9.5.8. Imágenes Fuente

Tanto el sistema de reproducción propuesto como el sistema de espacialización que lo maneja permiten obtener diversas configuraciones de Imágenes Fuente asociadas a Fuentes Conceptuales, cuyas relaciones se analizarán en el apartado donde se tratan los juegos perceptuales propuestos. En esta obra se pueden obtener entonces un conjunto diferenciado de Imágenes Fuente que puede ir desde 16 posiciones diferentes, propuesto como máximo en el sistema de reproducción, hasta un mínimo de 1 sola Imagen Fuente, cuando todas las señales se reproducen en un solo sector del sistema. En el medio se utilizaron configuraciones donde se obtenían 2 y 4 Imágenes Fuente, aunque el sistema permitía casi cualquier configuración.

Quizás una de las características más distintivas en este aspecto, el de las Imágenes Fuente, fue que el sistema de espacialización utilizado diseñado por Subía Valdez (2011) permitía no solamente la ubicación en determinado lugar entre los diferentes parlantes, sino que también podía dividirse todo ese espacio en las partes necesarias y además mover las fuentes de una forma muy

sencilla. A pesar de la posibilidad de tener fuentes móviles, esta opción solo se utilizó durante la reproducción de los montajes electroacústicos, en el resto de la obra solo se utilizaron posiciones fijas. Este hecho acentuó la idea del pabellón como instrumento único, ya que las posiciones fijas potenciaron la idea de que los sonidos salían del pabellón cuando este era “tocado”, mientras que en el caso de los montajes electroacústicos, el pabellón funcionaba como reproductor y no como generador de sonido.

Es interesante agregar que frente al pabellón (a 50 metros, aproximadamente) había una pared de edificios, que devolvía un eco, con bastante amplitud, de lo que generaba el sistema de reproducción, produciendo una Imagen Fuente adicional, repetición de los que se escuchaba al lado del pabellón. Este eco llegaba aproximadamente unos 145ms más tarde, y era más evidente para los espectadores que estaban del lado del pabellón próximo a los edificios, siendo no tan evidente para los que estaban del otro lado. Lo interesante es que ese eco era parte de la obra: desde un primer momento se tuvo en cuenta que iba a suceder ese fenómeno y aunque no se podía controlar en forma total en algunos casos se eligió direccionar algunos parlantes hacia los edificios para incrementar el eco. El uso de esta Imagen Fuente acústica expande la noción de Imagen Fuente propuesta por Kendall, que ya es un subproducto de las variables del entorno donde se reproduce la obra.

Por último hay que hacer una salvedad con respecto a este análisis sobre las Imágenes Fuente. En este caso en particular se tiene que tener en muy en cuenta la ubicación del espectador, ya que un espectador a una distancia muy cercana al pabellón (por ejemplo, a 2 metros) no podía percibir las diferentes Imágenes Fuente aquí propuestas, debido a que seguramente tanto su audición



como su atención serían “atrapados” por el parlante más cercano, sin dejar lugar a la apreciación de la totalidad. Así también un espectador muy lejano (por ejemplo, a 40 metros), percibía la estructura como un punto, sin posibilidad de escuchar las variaciones internas de la misma. Por esto se decidió que los espectadores debían estar a una distancia entre los 10 y los 20 metros de la estructura.

#### 9.5.9. Niveles de organización

Podemos observar diversos niveles de organización y variaciones en cada uno de esos niveles. Empezando por el nivel de organización más alto que propone Kendal, el de ESCENA, es interesante intentar explicar lo que ocurre en “oí(r) el ruido”, desde el punto de vista que no existe un contenedor convencional, ya que la ESCENA se planteó en una plaza, al aire libre, pero con características propias como la presencia de una avenida, y una pared de edificios de en uno de los lados. El nivel ESCENA plantea entonces una dificultad extra, y nuevamente depende de la posición del oyente. Esta quizás es una dificultad en el planteo de Kendall, ya que en una obra al aire libre se torna dificultosa la idea de plantear los límites del nivel de ESCENA. En este caso se puede plantear una ubicación determinada, pero la obra no posee límites claros en cuanto a este nivel de organización. En este caso se plantea un sector de la Ciudad como contenedor, sin límites posibles.

En cuanto a los niveles FUENTE y ENSAMBLE, también existen condiciones. Como se aclaró antes, el pabellón funcionó como múltiples FUENTES, pero que a la vez, al llegar a entenderlo como instrumento, puede ser considerado como ENSAMBLE. En el caso de las FUENTES, cuya configuración varía

según cada escena, el factor de la distancia de cada una de estas es sensible. El atributo espacial distancia, es quizás el más importante en esta obra. Es interesante aclarar un punto que se relaciona con la idea de CONTENEDOR planteada por Kendall y que ofrece un juego perceptual interesante. Cuando se produce el coral de voces sobre los gazebos de chapa, la Señal Fuente producida lleva impresa las características acústicas de reverberación metálica propia de ese espacio y esos materiales. Esa señal se transforma luego en una Imagen Fuente, que se reproduce en un RECINTO que contiene a la ESCENA, que en realidad no existe, o por lo menos sus márgenes no están claros. Así tenemos un juego perceptual organizativo muy interesante, donde en una ESCENA donde no hay RECINTO claro, se utiliza las cualidades de otro RECINTO, que además se encuentra dentro de la ESCENA: gazebo de chapa de zinc, dentro del pabellón, dentro de ese espacio de la Ciudad.

#### **9.5.10. Juegos con los agrupamientos perceptivos**

A continuación, se procederá al análisis de las 4 escenas de “Oí(r) el ruido” basándonos en la clasificación de juegos perceptivos de Kendall (2010).

##### **ESCENA 1:**

En la primera escena de la obra se plantea un juego perceptivo basado en la multiplicidad de Fuentes Conceptuales, propuestas por la ejecución de estructuras sonoras sobre los parantes de hierro, realizada por los músicos y bailarines que se ubican en el perímetro del Pabellón. En esta escena se plantea que las múltiples Fuentes Conceptuales se relacionen con múltiples Imágenes Fuente proporcionadas por los 16 parlantes funcionando con señales indivi-

duales. En esta caso, existe una correlación directa entre la información visual y la sonora, ya que la ubicación de los sonidos a lo largo del pabellón respeta la ubicación de los ejecutantes. Un rasgo interesante a describir es que en esta escena se hace un gradiente entre estructuras sonora agudas hasta estructuras con contenido espectral grave pasando por materiales que cubren el espectro completo. Esto genera una sensación de acercamiento-alejamiento de las fuentes sonoras debido a indicios espectrales (Capítulo 3, localización en distancia de fuentes sonoras) buscada directamente en la composición.

Es importante aclarar que sin duda la percepción de las relaciones entre Fuentes Conceptuales e Imágenes Fuente dependía de la posición de cada espectador. Al espectador de un extremo de la estructura le llegaba el sonido directo de la fuente más cercana y además, retrasadas en el tiempo (la fuente más lejana le llegaba 100 milisegundos más tarde), otras 15 fuentes, que además estaban atenuadas en amplitud por la distancia recorrida (aproximadamente unos 50dB SPL con menor amplitud que la primera fuente). Agregando el eco producido por los edificios, la información sonora que recibía cada uno de los espectadores, era una mezcla de varias Imágenes Fuente más cercanas, las fuentes retrasadas por la distancia a la que estaban separados los parlantes y el sonido que se reflejaba en los edificios, provocando un retraso sonoro físico importante, y generando una sensación de envolvimiento completa. En la obra no se renegó de esta mezcla de sonidos provenientes de varios puntos, y producto de las grandes distancias recorridas, sino que en algunos momentos de actividad seguida de silencio, funcionaba muy bien como información extra: las Fuentes Conceptuales regresaban procesadas por las características

arquitectónicas de este sector de la ciudad.

Analizando esta escena, se da cuenta entonces que en realidad la correcta forma de indicar esto basándonos en la taxonomía de Kendall es decir que se contó con múltiples Fuentes Conceptuales (además replicadas en espejo), pero las características del espacio público (el rebote con los edificios sumado a la gran distancia de los parlantes ubicados en las estructuras) generaban más de una Imagen Fuente, o por lo menos cada Imagen Fuente era aun más compleja que la que Fuente Conceptual.

Escena 2:

En esta escena se produce un cambio importante en cuando a la espacialización ya que solamente se utiliza la mitad de los parlantes del sistema de reproducción, produciendo que la imagen sonora se angoste.

Esto produce una nueva sensación de distancia pero solo para los espectadores que se encuentran lejos de los parlantes que están encendidos.

También cambia la relación entre las Fuentes Conceptuales y la imagen Fuente ya que hay músicos tocando figuras rítmicas en los parantes de metal y el gazebo de zinc, pero todos salen por los parlantes de la mitad del pabellón. Aquí el juego perceptivo se reduce a múltiples Fuentes Conceptuales - una Imagen Fuente, muy amplia por supuesto, pero una sola imagen al fin.

Sobre el final de esta escena la actividad de los músicos cesa y comienza a escucharse las grabaciones procesadas con granulador y pitch shifter. Es el único momento de la obra donde hay fuentes móviles, y son justamente grabaciones electroacústicas y no tomas en tiempo real. En ese momento se rompe nuevamente el esquema espacial para utilizar todos los parlantes, pero en es-

te caso se logran ubicaciones intermedias gracias al sistema de espacialización.

Escena 3:

Luego de las fuentes móviles, se vuelve a la producción de los ejecutantes, estos se ubican junto a los bailarines dividiendo al pabellón en 8 partes iguales, comenzando a realizar una coreografía junto a los bailarines, donde se ejecutan rítmicas lentas y sincopadas sobre el piso de madera, utilizando baquetas de madera.

Lo interesante de esta escena es nuevamente el juego perceptual. A cada uno de los ejecutantes de estas rítmicas sobre el suelo de madera se le asignan dos parlantes del sistema de reproducción obteniendo así una configuración múltiple entre Fuente Conceptual individual - múltiples Imágenes Fuente. Esto también juega nuevamente con la percepción de distancia, ya que cuando una fuente sonora se ensancha da la sensación de que se acerca.

Sobre el final de esta sección, algunos músicos se acercan a los gazebos y comienzan a cantar sobre estos. En este caso a cada voz, Fuente Conceptual, se le asignan 4 parlantes, por lo que las Imágenes Fuentes vuelven a ensancharse ocupando casi 12 metros de ancho cada una, armando un coro gigantesco. Lo interesante aquí es que los cantantes están a la izquierda del público pero se escuchan en la totalidad del pabellón. Si bien la relación entre Fuentes Conceptuales e Imágenes Fuentes es bastante sencilla (una Imagen Conceptual - una Imagen Fuente) es interesante pensar que dentro de la taxonomía que maneja Kendall podría haber un lugar para el análisis de cómo las Imágenes Fuentes afectan algunos atributos espaciales de las Fuentes Conceptuales, como es el caso del ancho o la distancia de los mismos.

Sobre el final de esta sección, se vuelven a reproducir montajes electroacústicos pero esta vez de grabaciones de las voces del coral. En ese momento se mezclan las voces en vivo con las grabaciones, pero luego estas últimas comienzan a moverse, y los cantantes dejan de cantar.

Escena 4:

En la última escena, los músicos se colocan nuevamente sobre la mitad izquierda del pabellón y ejecutan un motivo rítmico sobre los parantes de hierros, piso de madera y gazebo de cinc. Se utiliza todo el sistema de reproducción pero dividiendo los materiales a lo largo del pabellón. De esta manera cada material (metal, madera, zinc) ocupa un lugar en el espacio. Nuevamente se repite la configuración Fuente Conceptual individual - múltiples Imágenes Fuente.

## § 9.6 OCUPACIÓN INVISIBLE

### 9.6.1. Descripción General

“Ocupación Invisible” se estrenó dentro de la muestra Post it City - Ciudades Ocasionales (<http://www.ciutatsocasionals.net/>) en Buenos Aires en la Casa de la Cultura (Av. De Mayo 575), el 7 de julio de 2010 y en Montevideo, en el Espacio de Arte Contemporáneo (Arenal Grande 1930), el día 5 de octubre de 2010.

Esta muestra itinerante, acostumbrada a presentar trabajos de artistas visuales dentro de la fotografía y el video documental, los cuales reflejan la búsqueda de prácticas urbanas que aparecen, se desarrollan y desaparecen dentro de las ciudades. El espacio público es optimizado y parasitado por las prácticas emergentes, que, en la mayoría de los casos desaparece luego de un tiempo. En este contexto, “Ocupación Invisible” aparece como la única obra de la muestra puramente sonora, lo que a la vez la convierte, dentro de la muestra, en ocupante de un lugar ocasional. Se trató de una instalación sonora, autónoma e interactiva donde la idea giró en torno a las radios “clandestinas” de la Ciudad de Buenos Aires. La reseña de la obra indica textualmente:

“Denominados ‘pirata’ por los empresarios radiales, los proyectos político-culturales alternativos, populares y comunitarios de estas estaciones, conforman una pluralidad de voces con diversos ejes temáticos, en su mayoría sin fines de lucro y al borde del ruido. Su principal destinatario es la comunidad barrial que los rodea y sus emplazamientos permiten inferir la distribución

socio-económica de la ciudad. Este tipo de ocupación se caracteriza por lo ingrátido de su materia y lo difuso de sus límites, que dependen tanto de la potencia como del cruce con otras emisoras de frecuencia cercana.”

Por lo tanto, los materiales sonoros con los cuales se trabajó en Ocupación Invisible fueron grabaciones de radios clandestinas que invisiblemente ocupan un espacio, el del espectro radiofónico, de la ciudad. Mediante grabadores mp3, se recogió material de varias decenas de radios clandestinas porteñas, que luego fueron editadas para recoger diferentes temáticas y poder contar con material narrativo adecuado. Ocupación invisible se trató entonces de una instalación sonora interactiva. Instalación por que ocupó, durante el tiempo que duró la muestra, un espacio determinado. Sonora, porque el principal (y casi el único) elemento perceptual utilizado fue el sonido. Interactiva, por que se trabajó con una grilla de sensores que podían seguir el movimiento de los espectadores que circularan por allí, intentando simular la idea de Dial Humano.

#### **9.6.2. Diseño del espacio, sistema de reproducción y espacialización**

El espacio para “Ocupación Invisible” fue diseñado desde cero. Para esto se pensó desde un primer instante en una sala vacía y oscura donde se pudiera aprovechar la información sonora con la intervención de la menor cantidad de información visual. Para esto se partió de una sala de aproximadamente 6X4 metros construida con paneles de yeso, la cual tenía dos aberturas por donde podían entrar o salir los espectadores.

La instalación se pensó desde la perspectiva de poder sentir el movimiento de de las personas en este espacio visualmente vacío, pero sonoramente lleno



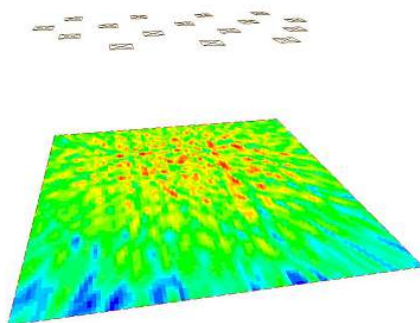
y, mediante esta información, disparar determinados montajes sonoros con un mayor o menor grado de procesamiento, que dependieran exclusivamente de la posición del espectador. Para esto se ideó un sistema que complementó un sistema de reproducción sonora, un sistema de sensado de movimiento y criterios de espacialización de fuentes sonoras.

### 9.6.3. Sistema de Reproducción

Para el sistema de reproducción debieron hacerse varias simulaciones y pruebas previas. Lo primero fue calcular la cantidad de canales de audio con los que contábamos. En total con el equipamiento adecuado llegamos a 16 canales de reproducción de audio independientes. Esto se logró gracias a la utilización de una placa de sonido MOTU mk1-828, la cual posee 8 salidas y a la cual se pudo añadir 8 entradas más mediante el acoplamiento óptico de un módulo pre amplificador de 8 entradas Behringer Pro-8 Ultragain ADA8000. Una vez que se definió la cantidad de canales independiente de salida, se pasó a la elección de los parlantes, donde buscamos una solución direccional, de potencia media y que sea fácil de montar. Elegimos para esto 16 parlantes Pioneer TS-A1673S, de tres vías, que presentaban un patrón bastante direccional en el rango de frecuencias utilizado. Para observar cual sería era la mejor forma de disponer los parlantes en el techo se procedió a realizar una simulación con el programa cadP2 (<http://www.cinemattech.info/manuals/Sound/Speakers/JBL/pages/software/cadp2.htm>) el cual permite realizar modelado electroacústico a partir de un array de parlantes. Se probaron diversas configuraciones con diferentes resultados. A continuación un resumen de las simulaciones realizadas por Daniel Hernández, asesor técnico de Buenos Aires Sonora:

### Disposición Hexágono Truncado

Los parlantes se ubicaron en una figura que se asemeja a un hexágono truncado y solapado. Se logra una distribución de fase más aleatoria con direccionalidad puntual, pero no funcionando para todo el espacio (Figura 9.7).



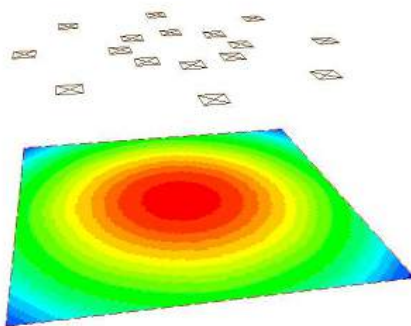
**Figura 9.7:** Diagrama de Radiación de la Disposición Hexágono Truncado – los colores indican un máximo Nivel de Presión Sonora (rojo – 115 dBSPL), hasta un mínimo Nivel de Presión Sonora (azul - 40dBSPL) Diseño y simulación: Daniel Hernández.

### Disposición Doble Octofonía Radial

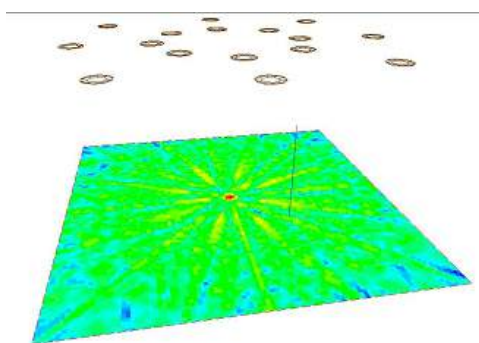
Tenemos dos círculos octofónicos cuyos ejes coinciden de manera que se observa una direccionalidad mucho más marcada al centro de la sala, de manera tal que toda la energía estará puntualizada en ese sector. Esto es una complicación porque lo que se buscaba para este proyecto específicamente era una direccionalidad independiente de cada parlante. Esto definir trayectorias circulares y rectilíneas con mayor precisión si coinciden los ejes de las octofonías, y aparenta tener menor cantidad de cancelaciones de fase (Figura 9.8). Si los ejes de las octofonías no coinciden, observamos que el sweet pot (punto focal) se hace más pequeño (Figura 9.9).

### Disposición Cuadrícula

Mediante esta disposición, que semeja una matriz de parlantes (Figura 9.10),



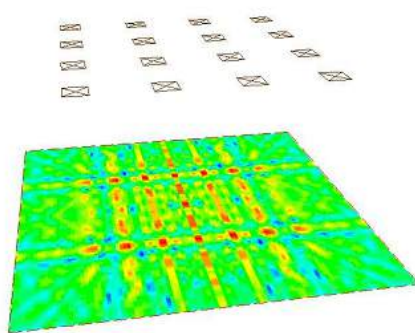
**Figura 9.8:** *Diagrama de Radiación de la Disposición Doble Octofonía Radial con Ejes Coincidentes – los colores indican un máximo Nivel de Presión Sonora (color rojo – 115 dB SPL), hasta un mínimo Nivel de Presión Sonora (40dB SPL) Diseño y simulación: Daniel Hernández*



**Figura 9.9:** *Diagrama de Radiación de la Disposición Doble Octofonía Radial con Ejes no Coincidentes – los colores indican un máximo Nivel de Presión Sonora (color rojo – 115 dB SPL), hasta un mínimo Nivel de Presión Sonora (40dB SPL) Diseño y simulación: Daniel Hernández*

se obtienen enfoques más direccionales y repartidos en toda la sala. Este es uno de los comportamientos que se buscaba para que cada sujeto escuche exactamente el parlante que tiene encima, y el diseño sonoro pueda realizarse de manera integral en cada una de las 16 fuentes, o sea tener 16 fuentes conceptuales puntuales. Como indicó el informe técnico propuesto por Daniel Hernández, esta era la mejor configuración para mantener un diseño sonoro direccional pero generaba una mayor complicación a la hora del armado de fuentes virtuales por las cancelaciones no tan aleatorias que se observan en otras configuraciones.

En definitiva la cuadrícula solucionaba el problema de la direccionalidad pun-



**Figura 9.10:** *Diagrama de Radiación de la Disposición Cuadrícula – los colores indican un máximo Nivel de Presión Sonora (color rojo – 115 dBSPL), hasta un mínimo Nivel de Presión Sonora (40dBSPL) Diseño y simulación: Daniel Hernández*

tual, y la de Fuentes Conceptuales perfectamente ubicadas en el espacio. Por lo tanto, el techo fue construido como una cuadrícula de parlantes ubicados de forma equidistantes.

#### 9.6.4. Sistema de sensado

La idea principal del sistema de sensado fue poder lograr que cada uno de los espectadores que ingresaban al espacio vacío se transformara en un dial humano, es decir que el movimiento del cuerpo fuera capaz de sintonizar cada uno de los montajes realizados con las grabaciones de las radios, y que estaban dispuestos en un lugar específico. Además, también se pensó en la emulación de una característica muy conocida del dial de radio, y que es justamente el hecho de que se debe buscar la emisora pasando por muchos estados donde se escucha ruido pero no sonido claro. Debido a la complejidad de utilizar cámaras de video para sensar el movimiento, específicamente por la poca altura del recinto, se procedió a utilizar un sistema de dos sensados diferentes: uno que determinara exactamente la posición de los espectadores (lo que dimos a llamar sensado fino), paralelamente, por que justamente la idea era que si

había más de un espectador cada uno escuchara el montaje específica de cada lugar, y otro sistema que detectara en qué forma se iba desplazando cada uno (sensado grueso). Se utilizaron entonces, sensores de distancia para detectar la ubicación, y sensores de movimiento para detectar hacia donde se movía cada espectador.

Los sensores de distancia que se utilizaron fueron unos Sharp GP2Y0A02YK los cuales proporcionan una lectura continua de la distancia de un objeto entre los 20 y los 150 cm. Estos sensores fueron ubicados en la carcasa de los parlantes que estaban en el techo (Figura 9.11), justo en el centro de los mismos, apuntando al piso para así, de esa manera, detectar exactamente cuando un espectador estaba en ese lugar. Al fin y al cabo el sensor funciona como barrera infrarroja entre el techo y el piso asegurando la detección.



**Figura 9.11:** *Se observa un parlante ubicado en el techo de la instalación con el sensor Sharp adosado, ambos apuntando hacia abajo.*

El sistema de sensado grueso fue diseñado a partir de sensores piroeléctricos (sensores ópticos de calor) de movimiento, los denominados sensores PIR. Estos sensores tienen la capacidad de detectar si alguien se estaba moviendo en un perímetro determinado y ajustable. Además, poseen un tiempo de reacción ajustable, lo cual permitía generar un determinado retardo con espectro

al movimiento, el cual fue utilizado para generar una estela sonora.

En la Figura 9.12, puede observarse un prototipo del sistema de sensado y reproducción, donde observamos tres parlantes colgados, cada uno con sus sensores Sharp, y en sus alrededores, conformando un cuadrado, los sensores PIR.

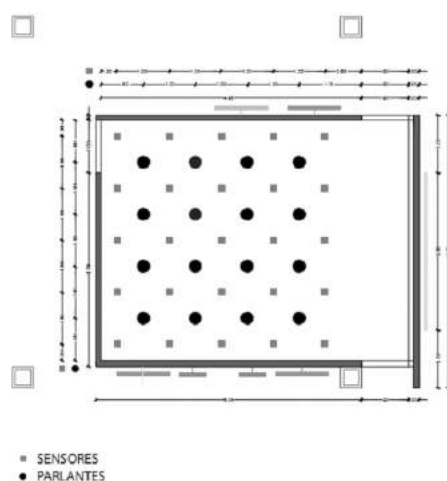


**Figura 9.12:** *Prototipo del sistema de reproducción y sensado de Ocupación Invisible. Se observa claramente la ubicación de los parlantes con sus sensores, y los sensores de movimiento (cubos blancos) que los rodean.*

En total se utilizaron 41 sensores, ubicados específicamente en todo el espacio como indica la Figura 9.13. Los círculos negros son los parlantes y sensores de posición, y los cuadrados rojos son los sensores de movimiento.

En una siguiente instancia se decidió la forma en que la información de todos estos sensores se procesara de tal manera que pueda ser útil al diseño. Para esto se recurrió a la plataforma de Hardware Abierto Arduino (<http://www.arduino.cc/es/>) la cual permite procesar varias entradas digitales y analógicas, las cuales pueden manejar los diferentes valores de tensión que entregaban los sensores. Además, este procesamiento se puede hacer de forma paralela, y por lo tanto se pueden utilizar los datos por separado. Pero lo más interesante de utilizar Arduino, es su capacidad de poder conectarse fácilmente a la PC y poder comunicarse con casi cualquier programa, entre ellos Pura Data. Pura Data (<http://puredata.info/>) es una herramienta

informática de código abierto que puede procesar información en tiempo real de varios medios, pero específicamente se utiliza para el procesamiento de audio. Pura Data puede, mediante comunicación con el puerto serie de la PC, obtener la información que Arduino a su vez obtiene de los sensores, y de esta manera reproducir determinados audios o procesarlos según el sensor, o los sensores, que estaban activos en determinado momento, y enviar información sonora a los parlantes ubicados en la sala. De esta manera Se pudo resolver fácilmente todo el array sensores – arduino – PC – Reproducción de audio.



**Figura 9.13:** *Diseño final del espacio de Ocupación Invisible. Diseño e Imagen: Hernán Kerlleñevich*

Sin embargo las cosas no se resolvieron tan fácilmente, ya que había que determinar una lógica de interacción, definida por la posición en el espacio de los espectadores.

#### 9.6.5. Diseño de interacción

Es importante mencionar algunas características del diseño de interacción de la obra debido a que está íntimamente ligado al diseño espacial y al uso del sonido en el espacio.

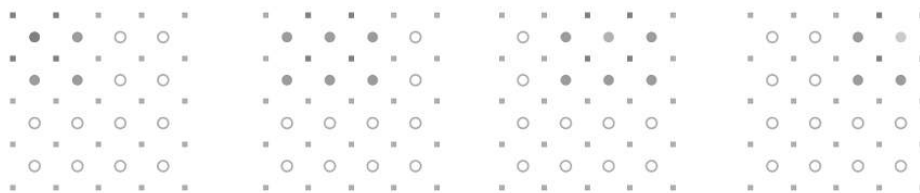
En primera instancia el diseño presentó comportamientos diferenciados. Co-

mo explicamos antes, la acción más simple que podía realizar la instalación era la detección de uno o varios espectadores bajo cada uno de los parlantes. De esta manera el espectador posicionado en un lugar en el espacio, podía escuchar un montaje sonoro directamente sobre su cabeza. Si se movía hacia otra posición de parlante, podía escuchar otro montaje sonoro. Este comportamiento, permitía en una primera instancia, determinar que montaje sonoro escuchaba un espectador y en qué lugar determinado era escuchado. La información sonora ocupaba un espacio determinado y por lo tanto, siempre que alguien se ubicaba en ese sector, escuchaba exactamente el mismo tipo de montaje. De esta manera esta instancia permitía un desarrollo espacial de las fuentes conceptuales, directamente relacionadas a las fuentes físicas. Una segunda configuración respondía a una interacción más compleja, ya que tenía en cuenta la cantidad de espectadores que había dentro de la instalación y en qué posición se encontraban. Entonces, cuando solamente había espectadores ubicados bajo cada uno de los parlantes, estos podían escuchar el montaje establecido para cada lugar de forma individual. La complejidad se producía cuando las personas se acercaban a un mismo punto de escucha, ya que los sensores puestos a modo de cerco, a medio metro de cada parlante, se iban activando y producían una interferencia en el montaje que se escuchaba en el parlante. Revisemos un poco cuales podían ser las combinaciones básicas de espectadores que interactuaban en el espacio.

Interacciones locales:

Los espectadores ubicados bajo un parlante, encendían un determinado montaje sonoro de forma limpia y sin ningún procesamiento.





**Figura 9.14:** *Diseño e Imagen: Hernán Kerlleñevich*

Interacciones zonales:

Si había más de un espectador dentro de un determinado rango, se activaban los parlantes que se encontraban dentro de ese rango, pero con un procesamiento que influía directamente en el entendimiento del audio, esta instancia fue llamada de interferencia.



**Figura 9.15:** *Diseño e Imagen: Hernán Kerlleñevich*

Los comportamientos de posibles combinaciones de espectadores interactuando entre ellos están basados en la Vecindad de Margolus (Margolus, 1984), el cual define reglas específicas de autómatas celulares. Estas reglas definieron 16 estados diferentes dependiendo de cuáles y cuantos sensores PIRs se encontraban activados en la interacción con los espectadores. Un diagrama simple de estos comportamientos se muestra en la Figura 9.16.

Interacción Global:

Esta interacción sucedía cuando no había nadie en el espacio, y en respuesta a esta situación se reproducía es una secuencia de voces que indicaban el dial y el nombre de algunas de las radios utilizadas, también desplegadas en el

1	1	0		2	0	1		3	0	0		4	0	0
	0	0			0	0			0	1			1	0
5	1	1		6	0	1		7	0	0		8	1	0
	0	0			0	1			1	1			1	0
9	1	1		10	0	1		11	1	0		12	1	1
	0	1			1	1			1	1			1	0
13	1	0		14	0	1		15	1	1		16	0	0
	0	1			1	0			1	1			0	0
REGLAS														

**Figura 9.16:** *Diseño e Imagen: Hernán Kerlleñevich*

espacio de la instalación.

#### 9.6.6. Análisis

Se procederá a realizar el análisis de esta obra utilizando la estrategia de análisis establecida por Kendall (2010) y desarrollada en el Capítulo 4, enfoques analíticos de la espacialidad en la música electroacústica.. La idea de esta obra fue emular una especie de “dial radiofónico humano”, en donde, como se explicó antes, los espectadores recorrieran un espacio vacío y oscuro y “sintonizaran con su cuerpo” diferentes montajes sonoros realizados a partir de grabaciones de radios clandestinas de la Ciudad de Buenos Aires. La estructuración temporal espacial de “Ocupación Invisible” dependió exclusivamente de las decisiones que tomaran los espectadores que recorrían la muestra. De esta manera, un espectador recorría la instalación, ejecutando los diversos montajes sonoros, los cuales ocupaban un lugar específico del espacio. Si bien la ubicación de las fuentes sonoras en el espacio estaba preestablecida, ya que cada parlante tenía definido uno de los montajes sonoros, de una temática específica, la percepción del espacio podía cambiar a partir de los

diferentes juegos perceptivos que ocurrían debido a las decisiones de recorrido de cada espectador, y también debido a la interacción ocurrida cuando varios espectadores recorrían la instalación al mismo tiempo. Los juegos perceptivos más interesantes estuvieron dados por el hecho de que cada montaje ocupaba un lugar específico en la sala. De esta manera los espectadores, a los cuales no se les daba indicaciones a priori, podían recorrer una vez el espacio, aprender y memorizar como se distribuían los montajes y realizar un nuevo recorrido, creando cada vez una nueva trayectoria, ligada especialmente al contenido conceptual de las fuentes sonoras, ya que cada montaje representaba una temática predeterminada. Un sistema de sensado (véase la descripción técnica en el apartado Sistema de Sensado de este capítulo) permitía diversos niveles de “sintonización”, donde solamente posicionamientos en ubicaciones exactas del espectador permitían que los registros de audio se escucharan de forma clara, de lo contrario aparecían interferencias sobre los sonidos. Se debe agregar que la interacción entre varios espectadores también producía el disparo de ciertos registros de audio y también la producción de interferencias. Como cabe indicar, la obra no estaba dividida en escenas, ni tenía un tiempo mínimo ni máximo de duración<sup>3</sup>, por lo tanto es imposible analizar, como si se hizo con las otras obras, una escena o la obra completa. Por esto, lo que se realizará es una descripción de las diferentes Señales Fuente, Fuentes Conceptuales e Imágenes Fuente que se presentaban en la obra, y por supuesto los diferentes juegos perceptivos que aparecían en los diversos recorridos que podían elegirse.

---

<sup>3</sup>Se trata, más propiamente hablando, de una instalación sonora

### 9.6.7. Análisis de los diferentes tipos de FUENTE

### 9.6.8. Señales Fuente

El material sonoro de la instalación estuvo compuesto por grabaciones de diversas radios de la Ciudad de Buenos Aires. Las grabaciones de estas radios se llevaron a cabo mediante grabadores y reproductores portátiles de mp3, los cuales tenían la posibilidad de sintonizar radios AM y FM y a la vez poder grabarlas mientras se realizaba la escucha. Con estos grabadores, se recorrió la Ciudad, buscando los puntos en donde se indicaba que cada radio era transmitida. Es interesante aclarar que muchas de estas radios tenían una señal que no pasaba de los pocos cientos de metros, y que en muchos casos varias radios ocupaban el mismo espectro de la señal, y solo se diferenciaban por unos pocos metros. De aquí también salió la idea del diseño de interacción. Una vez que se obtuvo este material sonoro, se comenzó la etapa de escucha, selección y edición del material a utilizar. Esto marcó profundamente la estética de la obra, ya que al final de este proceso se terminaron eligiendo unos cientos de archivos de audio por sobre miles, los cuales fueron elegidos por el interés musical sonoro que representaban. Los registros seleccionados fueron utilizados para realizar 16 montajes sonoros, con una temática determinada, lo que se observa en la Figura 9.17.

En esta tabla se muestra el número de canal de audio, la temática del montaje, y el contenido específico de cada montaje. Así, por ejemplo, si alguien activaba el sensor que estaba en el parlante 6, escuchaba grabaciones de radios de comunidades internacionales, específicamente las publicidades que éstas presentaban en sus programas. Lo interesante de esto, es que si en la instalación

confluían varios espectadores que activaban diversos sensores, las temáticas convivían en forma simultánea, produciendo una interferencia tanto espacial como sonora y disparando nuevas Señales Fuente, que se reproducían en los parlantes asignados. Por lo tanto se contó con 16 Señales Fuente, provenientes de archivos de audio y reproducidos por un sistema de 16 canales de audio (ver Sistema de Reproducción).

Canal de Audio	Temática	Contenido
1	Religiosa	Música; pastor evangélico
2	Deportiva	Discusión sobre fútbol
3	Social	Plantación de árboles
4	Religiosa	Música católica
5	Espiritual	Curanderos/sanadores
6	Comunidades Internacionales	Publicidades
7	Espiritual	Sexo tántrico
8	Religiosa	Pastora católica
9	Musical	Publicidad de boliches
10	Social	Salud
11	Religiosa	Música; pastor evangélico
12	Comunidades Internacionales	Música; ayuda social
13	Comunidades Internacionales	Relato de fútbol; bolsa de trabajo
14	Espiritual	Mecánica cuántica; problemas vocacionales
15	Religiosa	Música; rezos
16	Religiosa	Pastora católica; pastora evangélica

**Figura 9.17:** *Materiales sonoros utilizados en Ocupación Invisible. Listado proporcionado por Pablo Chimenti.*

### 9.6.9. Fuentes conceptuales

Las fuentes conceptuales están determinadas por una gran diversidad de sonidos: música, voces habladas, voces cantadas, spot radiales, publicidades. Lo interesante es que al tratarse de grabaciones radiales, es casi instantánea la conexión que un oyente puede hacer con los sonidos. La conexión de los espectadores con el concepto de la fuente sonora esta directamente asociada al sonido radial. Se trata de espacios mediatizados (según Smalley, 2007). Todos los sonidos que se reproducían eran Fuentes Conceptuales, por lo que era difícil

contabilizar cada una de las Fuentes Conceptuales que aparecían en la obra, debido a que los montajes sonoros establecidos como Señales Fuente duraban aproximadamente 60 minutos cada uno, si eso lo multiplicamos por 16 canales y además añadimos los procesamientos, la tarea de contabilizar cada una de las Fuentes Conceptuales es posible, pero infructuosa.

Sin embargo se puede establecer, como en el caso de “Oí(r) el ruido”, una clasificación dentro de las Fuentes Conceptuales generada a partir de los niveles de interferencia que aparecen cuando la “sintonización” no es precisa o cuando varios espectadores se interfieren entre sí. Cuando se producía algún tipo de interferencia por alguno de estos sucesos, el sonido de las Señales Fuente era procesado mediante modulación AM, efecto especialmente utilizado para emular los sonidos de una mala sintonización en una radio AM, y una ligera granulación, lo que también creaba una sonoridad muy conocida para alguien que alguna vez haya escuchado una radio. Por ejemplo, la voz de un locutor puede verse modulada de tal manera de que exista todavía una conexión entre el oyente y ella, pero con un grado de procesamiento que hace que la conexión cambie: sigue siendo la voz de un locutor radial, pero mal sintonizada. Se vuelven a establecer, entonces, niveles diferentes de Fuentes Conceptuales, lo que hace que la clasificación de Kendall se extienda aun más.

Por lo tanto se tienen 16 montajes sonoros, los cuales contienen una gran cantidad de Fuentes Conceptuales reconocibles (en su mayoría voces y música) y que además en determinadas situaciones aparecen con un nivel de procesamiento que corre la conceptualización de esos sonidos, pero sin perder la idea conceptual.

### 9.6.10. Imágenes Fuente

Las Imágenes Fuente que podían percibir los espectadores de la obra dependían nuevamente del recorrido que estos realizaran dentro de la instalación. Por lo pronto se pueden especificar como Imágenes Fuente a los sonidos que reproducen cada uno de los parlantes ubicados en el techo. Así, por lo pronto se contaría con 16 Imágenes Fuente. Si bien el sistema de reproducción establece posiciones fijas de sonidos no permitiendo movimientos ni trayectorias, es interesante observar como este supuesto estatismo de las Imágenes Fuente se rompe cuando los espectadores recorren la muestra. Existe un efecto interesante que vale la pena describir. El sistema de sensado diseñado para la obra presentó un retraso en la respuesta que al principio parecía ser un problema, pero que luego se transformó en una oportunidad de reforzar la inmersión de los espectadores dentro de la muestra. Este retraso al sensar la ubicación de los oyentes, producía una especie de “aureola” sonora que daba la sensación de que el oyente estaba dejando un rastro a su paso. Debido a que cuando un espectador pasaba rápidamente por los sensores, la “sintonización” no era exacta, este dejaba una estela de sonido AM (al que además se agregaba una leve granulación), generando la sensación de que, en realidad, existía una fuente virtual en movimiento: el dial humano sintonizaba espacialmente.

### 9.6.11. Niveles de organización espacial

Los niveles de organización espacial de “Ocupación Invisible” se pueden dividir en FUENTE, ENSAMBLE, SALA y ESCENA. Se pueden determinar entonces las 16 FUENTES a partir de lo que sale de cada parlante ubicado en el techo de la instalación. Asimismo en determinadas situaciones, por ejemplo

cuando 4 o más espectadores estaban recorriendo el lugar y coincidían en 4 parlantes que formaban un cuadrado, se establecía un ENSAMBLE formado por la cuadrafonía ubicada en el techo, que disparan con sus cuerpos. Otro aspecto interesante se producía por el hecho de que esta instalación está pensada tanto para ser escuchada por oyentes que están dentro del cuarto de los altoparlantes, como fuera del mismo.

En ese sentido, es obvio que la sala que contiene la instalación misma cumple función de SALA solo cuando los oyentes estaban dentro. Al hablar de ESCENA la cuestión es un poco más elaborada. Nuevamente la ubicación del oyente era fundamental, ya que este estaba casi siempre en movimiento entrando y saliendo de la SALA. El nivel ESCENA puede percibirse desde dentro de la instalación, y desde fuera de la instalación. Si un espectador está por fuera de la instalación, a unos metros de la SALA, escucha una especie de ENSAMBLE, cuyo ancho disminuye a medida que nos alejamos de él. Si el espectador se acercaba a la SALA, y escuchaba desde la puerta de la misma, la ESCENA se ensanchaba, y comenzaba a estar inmerso en la misma. Un último paso era escuchar la ESCENA, ya dentro de la SALA. Por lo tanto el oyentes podía apreciar la ESCENA tanto desde afuera de la SALA, escuchando el sonido generado por los oyentes que estaban dentro de ella, o generar el mismo los sonidos ubicándose dentro de la SALA. Se puede hacer un paralelismo con un oyente que escucha a una orquesta y un oyente que decida ubicarse entre los músicos de la orquesta, aunque en este caso no es el mismo el que genera el sonido.

Si el oyente está lo suficientemente lejos de la SALA, escuchará una FUENTE (cuando un oyente se aleja de un sonido, el ancho de este disminuye), a me-



didada que se acerca a la SALA, esta se parecerá más a un ensamble (en el caso que haya más de una persona en la instalación), cuando el oyente esta dentro de la SALA, ya puede generar sus sonidos y la generación de los niveles de organización espacial depende de sus movimientos.

#### **9.6.12. Juegos con los agrupamientos perceptivos**

Para hablar sobre los juegos perceptivos que se pueden encontrar dentro de “Ocupación Invisible” es conveniente también desarrollar las ideas y características espaciales propias que establece la obra.

El espacio de esta obra fue pensado, diseñado y construido “desde cero”, por lo cual se pudo elegir y controlar exactamente qué sucedía con el sonido en el espacio, y así que este forme parte exclusiva de la obra. Esta obra no podría existir si no existiera el diseño espacial. Las diferentes temáticas tuvieron un lugar específico en el cual desarrollarse, y de hecho la fuerte direccionalidad de los parlantes, que además se ubicaron justo encima de los espectadores, permitió poder escuchar directamente lo que de ellos salía. Sin embargo, también existía una mezcla espacial de los montajes cuando había más de un espectador, y también existía la posibilidad de que la posición espacial de una persona interrumpiera el discurso de la sección de la obra que se escuchaba. Si bien el sistema de reproducción no estaba centrado en permitir el movimiento de las fuentes sonoras, la ubicación fija de las fuentes permitía una apreciación de todo el espacio, reforzada además por la falta de información visual. Con respecto a esto, es interesante analizar la decisión de utilizar solamente información sonora para que los espectadores naveguen el espacio. Como se ha

estudiado en esta tesis (Capítulo 6, el rol de la visión en la percepción auditiva de distancia), la información visual puede afectar sobremanera la percepción auditiva del espacio en el que estamos inmersos, tanto por que nos entrega datos sobre las magnitudes de la sala, como por que afecta la percepción de la localización de fuentes sonoras tanto en acimut como en distancia (Calcagno et al 2012). En el caso de la percepción de distancia, la información visual puede mejorar directamente la performance del oyente, pero en este caso en particular se deseaba justamente lo contrario, que el espectador se vea inmerso en el “éter radiofónico virtual”, que la única información que tenga y las únicas referencias que tenga del espacio sean indicios sonoros. Cuando no tenemos información visual de referencia, las fuentes sonoras parecen más cercanas, y los ambientes parecen más pequeños. En el caso de “Ocupación Invisible” este efecto intentaba simular lo que sucede con las radios clandestinas: muchos lugares, muchas voces, muchos diales en un espacio muy pequeño. A la vez esta percepción del espacio ayudaba a la idea de interferencia, debido a la interacción entre los oyentes que estaban dentro de la instalación. Desde el punto de vista perceptivo, la resolución de no utilizar información visual para la instalación correspondió entonces a dos aspectos específicos: que se trate exclusivamente de una instalación sonora, y que además la falta de esta información potenciara la cercanía de las emisoras (en muchos casos solapadas) y las cercanías de las fuentes sonoras, y por lo tanto la interferencia se vuelva más efectiva desde lo perceptivo.

Desde el punto de vista de la relación entre las Fuentes conceptuales y las Imágenes Fuente, podemos definir tres relaciones diferentes. Anteriormente definimos las Fuentes Conceptuales a partir del contenido de cada monta-

je sonoro, y las Imágenes Fuente a partir de la ubicación espacial de cada parlante y algunos efectos espaciales producidos por el accionar del sistema de sensado. En un principio se pueden especificar 16 Fuentes Conceptuales que se relacionan con las Imágenes Fuentes correspondientes y que siempre ocupan el mismo lugar. En esta instancia se observa la relación una Fuente Conceptual - una Imagen Fuente. Sin embargo cada una de estas Imágenes Fuente se mezcla espacialmente cuando existe más de un espectador en la instalación, y de hecho si los espectadores estaban lo suficientemente cerca uno del otro, se producía una interferencia, que podía tener más de un resultado. Entonces si había dos espectadores, se tenían dos Fuentes Conceptuales (la de los montajes sonoros) con sus correspondientes Imágenes Fuente (del parlante correspondiente). Uno de los espectadores escuchaba con claridad su fuente, pero también escuchará la del otro espectador, y viceversa, y por lo tanto existen dos Fuentes conceptuales y dos Imágenes Fuente, aunque con diferentes grados de importancia para cada espectador, ya que los niveles de atención corresponderán a las señales que tienen el nivel más alto y el cual es mucho más direccional, una especie de efecto “coctel party” (Pollack y Pickett, 1957). Cuando existen más de dos espectadores, el efecto se magnifica, ya tenemos una Fuente Conceptual muy cercana, y otras más lejanas y menos direccionales, ya la distancia juega un papel importante en los niveles de atención. Un paso siguiente es cuando los espectadores están lo suficientemente cerca como para interferir las señales de cada uno. En ese caso las Fuentes Conceptuales tienden a fundirse debido a que la cercanía y la interferencia tienden a hacer que las fuentes tomen una sonoridad común. La fuente conceptual es ahora una sola, pero esta disgregada en varias ubicaciones o Imágenes Fuente,

este es uno de los juegos perceptivos que indica Kendall, una misma Fuente Conceptual partida en diferentes lugares del espacio. Esto se magnifica por la gran direccionalidad de las imágenes fuente. Este efecto se produjo con mucha frecuencia ya que el espacio ya se volvía pequeño cuando había más de 5 espectadores, y se podía escuchar como competían las Fuentes Conceptuales, y como se reconfiguraba la espacialidad de la obra al pasar el tiempo. Otro punto interesante fue que después de un tiempo de navegar la instalación, cada persona podía realizar un recorrido y configurar o reconfigurar el espacio y la forma en que se escuchaba la obra. A pesar de haber controlado la emisión de cada uno de los parlantes, los espectadores eran libres de aprender de la obra y generar su propia obra, mediante la elección de temáticas y espacios preestablecidos. La interacción de varios espectadores en simultaneo, creaban otras nuevas formas de escuchar los audios, generando otra vez, una obra nueva.

Por último, existía una última configuración que se daba cuando no había ningún espectador en la instalación. Cuando esto pasaba, la instalación comenzaba a disparar montajes sonoros seleccionados al azar que contenían los mensajes institucionales de algunas de las radios utilizadas, en general el número del dial, el nombre de la radio y alguna frase a modo de lema. Cada uno de estos montajes ocupaba uno de los parlantes, y por ende una ubicación específica. De esta manera si un espectador escuchaba desde afuera la instalación, según a la distancia en que se encontrara, podía ver un cubo (el entorno de la habitación), pero también escuchar una fuente puntual (la instalación desde afuera) que además iba alejándose y acercándose.

Ocupación Invisible es un claro ejemplo de que se puede configurar la espa-

cialidad de una obra musical, y que además ese espacio planificado puede convertirse directamente en un elemento fundamental de la misma.

## § 9.7 120 MÁQUINA LÍRICA

### 9.7.1. Descripción General

En conmemoración a los 120 años de producción ininterrumpida del Teatro Argentino de la ciudad de La Plata, el grupo Buenos Aires Sonora fue comisionado a realizar una obra. La idea principal se basó en la utilización de todos los espacios que conforman la realización de las óperas, desde los propios escenarios, hasta los talleres de realización, las salas de ensayo individuales y de la orquesta terminando en el espacio para el arte contemporáneo que tiene el teatro el TACEC.

Se propuso entonces investigar, desde lo sonoro, el funcionamiento de la máquina operística, integrada por el edificio y sus ocupantes, ofreciendo al público un recorrido distinto al que en general está acostumbrado. Entonces el recorrido se realizó por las entrañas mismas del teatro, desnudando la maquinaria “humano-tecnológica” que hace funcionar cada una de las óperas presentadas, como indica el siguiente comentario:

“120 Máquina lírica fue un recorrido por múltiples sentidos: por el interior de un teatro productor de arte lírico, por la voz humana, por los 120 años de historia del Teatro Argentino”

<http://martinliut.blogspot.com/2010/08/autobombo-120-maquina-lirica.html>

Así, el trayecto se planteó, para el espectador, como un laberinto de la escu-

cha: los sonidos de la lírica, la música y la producción operística se propagan por espacios anómalos, usualmente sin uso para el tránsito artístico. De esta manera Buenos Aires Sonora aportó la tecnología, que utiliza generalmente, para reunir en el espacio y el tiempo fragmentos de los 120 años de historia lírica del Teatro Argentino de La Plata.

120 Máquina Lírica se dividió en 8 escenas que se desarrollaron en diferentes espacios del Teatro Argentino, el Foyer, la sala Pettorutti, el segundo subsuelo, donde se encuentra la sala de ensayo para orquestas y otras salas menores, las escaleras que dan a la calle 53, los talleres de escenografía, los ascensores y el TACEC (Teatro Argentino Centro de Experimentación y Creación que funciona en el mismo Teatro Argentino). A simple vista se puede observar lo ecléctico de los espacios utilizados. Cada uno de estos espacios contiene información acústica individual, lo que produce que la obra sea modulada sonoramente por estas características. 120 Máquina Lírica, nació pensada de esta manera: espacios sin uso pero que poseen riqueza tímbrica y espacial particular, el espacio forma parte estructural de la obra.

Debido a la gran cantidad de espacios y dispositivos utilizados en esta obra, solo se procederá al análisis y descripción de la última escena de la obra, quizás la más compleja de todas. Este análisis contendrá las características del lugar utilizado, los sistemas de reproducción y espacialización utilizados y el análisis de la estructura sonora espacial.

### **9.7.2. Análisis “120 Máquina Lírica” – Escena Final**

Se procederá al análisis de la última escena de la obra, la cual se desarrolla entre la sala del TACEC (Teatro Argentino Centro de Experimentación

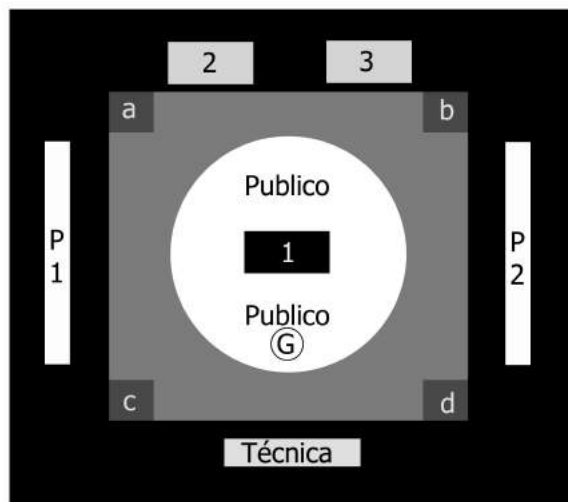
Contemporánea) y la sala Ginastera (Sala principal del Teatro Argentino), ubicada inmediatamente arriba del TACEC.

El análisis de la misma se realizará, como se ha indicado, mediante la metodología desarrollada por Kendall (2010), comentada en el Capítulo 4, enfoques analíticos de la espacialidad en la música electroacústica.

La escena se planteó como un “ensayo de una ópera”, con cantantes (bajo, tenor, mezzo y soprano), un director y un piano en la Sala Ginastera, el cual se transmitió “en vivo y en directo” por video a dos pantallas ubicadas en la sala del TACEC, donde se encontraba el público. En dicha sala se encontraban, además del público, una batería completa, un xilófono, un trío de metales (Trompeta-Trombón-Trompeta), el director de la escena del TACEC, la narradora-guía (que lleva al público por una especie “de visita guiada” durante toda la obra) y la mesa de control de sonido y procesamiento en tiempo real. En la Figura 9.18 y la Figura 9.19 se puede observar un diagrama de la ubicación de las fuentes y ensamble dentro de la sala del TACEC y cuatro fotografías que dan una idea de la puesta en escena:

Esta escena, la última de la obra, se presenta como una ópera intervenida por las nuevas tecnologías, y por la pluralidad de herramientas de procesamiento. Por lo tanto la misma comienza como un ensayo acústico propio de la ópera clásica (cantantes, piano y director) que con el correr del tiempo comienza a ser intervenido por instrumentos amplificados y a ser procesados a la manera de la música electroacústica. Se debe destacar que el desarrollo de la escena traslada al público, y por lo tanto a su perspectiva de una sala de ópera tradicional (la sala principal Alberto Ginastera), hacia la sala del Centro de Experimentación Contemporánea del Teatro Argentino, lo que simbolizará el





**Figura 9.18:** *P1-P2: Pantallas de video; a-b-c-d: Cuadrafonía; 1: Ensamble de metales (Tp, Tb, Tr); 2: Batería; 3: Xilófono; Técnica: Consola, Pc para procesar y espacializar; Público: El círculo blanco marca donde estaba el público. G: ubicación de grabación binaural, mirando hacia 1.*

nuevo lugar de la ópera moderna.

Esta última escena de la obra se extiende por casi 23 minutos conteniendo una introducción, donde se puede ver y observar las características de los momentos previos a un ensayo y el comienzo del mismo (duración aproximada: 8 minutos), el ensayo en sí, a la manera tradicional (duración aproximada: 3 minutos), desarrollo de la obra, con intervención de todos los actores (duración aproximada: 9 minutos) y una coda final (duración aproximada: 3 minutos).

El análisis de esta escena se realizará a partir de una grabación binaural que se realizó de una de las funciones, y que rescata la percepción espacial de la escena. Esta grabación se realizó mediante un grabador portátil Zoom HN4 y unos micrófonos intraurales de Sound Professional ubicados en las orejas de una persona que realizó todo el recorrido. Esta grabación binaural esta disponible en <http://estebanalcagno.com.ar/>.



**Figura 9.19:** *120 Máquina Lírica.* Arriba izquierda: Atrás, batería y xilófono; adelante, ensamble de metales y director de escena. Arriba derecha: Consola y PC para procesamiento y espacialización de las señales fuentes. Abajo izquierda: Pantalla de la derecha; Abajo derecha: pantalla izquierda, ambas muestran lo que sucede en la sala Ginastera.

### 9.7.3. Análisis de los diferentes tipos de FUENTE

#### 9.7.4. Señales Fuente

Se comenzará este análisis indicando cuales son las señales fuente que se utilizaron en esta escena. La producción de cada uno de los actores de la escena fue capturada con micrófonos individuales, por lo que las señales fuentes son las siguientes:

Sala Ginastera: Soprano, Mezzo, Tenor, Bajo, Director, Piano.

Sala TACEC: Batería (estéreo), xilófono (estéreo), Trompeta, Trombón, Trompeta y Voz Relatora.

Se deben agregar a estas, otras señales fuentes provistas por una computadora: textura granulada procesada en tiempo real (con control cambiante de tamaño de grano, altura de cada grano y separación entre granos) y dos mon-

tajes electroacústicos realizados a partir de grabaciones acústicas de operas de Giorgi Ligeti (“Le Grand Macabre”, el pasaje One, two, three, fivez ) y Giussepe Verdi (Coro de “Tosca”).

Por lo descrito arriba, en total se trabajó con 16 señales fuente. Cada una de estas señales llegaba a la consola principal, y eran enviadas a una computadora que realizaba procesamientos en tiempo real, y luego a otra computadora encargada de espacializar cada una de las señales fuente por separado. Una vez realizado el procesamiento y la ubicación espacial, se enviaban 4 canales nuevamente a la consola principal, que se encargaba de enviar estas señales a los amplificadores y luego a la cuadrafonía que rodeaba al público.

Vale agregar que existía otra Señal Fuente que se presenta en varios pasajes de la escena, pero que su origen es acústico: son pasos que se escuchan arriba de los espectadores en el TACEC, pero que son generados en el piso de madera la Sala Ginastera. Estos sonidos de pasos, golpes y corridas son generados adrede y constituyen otra Señal Fuente utilizada en la obra.

Si bien el marco de análisis de Kendall está restringido a la música electroacústica (señales de audio reproducidas en parlantes), tomaremos a las señales acústicas que producen los instrumentos como señales fuentes, debido a que también son señales que se convierten (se toman con micrófonos) para reproducirse en parlantes. Se observa también que la parte de estas señales que se escuchan por afuera del aparato de reproducción también funcionan como señales fuente y son muy útiles para generar juegos perceptuales interesantes.

### 9.7.5. Fuentes conceptuales

En general Kendall considera fuentes conceptuales a aquellos sonidos con los que el oyente tiene un vínculo de significado. Por lo tanto se pueden considerar fuentes conceptuales a aquellas que los espectadores reconocen como tales:

Las voces: Soprano, Mezzo, Tenor, Bajo, Director y Voz Relatora.

Los Instrumentos: Batería (el concepto de batería, aunque se la considera un ensamble), xilofón (un instrumento que no es tan reconocible como una batería), Trompetas, Trombón.

Las secuencias electroacústicas generadas por computadora (granuladas) no pueden considerarse fuentes conceptuales, porque en general no pueden asociarse con una fuente en sí. Aunque las granulaciones son bien reconocidas por los músicos electroacústicos, quizás el público no las reconozca, en todo caso cada uno la relacionará estos sonidos con alguna experiencia previa de escucha, por lo que no puede considerarse una fuente conceptual. Los montajes electroacústicos de las óperas de Ligeti y Verdi podrían considerarse en algún punto fuentes conceptuales. Sin embargo ambas señales tienen un procesamiento sumamente destructivo, donde solo se llega a reconocer una orquesta sinfónica y un coro, los grandes ausentes físicos en “120 Máquina Lírica”. Igualmente, cuando se escuchan estos montajes en la escena se logran reconocer las cuerdas y algunas voces, por lo que es interesante analizar cuanto de fuente conceptual tienen estas grabaciones. Se produce, en este último aspecto, una deliberada ambigüedad en la taxonomía de Kendall que podría enriquecer el marco conceptual.

### 9.7.6. Imágenes Fuente

Kendall propone que la Imagen Fuente es algo así como el producto sonoro final que percibe el oyente. Por supuesto, como vimos al principio de esta tesis, esto depende de las características del estímulo, las características del ambiente y también depende de la relación que puede llegar a tener con una Fuente Conceptual. También depende de la información visual que recibe el oyente pero, en ese caso, se puede hablar de juegos perceptuales. Sin embargo es interesante pensar en todas las cualidades de las que depende la Imagen Fuente, la percepción de los sonidos, y más específicamente en una escena auditiva tan compleja como la aquí presentada.

A lo largo de toda esta escena aparecen muchas Imágenes Fuentes asociadas a Fuentes Conceptuales. Debido a que el sistema de espacialización cuadrafónico basado en Ambisonics que se usó (ver Capítulo 8, Sección 5, GUISPA) en la obra permite la ubicación de las fuentes en casi cualquier localización de un espacio en dos dimensiones, es más constructivo analizar sección por sección de cada escena y observar los juegos perceptuales que se producen entre Fuente Conceptual e Imagen Fuente.

Otra característica a evaluar es que Kendall no trata en su taxonomía a las fuentes móviles, las cuales están muy presentes en esta escena. Sin embargo, la “animación” de las fuentes se puede asimilar a algunas de las técnicas de transformación de la imagen sonora que menciona (chorusing, granulación, disociación de regiones espectrales), en tanto puede producir, en efecto, una imagen-fuente que sea o no congruente con la-s fuente-s conceptual-es. De esta manera serán tratadas las fuentes móviles en este análisis. Se recorrerá la

clasificación de las Imágenes Fuente directamente en su relación con los juegos perceptuales que se analizaran más adelante.

#### 9.7.7. Niveles de organización

Ya planteados los diversos tipos de fuentes con los que cuenta la escena, se pueden observar diversos niveles de organización, que en algunos casos presentan características especiales.

Los niveles de organización planteados por Kendall, aparecen en esta escena, pero parecen mutar en las diferentes secciones de la misma.

Como esquema FUENTE se puede señalar a cada una de las voces (Soprano, Mezzo, Tenor y Bajo), Director, Voz Relatora, Trompetas y Trombón, Batería y Xilofón.

Sin embargo algunas de estas también funcionan como ENSAMBLE en algunas ocasiones: las voces como un coro, la batería como varios cuerpos separados, aunque la noción de batería la hace escuchar como un solo instrumento, los metales como ensamble de metales. Así también en determinados momentos se percibe que las conversaciones que se producen por parte de los cantantes, también funcionarían como esquema de ENSAMBLE, que además se acentúa debido a que los espectadores pueden observar por las pantallas a ese ensamble. Luego se observará que estos esquemas son unificados o rotos gracias a los juegos perceptuales aplicados. Tanto en su formato de FUENTE, como funcionando como ENSAMBLE, se aprecian algunos de los atributos espaciales propuestos por Kendall: Las voces presentan una distancia y una dirección, y se puede observar el ancho variable cuando funcionan como EN-

SAMBLE. La batería tiene un ancho determinado, gracias a que está tomada y reproducida en estéreo. Lo mismo sucede con el xilófono. En ambos casos se pueden establecer dirección y distancia.

En cuanto al esquema de RECINTO, se podría decir que esta parte de la obra se desarrolla en la sala del TACEC, pero sin embargo este no es el único recinto. Es necesario subir un nivel para explicar lo que en realidad sucede. El nivel superior está determinado por el esquema ESCENA, que contiene todos los actores de la misma y se desarrolla en dos RECINTOS separados: la sala del TACEC y la sala Ginastera. Esta organización puede ser percibida por el espectador tanto de forma auditiva como de forma visual. En este caso, el juego perceptual se da con los esquemas auditivos de RECINTOS. El TACEC contiene la escena final, pero también contiene sonora y visualmente a lo que está sucediendo en otra sala. A través de la mediatización de los sucesos que se producen en la sala Ginastera, la sala del TACEC tiene la doble tarea de ser RECINTO y CONTENEDOR, de la forma en que lo define Kendall, el espectador escucha la acústica de la sala Ginastera pero dentro de la sala del TACEC. Se podría pensar que este es un caso muy común en el cine y los audiovisuales, dado que se muestran entornos y recintos siempre dentro de otro recinto, que es el de la proyección. Pero en estos casos, la acústica del “recinto de audición” resulta neutralizada (a través de tratamiento acústico) y se crea la ficción de que el espectador es trasladado al ámbito acústico del relato cinematográfico. En cambio, en el caso antes descrito, el “recinto dentro de otro recinto” está puesto en evidencia y forma parte de un planteo estructural de la obra.

Como se ha descrito arriba, se producen complejos juegos perceptuales en

todos los niveles de organización planteados por Kendall.

#### **9.7.8. Juegos con los agrupamientos perceptivos**

Si bien se toma como marco teórico para el análisis de esta obra la taxonomía desarrollada por Kendall, es necesario realizar algunas aclaraciones que involucren los juegos perceptivos que aparecen entre las Fuentes Conceptuales y su relación con las Imágenes Fuentes que se conforman.

En principio, un juego perceptivo que involucra a toda la ESCENA, deviene en la relación que las FUENTES y ENSAMBLES tienen entre su condición visual y su condición sonora. Al tomar todas las señales fuentes y espacializarlas rompiendo la relación que estas tienen con las fuentes acústicas que las producen, se está generando una ruptura perceptual importante. Esto sucede por momentos con los instrumentos que hay en el TACEC: batería, xilófono y metales son tomados y movidos en el espacio rompiendo la relación con el objeto en sí. También existe otra ruptura entre información visual-sonora: como en general estos instrumentos pueden radiar sonidos a una amplitud muy alta, se produce el efecto siguiente: el espectador ve el instrumento, escucha el instrumento en ese lugar donde lo está viendo, pero también escucha al instrumento espacializado en la cuadrafonía, y esa posición virtual no siempre coincide con la localización visual del mismo. Esta ruptura sucede en casi toda la escena pero se potencia en algunos casos. Por las pantallas de video puede observarse lo que sucede en la sala Ginastera, puede verse a los cantantes cantando, pero los sonidos son ubicados en el espacio del TACEC mediante la cuadrafonía. La ruptura en este caso es total, y así aparece un juego perceptivo extremo. Esta ruptura visual-sonora es el primer juego perceptual que



aparece: las Fuentes Conceptuales sonoras son reforzadas por la información visual de esas Fuentes, pero son quebradas por las Imágenes Fuentes que se perciben en el TACEC mediante la espacialización virtual.

Partiendo de este marco, se analiza cuales juegos sobre agrupamientos perceptivos se realizan en la escena. Para esto se realizará un análisis basado en la escucha de la grabación binaural que se realizó de esta escena. Se dividirá la escena en 6 secciones y una Coda, sobre las cuales se indicaran los juegos sobre las relaciones entre Fuentes Conceptuales e Imágenes Fuente.

#### Sección 1 – Grabación: 0:00m - 5:50m

Esta sección plantea los momentos previos al ensayo operístico. Se escucha a la mezzo y a la soprano “templando” sus voces. Las voces funcionan como FUENTES individuales y sus Fuentes Conceptuales no son ambiguas. El juego perceptivo esta dado por los movimientos que se perciben en las Imágenes Fuente. Ambas voces están cantando pero pasan de estar estáticas a moverse rápidamente de izquierda a derecha, realizar intercambios de posiciones espaciales o envolver al público. En algunos casos se concentran ambas voces en un solo punto del espacio.

En esta primera parte se observan entonces, más allá del movimiento, el juego de única Fuente Conceptual - única Imagen Fuente (una voz por ubicación espacial), varias Fuentes Conceptuales - única Imagen Fuente (dos voces una ubicación espacial) y una Fuente Conceptual - varias Imágenes Fuente (una voz, múltiples localizaciones rodeando al espectador).

A su vez, en esta escena comienzan a aparecer otras FUENTES: una granulación con granos agudos y que van apareciendo aleatoriamente en el espacio,

sonidos de pasos por encima de la cabeza de los espectadores. Si consideramos a la granulación como fuente conceptual (básicamente no lo es, pero después de un lapso es eso, una granulación), aparece también la relación una Fuente Conceptual - varias Imágenes Fuente (un sonido de granulación, varias localizaciones espaciales).

A lo largo de esta sección también escuchamos algunos sonidos de batería. Es interesante remarcar que en la grabación se puede escuchar por momento golpes en los platillos sobre la derecha pero también al frente a la izquierda. La batería funciona como Fuente Conceptual, pero es descompuesta en dos ubicaciones espaciales. Esta escena sigue presentando estas organizaciones hasta que llega el director.

#### Sección 2 – Grabación: 5:50m – 11:18m

Al comienzo de esta sección aparecen el director y el pianista para comenzar el ensayo, los cuales se ubican a la izquierda al igual que cuando aparecen las demás voces, ubicándose de esta manera (muchas fuentes conceptuales una ubicación espacial) casi hasta el final de la sección, donde después de una indicación del director para que se dirijan a un lugar más amplio, todos empiezan a ubicarse en un arco más amplio del espacio. Pasando a una configuración de múltiples Fuentes Conceptuales - múltiples Imágenes Fuente. El efecto pasa por abrir el campo sonoro espacial cuando lo indica el director, pasando de una configuración a otra. Hay que aclarar que esta transformación es acompañada por la batería y el xilofón que también empiezan en la izquierda y luego se acomodan en arco frontal.

Desde el principio se percibe el hecho de que la granulación se hizo más den-

sa y empieza a girar en forma de círculo alrededor de los espectadores, y va desapareciendo prácticamente al final de la sección, preparándose para la siguiente sección.

### Sección 3 - Grabación: 11:18m – 20:11m

Esta sección comienza con el piano en el centro realizando un obstinado, un acorde que se repite siempre igual. Las voces se encuentran ocupando el arco frontal, generando una configuración de múltiples Fuentes Conceptuales - múltiples Imágenes Fuente. La batería se encuentra ubicada hacia la izquierda, mientras que el xilofón abarca el estéreo del frente, generando una configuración de una fuente conceptual - dos Imágenes Fuente. Los metales se ubican puntualmente en el centro de la escena, y se puede hablar de una configuración de varias Fuentes Conceptuales relacionadas con una sola Imagen Fuente.

La electrónica granulada sigue moviéndose alrededor del público, volviendo a ser más errática sobre el minuto 12:42.

Justo en ese momento se empieza a repetir el ciclo descrito anteriormente, pero con mayor amplitud y algunas modulaciones, pero manteniendo sus lugares en el espacio. El único cambio importante es el alejamiento que se produce entre las dos trompetas, ubicadas en el mismo sitio: una parece estar a mayor distancia de la otra. En el minuto 14:06, vuelve a empezar el mismo ciclo, aunque con mayor densidad cronométrica y con mayor amplitud. Vuelve la electrónica granulada que de vez en cuando gira en forma de círculos. Es en este momento donde las FUENTES se transforman en ENSAMBLE y proveen una sensación más envolvente debido a la densidad, aunque las FUENTES siguen manteniendo sus espacios.

En el minuto 15:37, comienza un juego perceptual desde lo sonoro ya que

se comienza a reproducir una grabación que contiene parte de una ópera de Giuseppe Verdi, más específicamente un duetto de voces. Lo interesante es que esta grabación se ubica a la izquierda de los espectadores, pero a la vez se armonizan las voces de la Soprano y Tenor (cantando muy agudo), añadiendo una capa, y ubicándolas en el mismo sitio que la grabación. De esta manera se obtiene una configuración que contiene varias Fuentes Conceptuales relacionadas con una sola Imagen Fuente, pero que además vienen de dos tipos distintos de Señales Fuente: una grabación y tomas de micrófonos.

Entre los minutos 16:52 y los 17:57 se produce el climax de esta escena final. Los metales cambian de configuración espacial y pasan de una ubicación puntual a moverse por el espacio en el arco frontal, de manera que pasan a ser Fuentes Conceptuales relacionadas con Imágenes Fuentes individuales. Es en este momento (16:52) donde aparece en escena una grabación que contiene información de una ópera de Ligeti procesada electrónicamente. La sensación sigue siendo envolvente ya que múltiples fuentes rodean y entregan información sonora desde casi todos los lugares. Las Fuentes Conceptuales con mayor movimiento son los metales, el resto se mantiene estático en su lugar. A lo lejos se escucha la textura producida por el procesamiento de la ópera de Ligeti. En el minuto 18:02, casi la única información sonora que queda proviene de la cinta, que presenta un formato envolvente, acompañado por los metales que vuelven a una configuración de una única Imagen Fuente, y la batería, siempre a la izquierda de la escena, además de la granulación.

Siguiendo en esta sección, ya al minuto 18:47, los metales en el centro de la escena tocan un motivo rítmico armónico. La batería sigue a la izquierda, al xilófono de la derecha, mientras el personaje que oficiaba de guía realiza un

relato en el centro derecho de la escena. El juego perceptivo se da entre el xilófono y la batería, ya que sus sonidos se asemejan mucho. De esta manera parece haber una Fuente Conceptual dividida en dos Imágenes Fuente.

Coda – Grabación: 20:11m – 22:52m

La última escena presenta la batería a la izquierda, el trombón en el centro, el xilofón en el medio, la electrónica granulada a la izquierda y al cantante Bajo realizando canto armónico. Lo más interesante que ocurre en este final es el juego que se realiza entre el xilófono y el cantante armónico. El xilófono se ubica en el centro pero está tocado con arco, que se entremezclan con los armónicos que el Bajo despliega en su canto. Da la sensación de que los sonidos del xilofón son los mismos que los armónicos del Bajo produciendo una ruptura en el espacio. Se podría decir que se tiene una Fuente Conceptual, dividida en dos Imágenes Fuente.

#### 9.7.9. Conclusiones

La música reproducida por altoparlantes no tiene una espacialidad natural en sí misma, como por ejemplo sí la tiene un instrumento acústico. Estas características espaciales deben diseñarse mediante el procesamiento de las señales sonoras, la elección del sistema de reproducción (cada parlante agrega sus características tímbricas a la señal reproducida) y el lugar donde el sonido se desenvuelve. Así, la elección de los espacios acústicos determina también una forma de darle entidad acústica a la obra musical. La producción de obras en sitios específicos (y momentos específicos; Liut, 2009b) genera una ruta para la explotación de las cualidades acústicas y espaciales de sonidos dentro

de esos ámbitos. Diferentes espacios generan diferentes sonoridades, y estas sonoridades proveen de carácter propio a los sonidos que en él se reproducen. En esos ámbitos, la cualidad espacial del sonido aparece en primer plano pudiendo generar marcos estructurales que dan un basamento para la experimentación musical, sonora y acústica. Buenos Aires Sonora basa su trabajo en estas cualidades e interpreta que, justamente, la espacialidad del sonido, el espacio del sonido, es parte fundamental de una composición musical, sea acústica o electroacústica. Las tres obras aquí presentadas demuestran que se puede planear, organizar y trabajar sobre estas cualidades, lo que aporta un punto más de organización musical que puede generar posiciones estéticas interesantes. Los espacios eclécticos en los cuales se desarrollan cada una de estas composiciones musicales, demuestra la búsqueda del grupo sobre una relación aun más profunda, estructuralmente, del sonido musical con los ambientes en donde estos se desarrollan. Esta búsqueda intenta jugar con estos espacios y recorrer los diferentes estadios donde el sonido se transforma, y utilizar estas características en función de la obra, y no solo como “el lugar” de la obra.

## REFERENCIAS BIBLIOGRÁFICAS

- BLESSER, B. y SALTER, L.R.: *Spaces speak, are you listening? Experiencing aural architecture*. The MIT Press, Cambridge, 2007.
- LABELLE, B.: *Background Noise. Perspectives on Sound Art*. London, 2006.
- LIUT, M.: «Arte Sonoro en el Espacio Público: condiciones para su desarrollo actual en Argentinas». *Revista Afuera, Estudios de Crítica Cultural*. <http://www.revistaafuera.com/NumAnteriores/print.php?page=04>. *PraxisUrbana. Liut. htm*, 2009a, 4(7).
- : «Notas al pie de la ciudad. Arte sonoro para sitios y tiempos específicos». *Revista LIS. Letra Imagen Sonido. Ciudad Mediatizada*, 2009b, 1(3), pp. 85–99.
- LÓPEZCANO, R.: «Arte Sonoro: procesos emergentes y construcción de paradigmas». *Sanchez Andrea, Leticia (ed.). Música, Ciencia y Pensamiento en España e Iberoamérica durante el siglo XX. CM-UAM*, 2012.
- MARGOLUS, N.: «Physics-Like models of computation». *Physica D.*, 1984, 10, pp. 81–95.
- POLLACK, I. y PICKETT, J.M.: «Cocktail party effect». *J. Acoust. Soc. Am.*, 1957, 29(11), pp. 1262–1262.
- SUBÍAVALDEZ, J.R.: «Real Time Custom Spatialisation Patch for a large dimension and non conventional area». *Proceeding of PDcon 2011, Berlín*, 2011.

---

TERCERA PARTE:  
CONCLUSIONES





---

---

## Conclusiones

---

---

En esta tesis se han presentado desarrollos y resultados en tres de los campos que estudian la percepción auditiva del espacio: la psicoacústica, el desarrollo de programas informáticos y la música electroacústica. Si bien de cada temática por separado se obtienen conclusiones valiosas, es interesante pensar también como cada uno de estos campos presentan zonas comunes y relaciones que potencian los aportes.

Por esto, las conclusiones de cada uno de los temas tocados en esta tesis serán presentadas de forma individual, pero también serán relacionadas entre sí, con el afán de tender puentes entre los diferentes campos que estudian el espacio auditivo.

### § 9.8 CONCLUSIONES SOBRE EL ESTUDIO DE LA PERCEPCIÓN AUDITIVA DE DISTANCIA

En cuanto al estudio de la percepción espacial del sonido, se ha hecho hincapié en el estudio de la percepción auditiva de distancia, un campo dentro de la psicoacústica que no ha sido explorado en profundidad. Para esto se han presentado los resultados de dos estudios donde se estudió la influencia de la visión y el contenido espectral de un estímulo sonoro en la percepción de

distancia. Con respecto a la influencia que la información visual puede tener sobre la percepción auditiva de distancia es interesante destacar varias conclusiones.

Los resultados obtenidos en el Capítulo 6 muestran que la presencia de información visual afecta a la percepción auditiva de distancia para estímulos con información espectral completa en el rango audible (ruido blanco). Sin pistas visuales la percepción de distancia es idéntica a la reportada en trabajos anteriores (Bekesy, 1949; Cochran et al, 1968; Simpson y Stanton, 1973; Bronkhorst y Houtgast, 1999; Zahorik, 2001, 2002; y Zahorik et al, 2005), o sea, una buena precisión en los primeros metros y una subestimación de las distancias a partir de los 3 metros.

Cuando existen pistas visuales limitadas pero de las cuales se sabe la distancia a la que se encuentran, se observan dos efectos interesantes: la distancia a la fuente sonora se sobrestima en casi todas las distancias testeadas y además la respuesta es más extendida que cuando no hay pistas visuales.

Otro aspecto interesante de los resultados reportados en el Capítulo 6 de esta tesis es que la información visual obtenida de pistas visuales con anterioridad, puede ser guardada en memoria para ser utilizada como una referencia espacial para estimar distancias de una fuente sonora en la oscuridad. Por lo tanto, una respuesta más precisa puede ser obtenida en la oscuridad cuando a los participantes se les permite inspeccionar visualmente el ambiente antes de realizar el experimento.

La visión podría proveer un contexto para los juicios de localización espacial auditiva, por lo que serviría como un organizador espacial auditivo (Warren,

1973). Esto puede ser soportado por la idea de que, debido a que las características espaciales de un ambiente pueden ser percibidas de forma más precisa por la modalidad visual que mediante la modalidad auditiva, se debería esperar que esta información que nuestro cerebro puede obtener mediante la modalidad visual pueda calibrar las respuestas obtenidas mediante el sistema auditivo.

En el trabajo presentado en esta tesis se sugiere un posible rol de la visión en la percepción auditiva de distancia (PAD): obtener información confiable acerca de las características espaciales (dimensiones) del lugar donde se producen los acontecimientos auditivos (en este caso los cambios en la distancia de la fuente), con el fin de calibrar la información que proviene de las señales auditivas relacionadas con PAD (en su mayoría relativas) y asignar un valor de distancia a la fuente de sonido percibido dentro de este espacio.

Esto también puede apoyarse en la idea de que la PAD y la percepción auditiva del tamaño de una sala pueden estar relacionadas mediante el indicio que otorga en ambos casos la reverberación. Un mayor tiempo de reverberación indicaría una sala más amplia. Si una fuente de sonido y un oyente se encuentran en la misma sala, la percepción del sonido de una fuente alejada más allá de las dimensiones de dicha sala solo sería posible dentro de una sala de mayores dimensiones (Cabrera et al 2005, 2006).

Una cuestión importante que se quiere tratar en este trabajo es la posibilidad de que la mejora en los juicios de PAD observadas en ambientes reverberantes podría ser explicada, además de por la razón entre señal reverberada y directa, por la asociación entre el tiempo de reverberación de una sala y su tamaño.

Existe abundante evidencia que apoya esta hipótesis. Mientras la presentación de un sonido sin reverberación produce la impresión de una fuente muy cercana (Mershon y King, 1975) el agregado de reflexiones lleva a la percepción de una fuente más distante (Mershon et al, 1989). Bronkhorst y Houtgast (1999) ha demostrado que un incremento en la cantidad de las reflexiones en la simulación de la reverberación de una fuente de sonido virtual resulta en un incremento en la distancia aparente. Nielsen (1993), a su vez reporta un incremento en los juicios de distancias aparentes en ambientes reverberantes y semireverberantes cuando se los compara con las respuestas en una cámara anecoica. Otros resultados similares fueron obtenidos en muchos estudios previos (Békésy, 1938; Butler et al, 1980; Wagenaars, 1990; Begault, 1992; Bronkhorst y Houtgast, 1999).

Muchos experimento de PAD observan una función compresiva para la distancia percibida VS la distancia real de la fuente sonora (Zahorik et al, 2005). La parte superior del rango restrictivo de las distancias percibidas es comúnmente conocido como “horizonte auditivo” (Bronkhorst y Houtgast, 1999; Zahorik, 2002). Este límite ha sido asociado con la igualdad que produce entre la energía del sonido directo y el sonido reflejado que provee una fuente sonora a determinada distancia (Sheeline, 1984).

Así, ha sido sugerido que la distancia del “horizonte auditivo” depende de las características acústicas del ambiente. Sin embargo, los resultados obtenidos en el presente trabajo (experimento 2 y 3) sugieren fuertemente que el “horizonte auditivo” puede también depender de factores no acústicos, como la presencia de múltiples pistas visuales antes y durante el experimento.

Si la hipótesis de que la PAD y la percepción del tamaño de la sala están relacionadas es verdadera, el sistema visual tendría una ventaja obvia sobre el sistema auditivo al proveer esta información a causa de que la percepción auditiva del tamaño de una sala depende de las características acústicas del ambiente, mientras que la visión es más efectiva al obtener información sobre el tamaño, forma, materiales, etc.

Si bien estas conclusiones se basan en datos duros obtenidos mediante experimentos en entornos sumamente controlados, en el ámbito de la música electroacústica se podrían utilizar los resultados obtenidos en este trabajo debido a que, en general, la audiencia está expuesta a mucha información visual que puede afectar a la percepción espacial de la escucha de una obra.

Imaginemos que un compositor piensa en una determinada disposición espacial de los sonidos que utiliza en su obra. Su idea es generar una determinada sensación espacial, de distancia o profundidad sonora. Si estos sonidos se reproducen en parlantes a la vista de los oyentes, podría comprometerse de forma importante el éxito de las sensaciones que el compositor pretende transmitir, debido a que una información visual que no es coherente con lo que se está escuchando podría cambiar la percepción de ese espacio.

Tener en cuenta como la visión puede establecer fuertes efectos perceptivos, también puede dar herramientas al compositor para que su trabajo sea más inmersivo (si este así lo requiriese) o para que los juegos perceptivos que intenta establecer sean más eficientes. Quizás un simple diseño de iluminación, o una distribución velada de las fuentes sonoras, generaría un ambiente más apto para la correcta percepción en profundidad de las fuentes. Un ejemplo de esto

puede apreciarse en el Capítulo 9, Sección 6, en donde se describe “Ocupación Invisible”, del grupo Buenos Aires Sonora. Allí se decidió de antemano una sala oscura (desprovista de información visual) como ambiente de la obra, lo que permitió asegurar que la percepción espacial del ambiente estaría basada en gran medida por la información sonora que se presentaba.

Otro ejemplo, contrario al anterior, sería el que se describe en el Capítulo 9, Sección 7, en el análisis de la obra “120 Máquina Lírica”, también de Buenos Aires Sonora, en donde se eligió para la última escena de la obra un entorno totalmente cargado de información visual y auditiva en muchos casos contradictoria, lo que permitió que los juegos perceptivos, muchas veces buscados en la música electroacústica, sean sumamente efectivos mediante el planeamiento a priori del tipo de información sonora-visual que los espectadores tienen disponibles.

Así, el estudio y diseño a priori de las interacciones entre información espacial auditiva y visual (o, aunque más no sea la conciencia de la importancia y el rol de estas) en el contexto de una obra musical o multimedia, puede llegar a ser crucial para la percepción de determinadas características de una obra.

Por otro lado, los resultados mostrados en el Capítulo 7 (el rol del contenido espectral en la percepción auditiva de distancia) permiten observar claramente que el contenido espectral de una fuente sonora puede afectar marcadamente la percepción de distancia auditiva de un oyente, en condiciones donde, si bien este conoce el ambiente en el que está inmerso, no recibe información visual de la fuente sonora. De esta manera se ha observado que estímulos que tienen un gran ancho de banda y que tienen información en todo el espectro auditivo,

tal como es el caso del ruido blanco y el ruido rosa, pueden ser localizados en distancia con una precisión muy alta. Sin embargo, en las mismas condiciones, estímulos con un contenido espectral reducido, como es el caso de los tonos puros, son más difíciles de ubicar, y dependen mucho de la zona del espectro donde se encuentren. Mediante los experimentos realizados en esta tesis se puede observar que estímulos de tonos puros ubicados en la zona baja-media del espectro (en este caso 500Hz y 1KHz) no pueden ser localizados correctamente en distancia por los oyentes. Sin embargo cuando los tonos puros se ubican en la zona alta del espectro (en este caso 2KHz y 4KHz) las respuestas se vuelven un poco más precisas pero sin dejar de subestimar distancias lejanas. En estos experimentos también se observó que el ancho de banda mínimo que se necesita para una buena localización en distancia de una fuente sonora depende de la zona del espectro que se enfatice. De esta manera para las frecuencias bajas (ruido rosa filtrado con un filtro centrado en 500Hz) se necesita un ancho de banda de más de una octava para lograr una respuesta de mediana precisión, sin lograr los mismos resultados que con los estímulos completos y complejos. Cuando se aumenta la frecuencia central del estímulo, se necesita un ancho de banda menor para lograr mejores resultados, por ejemplo para ruido rosa filtrado con un filtro centrado en 1KHz se logran buenos resultados con un ancho de banda de  $1/6$  de octava, logrando performances similares a las obtenidas con ruido blanco y rosa al llegar a un ancho de banda de una octava y media. Por último, con ruido filtrado con un filtro centrado en 2 y 4 KHz, con un ancho de banda de entre  $1/12$  y  $1/6$  de octava ya se logran respuestas lineales y muy similares a las obtenidas con los ruidos blanco y



rosa.

A pesar de que estos resultados parecen ser robustos, es válido aclarar que estos pueden ser muy dependientes de la forma y composición acústica de la sala en la que se reproducen los estímulos. Por lo tanto es imprescindible saber cuál es la respuesta en frecuencia de los entornos en los que se realizan los experimentos. En este caso, la sala donde se realizó el experimento es casi anecoica para frecuencias por encima de 4KHz, y muy reverberante para frecuencia bajas, por lo que también las diferentes características reverberantes del entorno podrían estar afectando a la percepción de distancia. Sin embargo, y a pesar de esta salvedad, puede verse claramente que, en todo caso, el centroide de la frecuencia de un estímulo puede afectar la percepción auditiva de distancia.

Los resultados obtenidos en este estudio son útiles para los compositores de música electroacústica y también para los diseñadores de programas de espacialización de audio, los cuales en general no tienen un control tan evidente de la distancia de una fuente sonora.

Para un compositor, el dato de que el contenido espectral de una fuente sonora puede presentar cambios importantes en la percepción de distancia puede ser muy útil a la hora de elegir determinadas instrumentaciones para determinadas obras o secciones de una obra. Haciendo un paralelismo con los resultados mostrados aquí se podría ejemplificar la diferencia en la localización de distancia (vale la pena aclarar que esto también sucede en acimut y elevación, como se ha indicado antes) de una flauta, que tiene un contenido espectral semejante a un tono puro con un ancho de banda no muy amplio, y un trombón, que

tiene un contenido espectral más rico. Esta diferencia podría ser utilizada sin duda para escoger una orquestación que remarque espacialmente en distancia, texturas compuestas por estos instrumentos. Así también, el hecho de que la percepción en distancia cambia con la ubicación de la energía del espectro se acerca a la idea de poder ubicar los componentes de una melodía a diferentes distancias, o lograr diversos planos con diversos instrumentos de una misma familia. Esto sería también posible no solo en la música instrumental sino también en la música electroacústica, en la que el ancho de banda y la ubicación del contenido espectral es aun más controlable.

Por supuesto, estos efectos deberían estar acompañados de los juegos perceptivos indicados con anterioridad, donde la información visual podría ejercer un efecto diferente sobre cada uno de los elementos que componen la escena. De esta manera, por ejemplo, dos instrumentistas que toquen el mismo instrumento, podrían generar un juego perceptivo entre la información visual y la auditiva si uno de ellos está a la vista de los oyentes y el otro no. Dependiendo la zona del espectro en el que estén tocando y de si los instrumentistas son vistos o no, se pueden producir ciertos “engaños” que nutrirían las posibilidades espaciales que siempre se buscan en una obra de música electroacústica o mixta.

Por otro lado, la importancia de las características acústicas de un entorno, o el conocimiento detallado de las mismas, permite un control más profundo sobre estos efectos. Conocer cómo afecta un ambiente determinado a los sonidos que estamos reproduciendo o ejecutando es sumamente útil tanto para el diseño de una obra como para su correcta reproducción o ejecución. Dos

casos interesantes sobre el efecto de las cualidades de un ambiente sobre una obra pueden encontrarse en esta tesis. En el caso de la obra “Oí(r) el ruido” del grupo Buenos Aires Sonora, al plantear la obra en un espacio público y abierto se estudió (de forma empírica) una reflexión del sonido que emanaba de los edificios que estaban en frente de la audiencia. Este rebote presentaba una réplica del sonido reproducido, filtrado y con una amplitud importante, que se utilizó en la misma obra para generar un juego perceptivo entre lo que sucedía por delante y detrás de los espectadores. De la misma manera, en la última escena de “120 Máquina Lírica” se realizó un juego perceptivo utilizando las características reverberantes de dos entornos diferentes, por lo cual los sonidos y su percepción, tanto en distancia como en los otros planos del espacio, variaba con cada elección.

## § 9.9 CONCLUSIONES SOBRE EL DESARROLLO DE APLICACIONES INFORMÁTICAS DEDICADAS A LA ESPACIALIZACIÓN DE FUENTES SONORAS

En el Capítulo 8 de esta tesis se ha comentado el desarrollo y aplicación de tres programas informáticos específicos que se enfocan en la posibilidad de crear, diseñar y desarrollar distintos tipos de cualidades espaciales a partir de diferentes técnicas y enfoques. Las aplicaciones informáticas diseñadas pretenden ser útiles para los compositores de música electroacústica que estén interesados en un control más exhaustivo de la espacialidad dentro de sus obras. Para tal fin se enfocó el desarrollo de estos programas sobre tres problemáticas diferentes:

- 1) La generación de reverberación mediante el uso de la técnica de convolución de archivos de audio pre grabados con respuestas a impulso de entornos reales o virtuales.
- 2) El diseño de trayectorias espaciales en tres dimensiones, con el fin de poder generar una “partitura de trayectorias” de fuentes sonoras en una obra, de tal manera de poder controlar el movimiento de las fuentes sonoras y su relación espacio-temporal.
- 3) Un espacializador de fuentes sonoras en tiempo real basado en la técnica Ambisonics (Véase Malham, 2009).

Para cada una de estas problemáticas se llegó a soluciones específicas que no pretenden ser únicas ni finales, pero que se encuentran disponibles para su utilización y constituyen un buen punto de partida para futuros desarrollos.

En el primer caso, se desarrolló Gtkonvolve, programa que permite la carga de archivos de audio y archivos de respuestas a impulso multicanal, y combinarlos de tal manera que permitan la creación de juegos perceptivos determinados. Este programa utiliza la técnica de convolución (mediante el método overlap/add, vease Moore, 1990) en tiempo diferido, lo que permite tiempos de procesamiento relativamente bajos y resultados finales muy interesantes.

Para el segundo punto se desarrolló 3DEV, una aplicación basada en el trabajo con matrices que permite el procesamiento de la información de las trayectorias en 3 dimensiones dentro de un espacio determinado. De esta manera, el usuario del programa puede diseñar una gran cantidad de trayectorias de fuentes sonoras, visualizarlas y editarlas, para luego poder observar su inter-

relación en el tiempo y en el espacio. Además se puede asociar un archivo de audio determinado, y así realizar una simulación gracias a la conexión de 3DEV a otros programas como, por ejemplo, Csound (Barry Vercoe et al, MIT).

Para el último punto, se trabajó con la plataforma de desarrollo de aplicaciones que procesan audio en tiempo real, Pure Data (Miller Puckette, UCSD), para la generación del programa GUISPA, el cual permite ubicar hasta 16 fuentes sonoras en un espacio determinado, y moverlas de forma individual o agrupadas, con trayectorias espaciales creadas en tiempo real mediante el ratón de la PC o una interfaz física MIDI. Una de las características interesantes que presenta GUISPA es el desarrollo de una interfaz de usuario similar a la de una consola de audio, lo que permite la rápida adecuación al sistema. Al tratar cada fuente sonora como si fuera un canal de audio, se permite al usuario individualizar las características de cada fuente y poder trabajarlas por separado, aunque después esta fuente sonora pueda formar parte, por ejemplo, de un ensamble (Véase Kendall, 2009). De esta manera podemos tener diferentes señales de audio provenientes de varias fuentes como un micrófono, un archivo de audio que se reproduce u otro medio, asignar una técnica de espacialización diferente para cada fuente (en este caso se prioriza el uso de Ambisonics), controlar su amplitud y su localización en el espacio elegido con una utilización óptima de los recursos del sistema.

Gracias a la aparición de tecnologías y lenguajes de programación cada vez más simples de utilizar y más sencillos de aprender, se puede prever que el desarrollo de programas ad hoc para espacializar fuentes sonoras en un en-

torno o sito específico se convierta en una práctica muy común por parte de los compositores de música electroacústica en los próximos años. El desarrollo de estas herramientas, cuya finalidad remite a la solución de problemáticas específicas a los que se puede enfrentar un compositor de música electroacústica, es una muestra de que es posible realimentar una obra musical desde la utilización y diseño mismo de las herramientas informáticas que se utilizan en ella. De esta manera, entender lo que se puede hacer mediante herramientas informáticas (en muchos casos de software libre) es una buena práctica por parte del compositor para resolver nuevas problemáticas a la hora de componer un espacio determinado.

Si bien los tres programas desarrollados están disponibles para su uso, solo el último mencionado, GUISPA, fue desarrollado y utilizado especialmente para una obra de música electroacústica. En el Capítulo 9, Sección 7 se comenta como se utilizó este programa en la obra “120 Máquina Lírica”, del grupo Buenos Aires Sonora, de manera tal que cumplió con las necesidades planteadas, al lograr generar un control preciso de las trayectorias de fuentes sonoras de diferentes orígenes. De esta manera GUISPA se convirtió en parte fundamental de la obra y se sirvió de la realimentación que otorgaron los ensayos de la misma a la hora de calibrar el sistema y mejorar su performance.

Otro vínculo interesante que se puede encontrar entre uno de estos programas y otras temáticas de la tesis aparece en la relación directa entre la percepción de distancia y el programa Gtikonvolve. Es interesante que mediante Gtikonvolve se pueda manipular la relación entre el sonido directo y el sonido reverberante (relación D/R) de una fuente sonora específica, ya que la relación

D/R es una de las pistas cruciales para la percepción auditiva de distancia. Esta característica propia de Gtikonvolve, muestra la forma en que se puede pensar el desarrollo de un programa informático para que cumpla las tareas que necesitamos. Mediante este artilugio, se puede simular la distancia de una fuente sonora dentro de un ambiente determinado no solamente generando variaciones de intensidad, sino también generando diversas relaciones entre el sonido directo y reverberante logrando una percepción de profundidad más robusta e indicando puntos en común con el resto de los temas que se tratan en esta tesis.

## § 9.10 CONCLUSIONES SOBRE EL ANÁLISIS DEL ESPACIO EN OBRAS ELECTROACÚSTICAS Y MIXTAS

En el Capítulo 9 de la presente tesis se realizó una extensa descripción de tres obras electroacústicas llevadas a cabo por el grupo Buenos Aires Sonora. En este capítulo se describieron y analizaron estas obras desde tres puntos de vistas específicos, pero que sin embargo poseen aspectos en común: las características históricas y arquitectónicas de los lugares donde se desarrolló la obra, lo que define los campos simbólicos y acústicos de la misma; el soporte tecnológico que se utilizó para llevar a cabo la reproducción de señales de audio, que en todos los casos lleva implícita las técnicas utilizadas para lograr la espacialidad de fuentes sonoras y, por último, el análisis específico de la relación entre la espacialidad del sonido y la estructura musical de cada obra, basado en el enfoque de análisis desarrollado por Kendall (2010).

Con respecto a esto último, vale aclarar que el enfoque de análisis de Kendall

(2010) fue útil y efectivo para describir de manera clara determinados aspectos espaciales de una composición o montaje sonoro, aun en situaciones de extrema complejidad como el que se trató en las obras analizadas de esta tesis. Es también interesante observar como mediante esta metodología se pudo encarar el trabajo sobre tres tipos de obras muy diferentes: una instalación sonora, una obra fuertemente audiovisual y una obra de grandes magnitudes al aire libre con la participación de bailarines y músicos.

Durante su utilización para el análisis de estas obras, se ha demostrado que, si bien la taxonomía original no está destinada al análisis de música y montajes sonoros para medios mixtos, esta puede ampliarse y proyectarse de manera afectiva para tales fines. Uno de los aspectos más importantes que surgen de esta ampliación es la posibilidad de que los oyentes estén expuestos a más de una imagen-fuente cuando escuchan una fuente acústica amplificadas y/o procesadas. Así por ejemplo, debido a los cambios que puede sufrir una fuente sonora por los diversos procesamientos que se pueden aplicar en ella, se pueden establecer diversos niveles de relación simbólica con esas fuentes sonoras. Los efectos de interrupción o incongruencia entre imagen fuente e imagen conceptual se pueden potenciar y aun producir mediante el agregado de información visual a la información sonora. La información visual agregada puede estar constituida por intérpretes, imágenes o animaciones, actuaciones, escenografía, etc. Por otro lado, se mencionó en el punto anterior que la presencia y actuación de fuentes acústicas aporta también información visual que participa en el juego perceptivo de los esquemas espaciales.

Las características de los juegos perceptivos planteados por Kendall y am-



pliados en esta tesis muestran la posibilidad de analizar obras de música electroacústica no solo mediante las características musicales tradicionales sino también mediante el aporte de la espacialidad que se presenta en la misma y el aporte que diversos tipos de información no sonora puede aportar a la percepción holística de la obra.

Este último capítulo de la esta tesis sirve como corolario de los temas desarrollados poniendo a disposición cada uno de los conocimientos revisados y adquiridos en su desarrollo. Se contribuye así a la explicación de los modos y circunstancias en los que la espacialidad del sonido puede adquirir una gran importancia en una obra musical electroacústica o mixta, hasta lograr ser parte estructural de la misma y darle un significado determinado.

El análisis profundo de las obras ha demostrado que la información de los tres campos estudiados en esta tesis (percepción sonora, técnicas y tecnologías, análisis musical) puede interactuar con las ideas compositivas de un músico electroacústico, tanto para cargar de significado a sus creaciones como para proporcionarle un sostén en el entendimiento acerca de la manera en que los oyentes perciben el espacio auditivo.

## REFERENCIAS BIBLIOGRÁFICAS

- BEGAULT, D.R.: «Perceptual effects of synthetic reverberation on three-dimensional audio systems». *Journal of the Audio Engineering Society*, 1992, **40**, pp. 895–904.
- BÉKÉSY, G.V.: «Über die Entstehung der Entfernungsempfindung beim Hören (On the origin of the sensation of distance in hearing)». *Akustische Zeitschrift*, 1938, **3**, pp. 21–31.
- : «The moon illusion and similar auditory phenomena». *American Journal of Psychology*, 1949, **62**, pp. 540–552.
- BRONKHORST, A.W. y HOUTGAST, T.: «Auditory distance perception in rooms». *Nature*, 1999, **397**, pp. 517–520.
- BUTLER R.A., LEVY E.T. y W.D., NEFF: «Apparent distance of sounds recorded in echoic and anechoic chambers». *The Journal of Experimental Psychology*, 1980, **6**, pp. 745–750.
- CABRERA D., CAPRA A. FARINA A., AZZALI A. y P., MARTIGNON: «Perceived room size and source distance in five simulated concert auditoria». *12th International Congress on Sound and Vibration Lisbon*, 2006.
- CABRERA D., KWAKE H.J., JEONGB D. y J.Y., KIMD: «Auditory room size perception for modeled and measured rooms». *Internoise Congress and Exposition on Noise Control Engineering*, 2005.
- COCHRAN P., THROOP J. y W.E., SIMPSON: «Estimation of distance of a source of sound». *American Journal of Psychology*, 1968, **81**, pp. 198–206.
- MERSHON, D.H. y KING, L.: «Intensity and reverberation as factors in the auditory perception of egocentric distance». *Perception and Psychophysics*, 1975, **18**, pp. 409–415.
- MERSHON D.H., LITTLE A.D. MCMURTRY P.L., BALLENGER W.L. y J.L., BUCHANAN: «Effects of room reflectance and background noise on perceived auditory distance». *Perception*, 1989, **18**, pp. 403–416.
- NIELSEN, S.H.: «Auditory distance perception in different rooms». *Journal of the Audio Engineering Society*, 1993, **41**, pp. 755–770.
- SHEELINE, C.W.: «An investigation of the effects of direct and reverberant signal interaction on auditory distance perception». *Tesis Doctoral Universidad de Stanford*, 1984.
- SIMPSON, W.E. y STANTON, L.D.: «Head movement does not facilitate perception of the distance of a source of sound». *American Journal of Psychology*, 1973, **86**, pp. 151–159.
- WAGENAARS, W.M.: «Localization of sound in a room with reflecting walls». *Journal of the Audio Engineering Society*, 1990, **38**, pp. 99–110.
- WARREN, D.H.: «The role of visual-auditory ‘compellingness’ in the ventriloquism effect: implications for transitivity among the spatial senses». *Cognitive Psychology*, 1973, **1**, pp. 114–133.
- ZAHORIK, P.: «Estimating sound source distance with and without vision». *Optometry and Vision Science*, 2001, **78**, pp. 270–275.
- : «Assessing auditory distance perception using virtual acoustics». *J. Acoustic. Soc. Am.*, 2002, **111**, pp. 1832–1846.
- ZAHORIK P., BRUNGART D.S. y A., BRONKHORST: «Auditory Distance Perception in Humans: A Summary of Past and Present Research». *Acta Acustica united with Acustica*, 2005, **91**, pp. 409–420.

---

## ANEXO



## - Apéndice A -

---

---

### Artículos publicados, presentaciones realizadas y financiamiento obtenido en el marco esta la tesis

---

#### § A.1 ARTÍCULOS PUBLICADOS

**Abregú E.L., Calcagno E.R. y Vergara R.O.** "*La distancia como factor estructural de la música*". Revista Argentina de Musicología, 2012. En Prensa.

**Calcagno E.R.** "*Diseño y construcción del espacio sonoro para sitios específicos*". Revista Sul Ponticello, 2012, 2(34).

**Calcagno E.R., Abregú E.L., Eguía M.C. y Vergara R.O.** "*The Role of Vision in Auditory Distance Perception*". Perception, 2012, 41(2), pp. 175-192.

**DiLiscia O.P. y Calcagno E.R.** "*Sonido Multicanal en composiciones acústicas espacialmente ricas*". Traducción en base al artículo homónimo de Robert Dow. En *Música y espacio: ciencia, tecnología y estética* (compiladores: Oscar Pablo Di Liscia, Juan Pampín, Gustavo Basso), 2010. Editorial Universidad Nacional de Quilmes. Argentina. pp. 161-165

## § A.2 PRESENTACIONES A CONGRESOS Y CONFERENCIAS

**Abregú E.L., Eguía M.C., Calcagno E.R., y Vergara R.O.** *El rol de la visión en la Percepción auditiva de distancia.* Congreso de la Sociedad Argentina de Neurociencias 2012, Huerta Grande, Córdoba.

**Spiousas I., Etchemendy E., Calcagno E.R. y Eguía M.C.** *Influence of two-dimensional sonic crystal on auditory depth perception for pink noise filtered bands.* Congreso de la Sociedad Argentina de Neurociencias 2012, Huerta Grande, Córdoba.

**DiLiscia O.P. y Calcagno E.R.** *3DEV: A tool for the control of multiple directional sound source trajectories in a 3D space.* En Proceedings de la Audio Linux Conference 2010. Utrecht, Holanda. <http://lac.linuxaudio.org/2010/download/lac2010proceedings.pdf>.

**Vergara O.R., Calcagno E.R. y Eguía M.C.** *The role of spectral cues and minimum bandwidth in the auditory perception of distance.* First Pan-American/Iberian Meeting on Acoustics 2010, Cancún, México. Publicado en J. Acoust. Soc. Am. 128(4), 2455-2455.

**Vergara R.O., Calcagno E.R. y Eguía M.C.** *Las características espectrales del sonido afectan la percepción de distancia de una fuente acústica.* I Reunión Conjunta de Neurociencias (IRCN), 2009. Huerta Grande, Córdoba.

**Calcagno E.R.** *Software Libre, Investigación Psicoacústica y música: Arduino + Pure Data bajo Linux*. Conurbania 2009, 1er. Ciclo de Conferencias sobre Software Libre en el Conurbano Bonaerense. UNQ.

### § A.3 FINANCIAMIENTO OBTENIDO

**Subsidio de apoyo a la investigación para estudiantes de grado e investigadores en formación (SAI2011)**. UNQ. “*Implementación de un dispositivo experimental multicanal con automatización remota*”. Director: Dr. Pablo Di Liscia. 2012 - 2013.

**Beca Conicet Tipo II**. LAPSO, UNQ. “*Desarrollo de un Entorno Acústico Virtual para el estudio de la percepción de movimiento y cambios de velocidad, sin desplazamiento angular, de una fuente sonora*”. Director: Dr. Manuel Eguía. Año 2011 - 2013.

**Subsidio de apoyo a la investigación para estudiantes de grado e investigadores en formación (SAI2009)**. UNQ. Proyecto: “*Desarrollo de Head Dummy para grabación binaural de sonido*”. Director: Dr. Pablo Di Liscia. Año 2010-2011.

**Beca ANPCYT (Agencia Nacional de Promoción Científica y Tecnológica)**. UNQ. “*El Espacio en la Composición Musical del Teatro Acústico*”. Director: Dr. Oscar Pablo Di Liscia. Año 2007-2011.

