

UNIVERSIDAD POLITECNICA DE VALENCIA

ESCUELA POLITECNICA SUPERIOR DE GANDIA

Grado en Ing. Sist. de Telecom., Sonido e Imagen

---



UNIVERSIDAD  
POLITECNICA  
DE VALENCIA



ESCUELA POLITECNICA  
SUPERIOR DE GANDIA

**“Espacialización sonora con *Wavefield Synthesis* y *Vector Base Amplitude Panning*. Estudio comparativo.”**

**TRABAJO FINAL DE GRADO**

Autor/a:  
**Pablo Gutiérrez Parera**

Tutor/a:  
**José Javier López Monfort**

**Valencia, Febrero 2013**



## Resumen

Los sistemas de sonido espacial *Wave Field Synthesis* (WFS) y *Vector Base Amplitude Panning* (VBAP) son capaces de reproducir sonidos localizados y ambientes sonoros. VBAP se basa en panoramizaciones por amplitud, mientras que WFS emplea *arrays* de altavoces para reproducir los frentes de onda de las fuentes sonoras.

La reproducción correcta de la distancia sonora, para percibir la distancia concreta a la que se sitúa la fuente sonora, es una característica no tan simple de generar por los sistemas de sonido espacial.

Para este trabajo se ha diseñado de un test de percepción subjetiva en el que se compara la capacidad de reproducción de la distancia sonora de los sistemas WFS y VBAP. En esta prueba se evalúan ciertos parámetros (distintas distancias, reverberación, tipo de sonido, ángulo de escucha) para estudiar algunas diferencias posibles en la reproducción de distancia de los dos sistemas.

Los datos obtenidos indican que aunque ambos sistemas son capaces de reproducir distancia sonora, WFS obtiene mejores resultados que VBAP.

Palabras clave: *Wave Field Synthesis*, *Vector Base Amplitude Panning*, distancia sonora, test percepción subjetiva

## **Abstract**

Spatial sound systems Wave Field Synthesis (WFS) and Vector Base Amplitude Panning (VBAP) are able to reproduce located sounds and sound environments. VBAP is based on traditional (amplitude) panning, but WFS uses speaker arrays to reproduce the wave fronts of the sound sources.

The right reproduction of the sound distance, to sense the right distance where the sound source is located, is not an easy characteristic to generate by spatial sound systems.

In this work a subjective perceptual test had been done to compare the ability to reproduce accurate sound distance of WFS and VBAP systems. Some factors (different distances, reverberation, sound type, hearing angle) have been studied with this experiment, to evaluate possible differences between the sound distance reproduction ability of both systems.

The data collected point out that both systems are capable to reproduce sound distance, but WFS has better results than VBAP.

Keywords: Wave Field Synthesis, Vector Base Amplitude Panning, sound distance, subjective perceptual test

## **Agradecimientos**

Este trabajo no hubiera sido posible sin la colaboración y apoyo de muchas personas y desde aquí quiero agradecer a todos su ayuda y asistencia.

En primer lugar quiero agradecer el apoyo de mi familia. Sin su confianza y colaboración este trabajo no existiría. Gracias especiales a Andrés Pérez por todo su trabajo programando en el entorno SuperCollider, su aporte de ideas, la colaboración en los montajes, pruebas y consejos. Muchas gracias también a la gente del Gtac por acompañarme a diario durante la realización de este trabajo. Es un placer trabajar con gente como ellos. No puedo dejar de agradecer a Luis Maciá sus colaboraciones todoterreno y a Fernando Domene sus clases particulares de Matlab. Gracias a Gabi, Luis, Fernando, Marian, Stefan, Jose Antonio, Jorge, Laura, Csaba, Berta, Montxo, Amparo, Daniela, Eleonora, Alfonso, Alex, Isa, Stefano, Emanuel, Sara, Héctor, Gemma, Gonzalo, Javier y Sandra por prestarse amablemente a hacer el test de percepción.

Muchas gracias especiales a Blas Payri por ayudarme con el análisis estadístico, a mi tutor José Javier López por su paciencia y apoyo, y a Berta Durán por continuar haciendo de musa y compartir este trabajo conmigo.

A todos ellos muchas gracias.



## Índice

Resumen - Abstract

Agradecimientos

1 – Introducción	1
1.1 – Interés y motivación	1
1.2 – Objetivo	2
1.3 – Estructura de la memoria	3
2 – Sistemas de reproducción de sonido espacial	6
2.1 – Estereofonía y sonido multicanal	6
2.2 – VBAP Principio de funcionamiento	8
2.3 – WFS Principio de funcionamiento	8
3 – Percepción de la distancia por medio del sonido	10
4 - Test de evaluación subjetiva: Comparación de la percepción de distancia con los sistemas WFS y VBAP	13
4.1 – Clasificación de los test psicoacústicos	13
4.2 – Objetivos del test	14
4.3 – Medios (equipo, sala y <i>software</i> )	15
4.4 – Diseño del test	18
4.5 – Implementación	25
4.6 – Procedimiento	37
4.7 – Resultados y análisis	41
4.8 – Posibles investigaciones futuras	50
5 – Conclusiones	51
6 – Bibliografía	53
Anexos	
I – Figuras adicionales (medios empleados)	I
II – Hojas de información para los sujetos	IV
III – Tablas de análisis estadísticos	VII





## 1 - Introducción

### 1.1 - Interés y motivación

Los sistemas de sonido espacial han desarrollado un gran auge en los últimos años. Con la popularización de los sistemas de sonido multicanal, gracias entre otras cosas a la aparición de formatos audiovisuales domésticos como el DVD o el Blu-ray, la sociedad en general ha adquirido cierto conocimiento sobre la escucha de sonido espacial. Ha sido la experiencia directa de estos sistemas de sonido multicanal (5.1, 7.1 ya sea en Dolby Digital o DTS) la que ha potenciado el interés de estos sistemas por el gran público. Este proceso se inició en realidad antes, cuando el público de los cines comenzó a experimentar el sonido multicanal.

El sonido multicanal no es algo nuevo, su aparición está ligada a la historia de la estereofonía y a la del desarrollo del sonido para cines (Holbrook, 1988). Hay que tener en cuenta que uno de los principales objetivos del sonido multicanal, y por ende del sonido espacial, es obtener una correcta sensación sonora espacial (una buena estereofonía) para una audiencia numerosa, es decir, buena estereofonía para muchos oyentes simultáneos. Un detalle interesante a tener en cuenta en el sonido espacial es la posición de escucha correcta o *sweet spot*. La estereofonía convencional de dos canales es capaz de reproducir espacios sonoros con bastante realismo siempre y cuando el oyente se sitúe en el punto exacto de escucha con respecto a los altavoces y éstos estén bien colocados formando un ángulo de 60° (Rec UIT-R BS.775-2, 2006). En cuanto el oyente se desvía de esta posición la escucha estereofónica se estropea, anulando el volumen espacial o creando efectos espaciales distorsionados. Así, para un grupo de varios oyentes sólo uno de ellos tendría una escucha correcta. Por este motivo surgen los sistemas de sonido multicanal, para expandir la posición de escucha correcta (*sweet spot*) a una zona más amplia donde muchos oyentes a la vez puedan experimentar correctamente las sensaciones del sonido espacial (estereofonía). En general, a mayor número de canales de sonido independientes la zona de escucha correcta o *sweet spot* se hace más amplia y más robusta (Holman, 1996) (Jackson, 2010). Además de esto, surgen una serie de ventajas adicionales que mejoran sensiblemente la capacidad de reproducción del sonido, que hacen de la escucha una experiencia mucho más rica y completa (Gutiérrez, 2008). En comparación con el estéreo de dos canales un sonido multicanal permite:

Una externalización de los objetos sonoros más fuerte y sencilla.

Una mejor percepción de los instrumentos.

Un mejor equilibrio de la sonoridad debido a un procesado binaural mejor.

Una mejor percepción del timbre debido a las correcciones de posición que permiten más frentes de onda.

Una percepción de distintos ambientes mucho más acentuada.

Una mejor directividad de las fuentes sonoras.

Mayor sensibilidad a los errores del canal o de procesado.

Los sistemas de sonido multicanal han sido responsables del auge del sonido espacial, pero no son los únicos sistemas capaces de generar sonido espacial. Existen otras aproximaciones, como son los sistemas de sonido estudiados en este trabajo, el *Wave Field Synthesis* (WFS) y el *Vector Base Amplitude Panning* (VBAP). En cierta medida se pueden considerar sistemas multicanales, dado que efectivamente emplean muchos canales, aunque por su naturaleza son considerados como una nueva generación de sistemas de sonido espacial. Un aspecto destacable es que estos sistemas de sonido son herederos de las intenciones de los sistemas de sonido multicanal previos, buscando igualmente una mejora en la reproducción del sonido espacial. Es interesante señalar que estos dos sistemas pueden proporcionar un *sweet spot* más amplio, liberando así a la audiencia de una posición de escucha rígida para percibir la estereofonía. De esta forma el sonido espacial va mejorando en uno de sus principales puntos fuertes. Además, las ventajas adicionales que se obtienen con los sistemas de sonido espacial indicadas antes, hacen que estos sistemas se conviertan en una de las vías principales para desarrollar la calidad y realismo de los sistemas de audio en general.

Así, dado el interés que ha adquirido la sociedad en general a nivel de usuario sobre los sistemas de sonido espacial, y dado también el interés de la industria en mejorar y ofrecer nuevos sistemas de sonido espacial (Slater, 2008) (Nagel, 2009), el estudio e investigación sobre los sistemas *Wave Field Synthesis* y *Vector Base Amplitude Panning* puede aportar información muy valiosa para el sector de la ingeniería de audio.

## 1.2 – Objetivo

Los sistemas objeto de estudio en este trabajo (WFS y VBAP) son una importante vía de desarrollo del sonido espacial. En base a esto en este trabajo se quiere establecer una comparación entre ellos que lleve al mejor entendimiento de sus capacidades y limitaciones. Comparar a fondo ambos sistemas es una tarea muy amplia, por eso se ha buscado un parámetro concreto que evaluar en ambos sistemas y que aporte información relevante. Para llevar a cabo esta comparación se ha elegido como herramienta de trabajo un test de percepción subjetiva con el que poder poner a prueba los sistemas sobre sujetos reales.

Así, el **objetivo** de este trabajo es diseñar y llevar a cabo un test de percepción subjetiva que permita establecer una comparación fundamentada entre los sistemas de sonido *Wave Field Synthesis* y *Vector Base Amplitude Panning*, preferiblemente sobre algún parámetro concreto que tenga interés particular para el sonido espacial.

Después de un tiempo de documentación se llegó a la conclusión de que un parámetro especialmente interesante para la reproducción de sonido espacial es la capacidad de reproducción de **distancia sonora**, es decir, la percepción de la distancia a la que se sitúa la fuente del sonido o la distancia del oyente al evento sonoro.

En cualquier campo sonoro existe una sensación de profundidad espacial. Esta sensación es la responsable de la percepción de la perspectiva en una escena acústica (Rumsey, 2002). El mismo patrón se cumple para una escena acústica reproducida por medios artificiales. La profundidad se considera un atributo esencial para la percepción sonora espacial de la escena acústica, y en esta profundidad cumple un papel primordial la distancia sonora de la fuente.

Según (Wittek, 2004), la recreación fiel de la profundidad de una escena sonora es el objetivo de referencia de los sistemas de sonido espacial. Para conseguir este objetivo es crucial la reproducción correcta de la distancia sonora.

Sistemas como WFS y VBAP, reproducidos sobre arrays de altavoces dispuestos envolviendo a los oyentes, tienen una gran capacidad para posicionar sonidos en diferentes ángulos del plano horizontal alrededor del oyente (Berkhout, 1992) (Pulkki, 2001). Sin embargo su capacidad para situar sonidos en diferentes distancias no ha sido comparada, pudiendo resultar un criterio importante a la hora de conseguir un sonido espacial más completo.

### 1.3 – Estructura de la memoria

La presente memoria, más allá de esta introducción, se encuentra organizada en capítulos de la siguiente manera:

Capítulo 2 – Sistemas de reproducción de sonido espacial:

2.1 – Esterofonía y sonido multicanal

En este apartado se repasan brevemente algunos de los sistemas de sonido más importantes, teniendo en cuenta su capacidad para generar sonido espacial. Se repasa el concepto de esterofonía, en qué consisten los sistemas de sonido multicanal y las distintas aproximaciones al sonido espacial.

## 2.2 – VBAP Principio de funcionamiento

*Vector Base Amplitude Panning* es el otro sistema de sonido espacial a comparar en este trabajo. Brevemente se recoge cómo funciona y sus posibilidades.

## 2.3 – WFS Principio de funcionamiento

El sistema Wave Field Synthesis es uno de los sistemas a comparar en este trabajo. Aquí se resume en qué consiste su aproximación al sonido espacial y en esencia cómo funciona esta tecnología.

## Capítulo 3 – Percepción de la distancia por medio del sonido

La distancia es la característica espacial clave que se utiliza como elemento de comparación entre los sistemas de sonido espacial en este trabajo. En este apartado se enumerarán los distintos factores responsables de la sensación de cercanía o lejanía de un sonido.

Capítulo 4 – Test de evaluación subjetiva: Comparación de la percepción de distancia con los sistemas WFS y VBAP

Este capítulo constituye el grueso de este trabajo y es la aplicación práctica que se ha llevado a cabo para responder al objetivo inicial.

### 4.1 – Clasificación de los tests psicoacústicos

Se hace un resumen breve de los distintos tipos de tests psicoacústicos.

### 4.2 – Objetivos del test

Partiendo del objetivo inicial del trabajo se enumeran una serie de objetivos concretos que se pretenden dilucidar con el test de percepción subjetiva.

### 4.3 – Medios (equipo, sala y *software*)

Para llevar a cabo el test ha sido necesario usar diferentes materiales y recursos, ya fuesen para reproducir el sonido, para realizar medidas y ajustes, o para analizar los resultados obtenidos. En este apartado se detallan todos los materiales empleados para hacer funcionar el test.

### 4.4 – Diseño del test

Se describe cómo se plantea el test: qué preguntar y por qué, qué factores y cuántos sonidos aplicar, mecánica y principios de funcionamiento.

### 4.5 – Implementación

Aquí se describe cómo se ha llevado a la realidad el diseño planteado, incluyendo el montaje final, medidas y ajustes.

#### 4.6 – Procedimiento

Una vez listo el sistema se puso en práctica con diferentes sujetos que respondieron a las preguntas planteadas.

#### 4.7 – Resultados y análisis

Con todas las respuestas recopiladas se procedió a estudiar estadísticamente los resultados, obteniendo diferentes gráficos y parámetros estadísticos. Estos datos son la base del análisis posterior.

#### 4.8 – Posibles investigaciones futuras

Se plantean diferentes líneas de investigación que, tras la experiencia del test planteado en este trabajo, pueden continuar aportando datos relevantes sobre el mismo tema.

### Capítulo 5 – Conclusiones

Es en este capítulo donde se recogen todas las conclusiones derivadas de la realización del test y del análisis de los datos obtenidos.

### Capítulo 6 – Bibliografía

Enumera todas las referencias citadas y consultadas para la realización de este trabajo.

### Anexos

Se incluyen una serie de documentos adicionales que pueden ser interesantes para la mejor comprensión del trabajo realizado. (I-Figuras adicionales de los medios empleados, II-Hojas de información para los sujetos que realizaron el test, III-Tablas de análisis estadísticos)

## 2 – Sistemas de reproducción de sonido espacial

### 2.1 – Estereofonía y sonido multicanal

El objetivo de todo sistema 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 de uno real.

A lo largo de los años han ido surgiendo multitud de sistemas de reproducción de sonido, los cuales aseguraban conseguir una inmersión del oyente en el ambiente sonoro cada vez más real. Sin embargo, aun existiendo multitud de sistemas, todos ellos se basan únicamente en tres técnicas diferentes o, a lo sumo, en una combinación de estas (Theile & Wittek, 2004). Las técnicas en cuestión son:

- Sistemas basados en el *efecto phantom*
- Reconstrucción binaural de las señales en los oídos
- Síntesis del campo sonoro alrededor del oyente

Cuando dos o más fuentes sonoras son coherentes el cerebro interpreta que el sonido proviene de una única fuente por lo que únicamente detecta una posición que viene a ser una media de las posiciones de cada fuente ponderada en función del nivel sonoro de cada una de ellas. A esta suma de localizaciones es a lo que se conoce como el **efecto phantom** ya que proporciona una posición ficticia desde la que aparentemente proviene el sonido. (Ver figura 2.1-1)

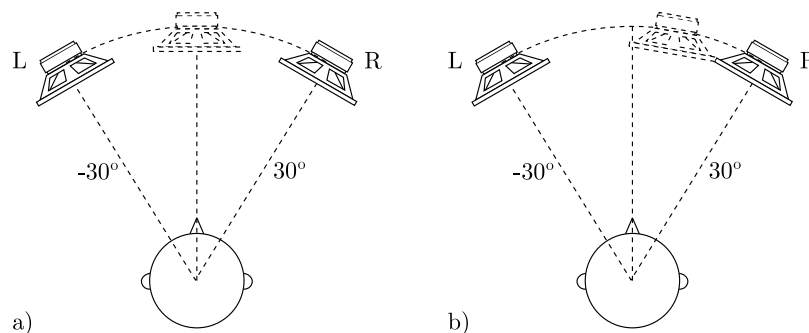
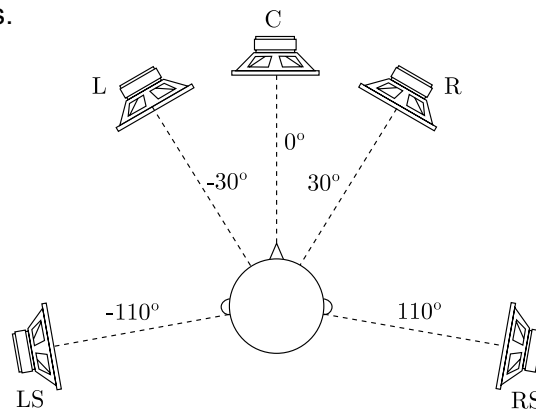


Figura 2.1-1 Efecto *phantom*. a) Imagen *phantom* centrada. b) Imagen *phantom* escorada a la derecha

La palabra *estéreo* en castellano tiene su origen etimológico en el griego *stereós* que significa sólido. Aplicada al sonido, es decir, **sonido estereofónico**, lo que se pretende es añadir al sonido la característica de lo sólido, esto es, con volumen y por tanto espaciosidad. Así al decir sonido estereofónico lo que se está diciendo es “sonido sólido”, que tiene características de profundidad.

Un **sistema estéreo convencional de dos canales**, izquierdo y derecho (L y R) permite crear imágenes *phantom* entre las posiciones de los dos altavoces. La separación entre los altavoces no debe ser superior a  $60^\circ$  para poder mantener una imagen estable (Theile G. , 1990). La forma de situar la imagen sonora es mediante una panoramización de amplitud, esto es, una distribución del nivel de la señal de forma ponderada entre los dos canales.

A lo largo de los años se han ido añadiendo progresivamente más canales al par estéreo con el fin de mejorar la sensación espacial de las escenas sonoras. Cuando los sistemas estereofónicos tienen más de dos canales, se les conoce con el nombre de **sistemas multicanal**. Existen múltiples formatos multicanal (Nakahara, 2005) cada uno de ellos con ciertas características distintas. La configuración 3/2 (3 altavoces delanteros y 2 traseros), conocida también como 5.1 (cinco canales completos más uno opcional con ancho de banda limitado a las frecuencias graves LFE) se convirtió en un estándar que representaba un compromiso entre las necesidades de obtención de una sensación espacial óptima y una implementación práctica que fuera compatible con el sistema estéreo (Rumsey, et al., 2001). Idealmente todos los altavoces deben situarse en un círculo centrado en la posición del oyente (ITU-R. BS 775-2, 2006) (ver figura 2.1-2) Este tipo de sistemas de sonido multicanal sigue haciendo uso del efecto *phantom* (panoramización por nivel entre altavoces) para situar las fuentes sonoras.



**Figura 2.1-2 Configuración de altavoces del sistema multicanal 5.1 según ITU-R BS.775-2**

Los sistemas de **reconstrucción binaural** tratan el problema de la reproducción de sonido desde un punto de vista distinto. Aunque la escena sonora sea muy compleja el oyente únicamente va a captar de ella aquello que llega a sus oídos. Por tanto, parece lógico diseñar un sistema que reconstruya exacta y únicamente la señal que llega a los oídos del oyente. Para esto se emplea la “Función de Transferencia Relativa a la Cabeza” (Head Related Transfer Function, HRTF). La HRTF comprende todos los aspectos físicos de la localización del sonido para una posición concreta de la cabeza (diferencia de tiempos interaural, sombra de la cabeza, difracciones en los pliegues de la oreja, canal auditivo, ...).

Una vez se conocen las HRTF para ambos oídos y la posición de una fuente virtual en el espacio, se pueden sintetizar las señales binaurales correspondientes a los dos oídos para un sonido emitido por dicha fuente.

## 2.2 – VBAP Principio de funcionamiento

El método *Vector Base Amplitude Panning* (VBAP) fue desarrollado por V. Pulkki en (Pulkki V. , 2001a) y es un método de reconstrucción de fuentes virtuales que emplea únicamente la panoramización en amplitud, igual que el sistema estéreo, de hecho, en el caso de escenas en dos dimensiones el VBAP no es más que una reformulación del sistema estéreo (Pulkki & Karjalainen, 2001). La principal novedad que aporta el VBAP con respecto al estéreo es que extiende la panoramización en amplitud a escenas en tres dimensiones (Pulkki V. , 2001b).

El funcionamiento del VBAP es bastante sencillo y efectivo, en el caso de tres dimensiones debemos rodear el escenario con una superficie ficticia que deberemos triangular y colocar un altavoz en cada vértice. La síntesis de una imagen *phantom* se consigue en dos fases, en primer lugar es necesario determinar en qué triángulo se encuentra situada la fuente para posteriormente calcular la ganancia que debería aplicarse a cada uno de los tres altavoces del triángulo. La forma de determinar la ganancia de cada altavoz se consigue expresando la posición de la imagen *phantom* como una combinación lineal de las posiciones de los tres altavoces, esto es, las posiciones de los altavoces forman una base. (Ver figura 2.2-1).

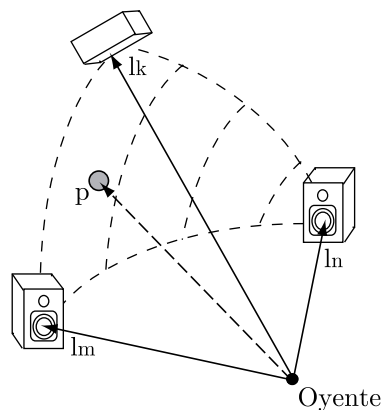


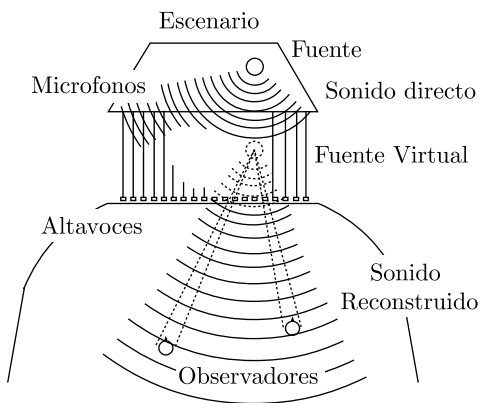
Figura 2.2-1 Descripción geométrica del sistema VBAP

## 2.3 – WFS Principio de funcionamiento

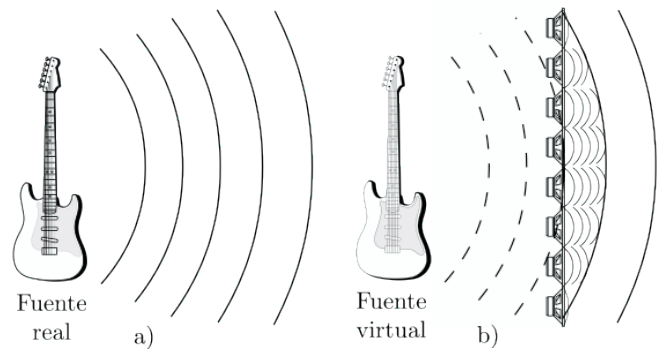
El concepto que más intuitivamente describe el funcionamiento de *Wave Field Synthesis* (WFS (Síntesis del Campo de Ondas) es el de la “cortina acústica” (Snow, 1953), (ver figura 2.3-1). El funcionamiento de la cortina acústica es el siguiente: si delante de un escenario



colocamos una cortina a modo de telón sobre la cual hubiera infinidad de micrófonos, seríamos capaces de registrar el sonido que viaja desde el escenario hasta el patio de butacas. Si ahora colocamos otra cortina sobre la cual hay dispuestos tantos altavoces como micrófonos teníamos en la cortina anterior, podríamos enviar el sonido registrado por cada micrófono a su altavoz correspondiente, con lo cual al otro lado de la segunda cortina estaríamos reconstruyendo la escena sonora original. Este concepto intuitivo fue descrito formalmente por Berkhout mediante la teoría de ondas a finales de la década de los 80 (Berkhout, 1987; 1988), comenzando a emplearse el término *Wave Field Synthesis* a principios de la década de los 90 (Berkhout, Vogel, & de Vries, 1992) (Berkhout & de Vries, 1993).



**Figura 2.3-1 Cortina acústica**



**Figura 2.3-2 Aplicación del principio de Huygens**  
**a) campo sonoro emitido por una fuente real**  
**b) campo sonoro sintetizado mediante WFS**

WFS se fundamenta en el principio de Huygens (Huygens, 1678). Este principio establece que el frente de ondas radiado por una fuente sonora se comporta a su vez como un conjunto infinito de fuentes puntuales distribuidas a lo largo del frente de ondas, a las cuales se las denomina fuentes secundarias. Por tanto, es posible describir la radiación del campo sonoro mediante este conjunto de fuentes secundarias. La forma directa de recrearlas es mediante altavoces, que se deben situar en la misma posición de las fuentes. Ya que la colocación de los altavoces formando la curva del frente de ondas, que además varía al moverse la fuente, no es práctica, se los coloca formando una línea recta y se compensa su desplazamiento mediante la aplicación de una ganancia y retardo (ver figura 2.3-2).

Recientemente, en (Spors, Rabenstein, & Ahrens, 2008) se ha revisado todo el desarrollo teórico extendiéndolo para el caso tridimensional. También se pueden encontrar muy buenas descripciones en (Verheijen, 1997) (Wittek, 2007).

### 3 – Percepción de la distancia por medio del sonido

El sistema auditivo humano es muy sofisticado y tiene la capacidad de analizar y extraer la mayor parte de la información espacial referente a un sonido con tan solo sus dos oídos. Sin embargo el proceso de localizar una fuente sonora es dinámico y a menudo se complementa por medio de otros estímulos sensoriales. Es sobre todo la visión la que también aporta información espacial, y de la combinación de ambos sentidos surge la percepción espacial completa de nuestro entorno. Aún así, el oído por sí solo tiene una percepción espacial muy completa y la capacidad de identificar la dirección de una fuente sonora proveniente de cualquier punto del espacio tridimensional es sorprendentemente precisa (Blauert, 2005). Pero la localización espacial sonora viene determinada, además de por la dirección de la fuente, también por la distancia a la misma.

Los principales factores responsables de la percepción de la distancia por medio del sonido (también llamados *indicios* acústicos) son:

- Nivel de presión sonora
- Relación sonido directo-sonido reverberante
- Espectro
- Diferencias interaurales

**Nivel de presión sonora:** En general, la presión sonora para una posición concreta decrece cuando aumenta la distancia entre el oyente y una fuente sonora de nivel constante. Sin embargo, esta relación depende de características del entorno o también de la fuente sonora, como por ejemplo la directividad de la fuente. En condiciones ideales (una fuente puntual en campo libre) la disminución del nivel de presión sonora en función de la distancia sigue la ley del cuadrado inverso, por la que se produce una reducción de 6dB del nivel de presión cada vez que se dobla la distancia a la fuente. En un entorno reverberante esta reducción será menor, dependiendo de la difusión del campo acústico. El nivel de presión sonora se considera el factor más claro y evidente para determinar la distancia mediante la escucha (Mershon & King, 1975).

**Relación sonido directo - sonido reverberante:** En entornos con superficies reflectantes, la proporción de energía que llega al oyente directamente (sin tocar ninguna de las superficies reflectantes) con respecto a la energía acústica que le llega tras contactar con las superficies reflectantes (sonido reverberante) decrece sistemáticamente al aumentar la distancia a la fuente. Es decir, cuando la fuente sonora está cerca al oyente éste percibe

más cantidad de sonido directo que reverberante, y a medida que la fuente se aleja la cantidad de sonido directo que percibe se reduce, aumentando el sonido reverberante. Este factor es por tanto determinante en espacios cerrados, donde siempre hay un cierto grado de reverberación. Para una habitación dada, la energía reverberante en función del tiempo viene determinada por el tamaño de la sala y por las superficies reflectantes de la misma. En muchos espacios abiertos también se producen reverberaciones, y por tanto la relación sonido directo-sonido reverberante también varía con la distancia.

**Espectro:** El contenido espectral de un sonido también da información sobre la distancia de la fuente. Para distancias mayores de unos quince metros las propiedades absorbentes del aire atenúan las altas frecuencias. Aunque el efecto puede considerarse pequeño, del orden de 3 ó 4dB de pérdida cada 100 metros a 4kHz (Ingard, 1953), el resultado es perfectamente percibido por el oído. Por otro lado, cuando la fuente sonora está muy cerca de la cabeza del oyente se puede producir un aumento de las bajas frecuencias, debido a la curvatura del campo sonoro (von Békésy, 1960).

**Diferencias interaurales:** Las diferencias relativas en las características de la onda acústica al llegar a los oídos aportan mucha información para la localización espacial de fuentes. Estas diferencias se pueden clasificar en diferencias interaurales de tiempo (ITD) y diferencias interaurales de nivel (ILD). Cualquier sonido alcanza cada oído con una diferencia temporal, debido al distinto camino que toma para llegar a un oído u otro, esto produce las diferencias interaurales de tiempo. Este ligero desfase es aprovechado por el sistema auditivo para calcular la posición en la que se sitúa la fuente. Las diferencias de tiempo se hacen mayores cuando la fuente sonora se acerca mucho a la cabeza, debido al tamaño y forma de la misma. Por otro lado, las diferencias interaurales de nivel son debidas a variaciones en intensidad de la señal acústica que llega a cada uno de los oídos. El sonido sonará más fuerte en el oído más próximo a la fuente, no tanto por la distancia entre oídos sino por el denominado *efecto sombra* que produce la cabeza. Este efecto se incrementa cuanto más alta es la frecuencia. Por debajo de 1-1.5kHz aproximadamente, las ILD dejan de ser relevantes como indicio de localización espacial porque las ondas comienzan a difractarse alrededor de la cabeza en lugar de reflejarse, evitando el efecto sombra.

Las ITD e ILD se consideran indicios muy importantes para la localización espacial sonora, y cada uno de ellos tiene su rango específico de operación en el mecanismo global de localización. Las ITD prevalecen a frecuencias bajas, hasta la mencionada frecuencia de corte de 1-1.5kHz. Además las ITD son capaces de dominar a las ILD siempre que el contenido en bajas frecuencias sea suficiente, como por ejemplo en señales de banda

ancha (Wightman & Kistler, 1992). Las ILD son significativas en ausencia de bajas frecuencias, especialmente para señales de banda estrecha (Middlebrooks & Green, 1989). Esta complementariedad funcional se suele denominar *time-intensity trading* (negociación tiempo-intensidad).

Las diferencias interaurales tienen una relación directa con la forma del frente de onda. Así, si la fuente está cerca del oyente el frente de onda aún conservará su curvatura, produciendo diferencias interaurales concretas y distintas de las que se producirían si el frente de ondas fuese plano debido a una mayor distancia de la fuente sonora (Zahorik, 2002).

Cuando se mueve la cabeza para buscar mejor la localización de un sonido se está haciendo uso de las diferencias interaurales, forzando sus máximos y mínimos e integrando esta información en el tiempo. Lo mismo sucede cuando la fuente sonora está en movimiento con respecto al oyente. Este ejercicio activo también se conoce como paralaje acústico.

Existe otro factor adicional de suma importancia para la localización espacial sonora y en concreto para la determinación de la distancia, que es el **conocimiento de la fuente sonora y del entorno acústico**. Este factor se puede considerar como complementario a todos los anteriores y en muchos casos es determinante para la correcta localización. La familiaridad con el tipo de fuente sonora en cualquier caso siempre ayuda a calcular las distancias, dado que se puede hacer una estimación de la potencia acústica de la fuente, así como del contenido espectral y temporal de la misma.

Los factores antes descritos se combinan y procesan de manera conjunta en un proceso psicológico cognitivo, dando finalmente como resultado la sensación de distancia a la fuente.

## 4 – Test de percepción subjetiva: Comparación de distancia con los sistemas de sonido WFS y VBAP

### 4.1 – Clasificación de los tests psicoacústicos

La psicofísica es una rama de la psicología que estudia la relación entre los estímulos físicos y las sensaciones o percepciones que generan (Ehrenstein & Ehrenstein, 1999). Dentro de la psicofísica se engloba la psicoacústica, que concentra su estudio en la percepción del sonido. Una de las principales herramientas de trabajo de estas disciplinas son los tests de percepción subjetiva, con los que se recopilan datos de muchos sujetos y se obtienen resultados concretos. Existen diferentes clasificaciones de métodos y tests psicofísicos (Gescheider, 1997) (Nunnally & Bernstein, 1994), pero en psicoacústica está muy aceptada la siguiente clasificación de métodos para cuantificar la impresión sonora (Bech & Zacharov, 2006):

**Escala directa:** Se basan en pedir al sujeto que realiza la prueba una conversión directa de la sensación sonora en una magnitud sensorial preestablecida en una escala

- Nominal: El sujeto debe decidir si dos estímulos son equivalentes o no.
- Ordinal: El sujeto debe decidir si la magnitud de la sensación es mayor o menor que la de otro estímulo, para un atributo concreto.
- Intervalos: El sujeto debe especificar sobre cada estímulo cuánto le parece mayor o menor con respecto a otro estímulo. Los datos obtenidos darán una lista de ítems ordenados, donde las distancias entre los ítems están especificadas. Los valores absolutos de la escala no son conocidos por los sujetos.
- Ratio: Es similar al método de escala por intervalos con la diferencia de que el valor absoluto de la unidad de medida de la escala sí es conocido.

**Escala indirecta:** Se mide la sensación mediante la capacidad de los sujetos de discriminar entre distintos estímulos

- Estímulo constante: El sujeto debe indicar cuál de entre dos estímulos da la mayor sensación, incluyendo un estímulo de referencia. Con las respuestas se calcula la probabilidad de que un estímulo sea mayor que otro y a partir de aquí se obtienen umbrales de percepción mínimos.
- Método de los límites: Se buscan los umbrales superior e inferior de percepción a partir de comparaciones entre una referencia y un estímulo modificado por pasos fijos.
- Método del ajuste: El sujeto debe ajustar una variable del estímulo hasta hacer coincidir su percepción con una referencia.

## 4.2 - Objetivos del test

En el test llevado a cabo en este TFG se ha comparado la capacidad para reproducir distancia sonora de los sistemas *Wave Field Synthesis* y *Vector Base Amplitude Panning*.

Ambos sistemas están especialmente diseñados para la reproducción espacial sonora, por esto resulta de interés comprobar cómo se comportan a la hora de reproducir distancias. Aunque ya han pasado años desde la concepción de ambos sistemas de sonido, siguen teniendo en gran medida un uso restringido a la investigación, siendo cualquiera de ellos una posible vía de futuro en la mejora y difusión de los sistemas de sonido espacial. De aquí también surge el interés de comparar en una misma prueba ambos sistemas.

Además, debido al número de altavoces y al procesado que cada uno de los sistemas necesita para su funcionamiento, sería útil conocer si existe una diferencia importante o no en su capacidad de reproducir distancia sonora. Es decir, dado que WFS requiere de más medios que VBAP, puede ser útil saber si para la reproducción de la distancia sonora merece la pena utilizar un sistema que como WFS requiere de más medios, o si con uno más sencillo como VBAP se obtienen los mismos resultados.

El objetivo inicial de este trabajo es diseñar y llevar a cabo un test de percepción subjetiva que permita establecer una comparación fundamentada entre los sistemas de sonido *Wave Field Synthesis* y *Vector Base Amplitude Panning*, concretada en una comparación sobre su capacidad para reproducir distancia sonora.

Partiendo de esta premisa y teniendo en cuenta el test de percepción subjetiva diseñado, se pueden establecer una serie de objetivos concretos a los que buscar respuesta. Así, con este test de percepción subjetiva se pretende dilucidar:

- si alguno de los dos sistemas es capaz de reproducir sensaciones sonoras de distancia con la suficiente fidelidad y claridad
- si para estos sistemas es determinante o no en la percepción de la distancia el tipo de sonido reproducido
- si para estos sistemas influye o no en la percepción de la distancia la localización angular
- si para alguno de los dos sistemas ayuda o no en la percepción de la distancia una reverberación adicional

- en definitiva, a la vista del comportamiento frente a todas estas características, evaluar si alguno de los dos sistemas tiene una capacidad global mejor para reproducir distancia sonora.

### 4.3 - Medios (equipo, sala y software)

Para llevar a cabo el test de percepción subjetiva se han empleado material e instalaciones del iTeam (Instituto de Telecomunicaciones y Aplicaciones Multimedia), situado en la Ciudad Politécnica de la Innovación de la UPV en Valencia.

Los medios empleados para llevar a cabo el test de percepción subjetiva comprenden el equipo para la reproducción y funcionamiento del test, la sala donde se realizó el test, los equipos de medida para comprobaciones y ajustes, y el software para la síntesis de sonido, procesado y análisis de los resultados.

**El equipo para la reproducción y funcionamiento del test** constaba de:

- 14 amplificadores de sonido multicanal (7.1 canales) YAMAHA RX-V650 (ver anexo I figura A-1)
- 4 tarjetas de sonido de 24 entradas y 24 salidas MOTU 24IO (ver anexo I figura A-2)
- 8 arrays de altavoces con 8 *tweeters* Beyma T2010 y 8 *woofers* Beyma 5MP60/N (ver anexo I figura A-3)
- 1 altavoz adicional con 1 *tweeter* Beyma T2010 y 1 *woofer* Beyma 5MP60/N (ver anexo I figura A-4)
- Ordenador para sintetizar los sonidos en WFS y VBAP
- Ordenador portátil con mando de videojuegos para control del sujeto que realiza el test
- Tela acústica (10 x 2,80m) para ocultar los altavoces a la vista

Los arrays encargados de reproducir los sonidos sintetizados por los sistemas WFS y VBAP constan de 8 altavoces (con *tweeters* de 1 pulgada y *woofers* de 5,5 pulgadas) separados entre sus centros acústicos 18 cm. (Ver anexo I figura A-3)

Todos los altavoces se ajustan a la recomendación (Rec UIT-R BS.1116-1, 1997) (*Métodos para la evaluación subjetiva de pequeñas degradaciones en los sistemas de audio incluyendo los sistemas de sonido multicanal*) según la cual se especifican unas

consideraciones sobre los dispositivos de reproducción que van a intervenir en una prueba de evaluación subjetiva. Entre estas condiciones se incluyen características electroacústicas como la respuesta en frecuencia, directividad, distorsión, retardo o dinámica. Una de las consideraciones generales básica que afectaban al diseño del test de percepción subjetiva era la necesidad de que todos los altavoces empleados en el experimento fuesen iguales. Dado que era necesario utilizar un altavoz adicional (que se usaría como referencia), externo a los altavoces integrados en los arrays que reproducirían WFS y VBAP, hubo que construirlo expresamente. Para cumplir con los requisitos de igualdad se utilizaron el mismo tipo de *tweeter* de 1 pulgada, *woofer* de 5,5 pulgadas y filtro pasivo que empleaban los arrays ya montados del laboratorio. Se construyó la caja acústica a la que se le añadió material absorbente en el interior y sobre la que se montaron el *tweeter* y el *woofer* ya comentados, además de su correspondiente filtro pasivo (ver anexo I figura A-4).

La **sala** donde se realizó el test fue la “sala de ensayos de audio y video” del laboratorio del Gtac (Grupo de Tratamiento de Audio y Comunicaciones) integrado en el iTeam, instituto de investigación perteneciente a la UPV. Dicha sala está acondicionada para la escucha y dimensionada para pruebas y ensayos audiovisuales (ver anexo I figura A-5). Gracias a este tratamiento la sala cumple con los parámetros especificados en la norma (EBU Tech 3276-2nd Edition, 1998) (*Listening conditions for the assessment of sound programme material: monophonic and two-channel stereophonic*) así como su suplemento 1 para sonido multicanal (EBU Tech 3276-E, 2004). En concreto nuestra sala tiene un  $T_{60}$  a 1000 Hz < 0,25 s y unas dimensiones de 9,1 x 4,45 x 2,64 m, lo que le confiere una planta de 40,5 m<sup>2</sup> y un volumen de 107 m<sup>3</sup>.

Aunque como se ha dicho nuestra sala se ajusta a los parámetros descritos en la norma (EBU Tech 3276-2nd Edition, 1998), resulta imposible ajustarse a algunas de las otras consideraciones que contempla la norma. Esto es debido simplemente a que no hay descritos parámetros que tengan en cuenta las características propias de los sistemas de sonido WFS ni VBAP, ni siquiera en el suplemento multicanal a la norma. Por tanto no se pueden respetar los parámetros que especifican la ubicación de los altavoces o la posición del oyente respecto a estos, porque no son congruentes con nuestro experimento.

Los **equipos de medida para las comprobaciones y ajustes** del montaje fueron:

- Micrófono omnidireccional Earthworks M30 (ver anexo I figura A-6)
- Tarjeta de sonido externa MOTU Traveler (ver anexo I figura A-7)



- Ordenador portátil
- Sonómetro Brüel & Kjaer 2250 (ver anexo I figura A-8)

El **software para la síntesis de sonido, procesado y análisis de los resultados** fue:

- SuperCollider 3.5.5 para la síntesis de sonido en WFS y VBAP, así como el control del test.
- Adobe Audition CS6 y Matlab 2012b para la edición de los sonidos empleados y la ecualización de los sistemas.
- SPSS (*Statistical Product and Service Solutions*), Matlab 2012b y Excel 2010 para el análisis de los datos.

SuperCollider es un entorno y un lenguaje de programación para la síntesis de audio en tiempo real y para la composición algorítmica. Consiste en un lenguaje interpretado orientado a objetos que funciona como cliente de red, para comunicarse con un sintetizador de sonido en tiempo real que funciona como servidor. Actualmente SuperCollider es un proyecto de *software* libre (*open source* GPL) que emplean tanto músicos como científicos y artistas para desarrollar sus trabajos (SuperCollider.net)

Adobe Audition CS6 es un programa de edición y mezcla de sonido profesional que incluye todo tipo de recursos de audio comunes. Filtros, ecualizadores, panoramizadores, procesadores de dinámica, herramientas para la edición, sincronización, limpieza del sonido, etc.

El programa Matlab 2012b maneja el lenguaje de alto nivel Matlab y proporciona un entorno interactivo para el cálculo numérico, la visualización de datos y la programación. Mediante Matlab es posible analizar datos, desarrollar algoritmos y crear modelos o aplicaciones. (Sitio web de Matlab)

Dentro del famoso paquete de herramientas ofimáticas Office, encontramos el programa Excel, muy útil para gestionar archivos de datos numéricos, realizar cálculos y representaciones gráficas.

SPSS Statistics (*Statistical Product and Service Solutions*) es un paquete de *software* para el análisis estadístico de datos (Sitio web de IBM SPSS Statisticis). Entre otros muchos factores estadísticos permite calcular parámetros de fiabilidad, relaciones cruzadas, análisis de varianza, correlaciones, medias, regresiones, etc.

#### 4.4 – Diseño del test

En base a la clasificación de los tests psicoacústicos descrita en el apartado 4.1, y después de leer y comparar diferentes artículos sobre pruebas de evaluación de la distancia con el sonido, se tomó la decisión de construir un test de percepción con respuestas de **escala directa por intervalos**. Este tipo de prueba permite realizar análisis mediante métodos cuantitativos, basados en la asunción de una distribución normal, permitiendo hacer pruebas-t y ANOVAs (Bech & Zacharov, 2006).

Además, debido a que no existe una normativa específica para realizar test de percepción de la distancia con el sonido, se tomaron como referencia otras normas y recomendaciones que tienen que ver con evaluación subjetiva de sonido (Rec UIT-R BS.1116-1, 1997), (Rec UIT-R BS. 1284, 2003), (EBU Tech 3276-2nd Edition, 1998).

La mecánica básica del test consistió en **comparar los sonidos reproducidos por los sistemas WFS y VBAP con una referencia externa**. Los sistemas WFS y VBAP sintetizaban las distintas distancias por medio de unos arrays de altavoces situados en una posición fija con respecto al oyente, mientras que la referencia externa se reproducía por un altavoz independiente situado a una distancia fija. De esta manera, cada sonido sintetizado a una distancia concreta (ya fuese por WFS o VBAP) debía ser comparado y evaluado con otro sonido reproducido a una distancia fija por el altavoz externo. Esta mecánica de comparación con una referencia externa real ha sido utilizada en otros experimentos de evaluación subjetiva de la distancia sonora (Kerber, Wittek, Fastl, & Theile, 2004) (Nielsen, 1991) (Wittek, Kerber, Rumsey, & Theile, 2004).

Los arrays de altavoces de que se disponía tenían un tamaño de 144 cm de largo. Se querían colocar los arrays en una posición fija alrededor del oyente, con éste situado en su centro. Dadas las dimensiones de la sala, con un pequeño cálculo se podía comprobar que la disposición más óptima de los arrays era en forma de octógono alrededor del oyente. (Ver figura 4.4-1)

A partir de la fórmula del diámetro inscrito en un polígono regular de n lados (siendo “a” la longitud del lado):

$$d_{inscrito} = a \frac{1}{\operatorname{tg}\left(\frac{\pi}{n}\right)} \quad \rightarrow \quad d_{inscrito}^{octógono} = 144 \frac{1}{\operatorname{tg}\left(\frac{\pi}{8}\right)} = 347,64\text{cm}$$

Y sumándole la medida del fondo de los arrays:

$$347,64 + 2 \cdot 29 = 405,64\text{cm} \quad \text{que es una longitud que se adapta bien en nuestra sala.}$$

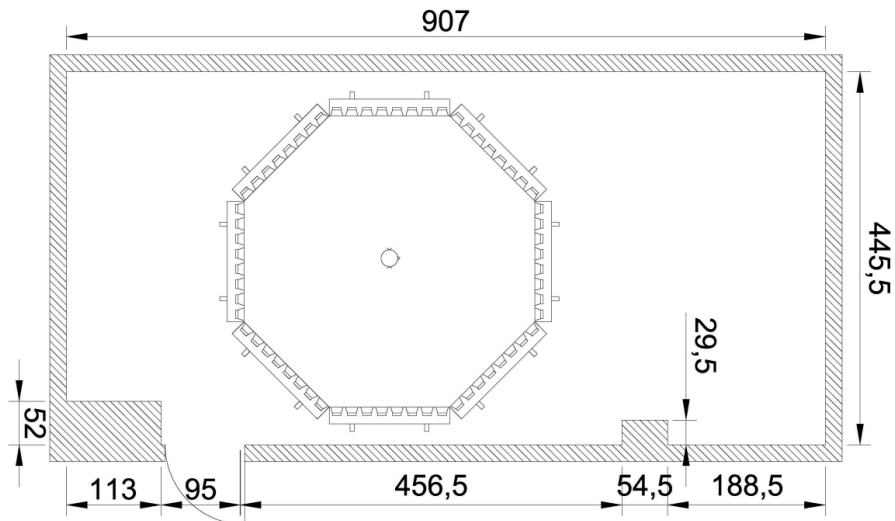


Figura 4.4-1 Esquema de la sala con arrays en forma de octógono

Para realizar el test se diseñó una **interfaz gráfica** que guiara a los sujetos durante el proceso. Con esta interfaz gráfica los parámetros internos del test se mantendrían ocultos (test ciego) mientras que el sujeto podía estar atento a la tarea encomendada. Por motivos prácticos se pensó en mostrar los sonidos a evaluar de dos en dos en el interfaz gráfico. Podían haberse presentado de cinco en cinco, pero podía inducir a confusión en el proceso de comparación con la referencia, o podían haberse presentado de uno en uno, pero hubiera resultado demasiado lento.

Se tomaron **siete distancias** (posiciones) a comparar con la posición de referencia: tres delante y tres detrás de la referencia, más otra posición coincidente con la distancia de referencia. Estas posiciones no iban a ser evaluadas por los sujetos que hiciesen el test mediante la asignación de un valor concreto en metros. De hecho hay experimentos que no recomiendan la respuesta de los sujetos mediante la asignación de un valor numérico concreto en metros, debido entre otras cosas a que no todo el mundo tiene una buena apreciación de cual es la medida exacta de un metro (Neher, Rumsey, & Brookes, 2002). Por este motivo se optó por una escala fija de siete posiciones con valores subjetivos: muy cerca, más cerca, un poco más cerca, igual a la referencia, un poco más lejos, más lejos, muy lejos. (Ver figura 4.4-2)

Las **cuestiones** que cada sujeto sometido al test debía responder eran dos:

- Indica a qué distancia te parece que está el sonido
- ¿Qué seguridad tienes?

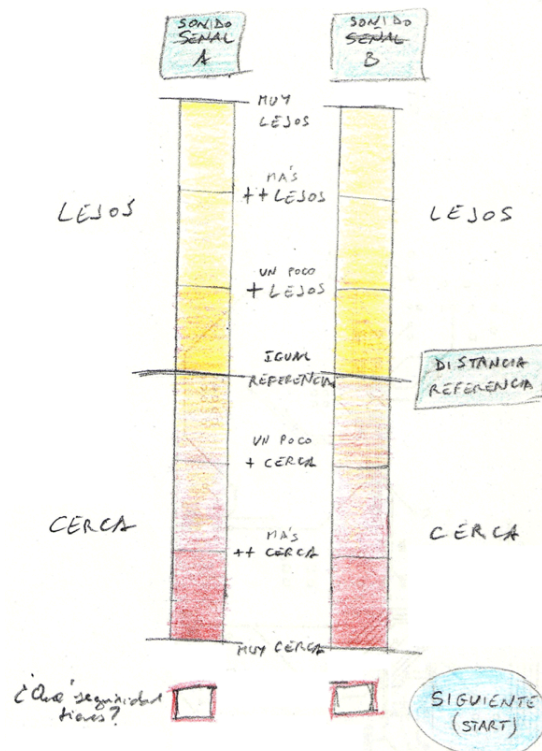


Figura 4.4-2 Boceto de la interfaz gráfica

Para la primera cuestión los sujetos debían responder cuánto de lejos o de cerca les parecía que estaba el sonido a evaluar con respecto al sonido de referencia, y podían elegir entre uno de los siete valores de la escala descrita antes (muy cerca, más cerca, un poco más cerca, igual a la referencia, un poco más lejos, más lejos, muy lejos).

Con la segunda cuestión los sujetos debían indicar la seguridad que tenían en su respuesta, o la confianza en su respuesta que les daba el sonido escuchado, valorándolo en una escala del 1-poco al 5-mucho. Con esta segunda cuestión se pretendía obtener alguna medida para diferenciar la confortabilidad en la escucha de cada sistema, o para detectar alguna distancia en la que hubiese más dudas. En definitiva, se buscaba evaluar la robustez de los sistemas y las distancias sintetizadas.

Cada uno de los sonidos debía tener distintas características que ayudasen a responder las cuestiones planteadas en los objetivos del test (ver apartado 4.2). Así, las **características de los distintos sonidos** eran consecuencia directa de varias decisiones tomadas y controladas en el diseño del test:

- Valor objetivo de las distancias a sintetizar
- Tipos de sonido atendiendo a su naturaleza
- Uso de reverberación
- Ángulo de escucha

La primera característica a determinar era el **valor objetivo de las distancias a sintetizar**, en metros. Estos valores de distancia se relacionan directamente con dos elementos a tener en cuenta: por un lado el tamaño físico real de la sala donde se realizó el test, y por otro lado la relación del nivel sonoro entre las distintas posiciones. Atendiendo a estos dos elementos las siete distancias escogidas fueron: 1,74 – 2,46 – 3,47 – 4,9 – 6,92 – 9,77 – 13,81 (metros), siendo 4,9 m la posición de referencia (ver figura 4.4-3). El punto de partida para elegir estas distancias fueron las medidas reales de la sala y la posición real del array de altavoces. Según esto la primera posición (la más cercana al oyente) debía ser 1,74 m, ya que era la distancia entre la posición del oyente en el centro de los arrays y el borde de los arrays. La otra posición determinante era la distancia del oyente al altavoz de referencia, que se situó cerca del final de la sala a 4,9 m. De esta manera la posición de referencia, que debía coincidir con la del altavoz físico de referencia, se ajustaba a las dimensiones de la sala. A partir de estas medidas reales se tuvo en cuenta el segundo criterio de la relación de nivel sonoro entre las distintas distancias. Como ya se comentó en el apartado 3.2, el nivel de presión sonora es el principal factor que empleamos para discernir la distancia de un evento sonoro. Los valores de distancia escogidos siguen una relación logarítmica según la ecuación  $distancia_i = \left[ 0,2142 \cdot e^{0,3473(i+6)} \right] + 0,02$  (Los valores de las constantes se obtienen al ajustarse a las medidas reales de la sala). Esta distribución de distancias logarítmica está calculada para obtener una atenuación de 3dB para cada distancia más lejana que la anterior. Siguiendo este criterio, empleado en otros experimentos (Zahorik, 2002), se consigue una atenuación de nivel que se asegura que sea percibida por los sujetos que realizan el test.

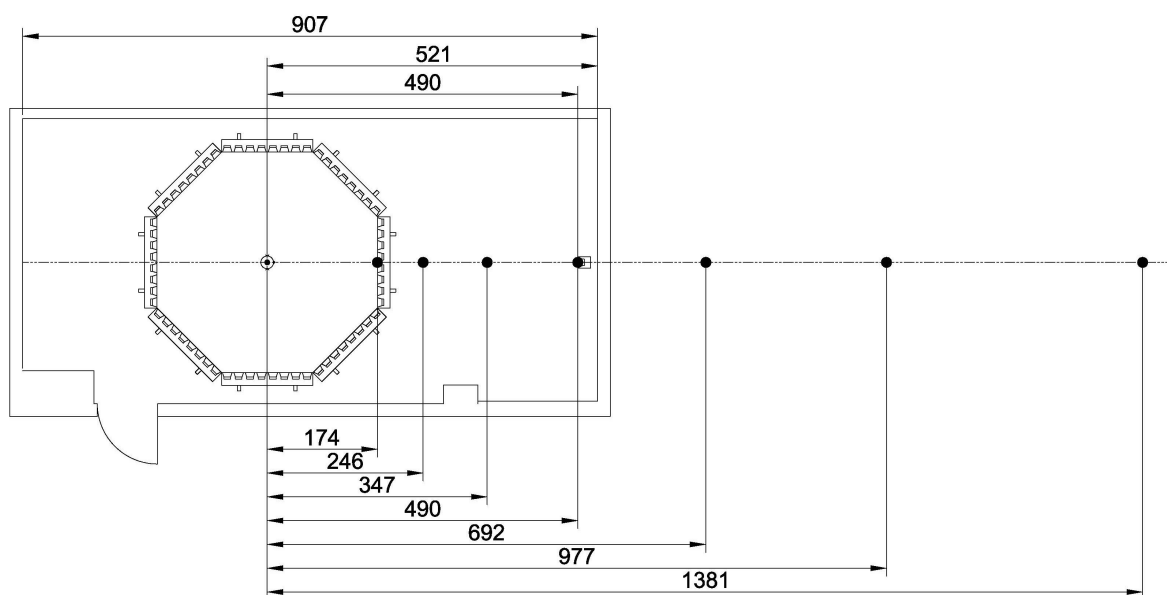


Figura 4.4-3 Distancias de evaluación del test

La elección de estas distancias también implica desechar la atenuación en altas frecuencias como factor para localizar en distancia. Nuestra distancia más lejana es de 13,81 m, y es partir de 15 m cuando se considera relevante la atenuación de las altas frecuencias (ISO 9613-1:1993(E))

Otra característica importante del sonido que puede afectar a la percepción de la distancia es el **tipo de sonido atendiendo a su naturaleza**. Según el sonido sea o no conocido por el oyente, según sea un sonido sintético o uno existente en la naturaleza, según el contenido espectral del propio sonido, la percepción de la distancia se puede ver influida. Así, se eligieron los siguientes tipos de sonido:

- Ruido Rosa
- Voz
- Instrumento acústico → Guitarra
- Puerta cerrándose

El **ruido rosa** es un tipo de sonido sintético de amplio espectro en el que la energía se distribuye uniformemente según bandas de octava, de manera que su distribución en frecuencia es constante en una escala logarítmica. El oído humano utiliza ese tipo de escala a través de bandas de octava como patrón de integración de señales en frecuencia, por lo que la percepción de frecuencias será más fiable para bajas frecuencias. Resulta interesante además el hecho de que al tratarse de un sonido sintético, no se tiene una referencia real de cómo debe sonar según su localización. El ruido rosa empleado era una ráfaga de 1 s de duración con transitorios inicial y final de 50ms para evitar posibles distorsiones en la reproducción.

La **voz** es un caso particular de señal sonora a la que todo el mundo se encuentra muy habituado. En este tipo de señales la mayor parte de la energía se encuentra por debajo de los 7kHz, y la mayor parte del contenido que aporta información sobre la posición de la fuente se sitúa en la zona entre 1-3kHz (Begault, 2000). La muestra de voz empleada era la grabación de las sílabas “ba-be-bi-bo-bu”, de una duración de 2,6 segundos. Esta muestra es interesante dado su contenido vocálico diverso y la presencia de la consonante bilabial “b” con cierto contenido impulsivo.

Se utilizó también un instrumento acústico común como es la **guitarra** para que con toda probabilidad todos los sujetos hubieran tenido la experiencia previa de haber oído una guitarra acústicamente, es decir, al natural, sin ningún tipo de amplificación electroacústica, y así resultase un sonido familiar. Además, al tratarse de un instrumento de cuerda, el sonido resultante está compuesto de varios sonidos impulsivos (amplio espectro), uno por

cada cuerda rasgada. El sonido impulsivo, con transitorios muy abruptos y contenido espectral continuo, resulta tener muy buena localización espacial (Blauert, 1996). La muestra de guitarra empleada era la grabación de un acorde flamenco en re bemol, de 2,8 s de duración, donde 0,5 s eran el ataque del acorde y 2,3 s la parte correspondiente a la relajación de las cuerdas.

De la misma manera que con la guitarra, el sonido de una **puerta cerrándose** es un sonido impulsivo conocido, en este caso de muy corta duración y con bastante contenido en bajas frecuencias. La muestra empleada tenía una duración de 0,5 s.

Todos los sonidos empleados en el test eran grabaciones mono a 44100Hz y 16 bits.

La **reverberación** es otra característica importante descrita para la localización sonora en distancia (Begault, 2000) (Blauert, 1996). En nuestro test de percepción subjetiva se presentaron los mismos estímulos sonoros con y sin reverberación. Se decidió que la reverberación escogida fuese lo suficientemente suave como para no establecer una diferencia abrupta entre los sonidos sin y con reverberación.

La última característica tenida en cuenta en el diseño de nuestro test fue el **ángulo de escucha**. La dirección de procedencia del sonido determina con respecto al oyente un ángulo que influye enormemente en la localización espacial (Begault, 2000) (Blauert, 1996). Según este ángulo la información que llega a cada oído se presenta ligeramente distinta provocando diferencias interaurales de tiempo (ITD) y de nivel (ILD). En nuestro experimento se eligieron dos ángulos de procedencia del sonido a estudiar, 0° es decir, escucha frontal, y 90°, escucha lateral. La elección de los 90° se tomó para tener una representación de la máxima diferencia interaural.

Teniendo en cuenta todas las características descritas de los distintos sonidos a emplear en el test, se puede calcular el **número de estímulos** a presentar a cada sujeto:

$$7 \begin{matrix} \text{distan} \\ \text{cias} \\ \text{sin} \\ \text{tetizadas} \end{matrix} \times 4 \begin{matrix} \text{tipos} \\ \text{de} \\ \text{sonido} \end{matrix} \times 2 \begin{matrix} \text{con} \\ \text{sin} \\ \text{reverberación} \end{matrix} \times 2 \begin{matrix} \text{ángulos} \end{matrix} \times 2 \begin{matrix} \text{sistemas} \\ \text{WFS} \\ \text{VBAP} \end{matrix} = 224 \text{ estímulos}$$

Además de estos estímulos se incluyeron varios sonidos para que actuaran como **referencias ocultas**. Las referencias ocultas son un medio de control útil para discriminar un cierto grado de fiabilidad de los sujetos que realizan el test. Si un sujeto no vota correctamente la mayoría de las referencias puede ser porque esté votando aleatoriamente, porque no haya entendido correctamente el test o porque por algún motivo no escucha correctamente. Dado que la mecánica de nuestro test de percepción subjetiva consiste en comparar cada sonido o estímulo sintetizado a una distancia por WFS o VBAP, con el

mismo sonido reproducido por el altavoz de referencia, las referencias ocultas de nuestro test consistían en una muestra de cada tipo de sonido reproducida por el mismo altavoz de referencia. Es decir, un sonido de referencia (reproducido por el altavoz de referencia) se comparaba de manera oculta (ciega), sin que los sujetos lo supiesen, con la propia referencia del test, que era en cualquier caso el mismo sonido reproducido por el altavoz de referencia.

Estas referencias ocultas también son útiles para determinar un cierto grado de fiabilidad del propio test, ya que lo lógico es que estos estímulos se voten de manera correcta (es decir, que los sujetos acierten la distancia y la voten igual que la referencia) en la mayor parte de los casos. Se incluyeron un total de 8 referencias ocultas para cada ángulo de escucha (0° y 90°). Estas referencias eran cada uno de los tipos de sonido (ruido rosa, voz, guitarra y puerta) que se presentaban ocultos mezclados con el resto de estímulos dos veces:  $4_{\substack{\text{tipos} \\ \text{de sonido}}} \times 2_{\text{repeticiones}} = 8 \text{ referencias ocultas}$  para cada ángulo.

Por razones prácticas el número de estímulos se dividió en dos partes, presentando en una primera parte los estímulos a 0° del plano horizontal con respecto al oyente, y en una segunda parte a 90° del plano horizontal. Así el número de estímulos presentado en cada parte era de 112 + 8 referencias ocultas.

Por tanto, el **número total de estímulos** a presentar a cada sujeto era de 240 estímulos, presentados en dos partes de 120 estímulos para cada ángulo de escucha.

$$\left( 112_{\text{estímulos}} + 8_{\substack{\text{referencias} \\ \text{ocultas}}} \right) \times 2_{\text{ángulos}} = 240 \text{ estímulos}$$

Se realizó una **aleatorización de los estímulos** para que su ordenación no afectase a los resultados y para que los posibles efectos de aprendizaje, cansancio o interacción entre estímulos contiguos se dispersasen lo suficiente como para obtener resultados válidos (Bech & Zacharov, 2006). Como ya se ha comentado antes, por razones prácticas se presentaron primero los estímulos a 0° y después a 90°. Era necesario por dos motivos, primero porque era posible girar al oyente 90° e imposible girar el montaje del test, y segundo para establecer una pausa de descanso dado el gran número de estímulos totales. Se eligió además el ángulo como característica para dividir el test en dos partes porque al ser un parámetro muy evidente de identificar no se estaba forzando ninguna influencia a la hora de dar las respuestas.



La elección de los sujetos que realizarán la prueba es también un factor importante en el diseño de un test de percepción subjetiva. En una primera fase de **preselección de sujetos**, se pretende de aquellos que participen, no que sean expertos en sonido, sino que tengan experiencia en la audición del sonido por sí mismo, ya sea porque hayan hecho más test de percepción subjetiva o porque por afición, profesión o formación estén acostumbrados a prestar especial atención al sentido del oído. Además se estima que con un mínimo de 20 sujetos el tratamiento estadístico de los resultados obtenidos puede tener suficiente fiabilidad (Bech & Zacharov, 2006) (Rec UIT-R BS.1534-1, 2003).

#### 4.5 – Implementación

Llevar a la práctica todo el diseño previo del test supuso el desarrollo de aplicaciones *software* y el montaje de todo el sistema con sus ajustes necesarios.

Para llevar un control unívoco de cada uno de los distintos sonidos que se presentaban en el test, se hizo una **caracterización numérica de cada estímulo y de sus respuestas asociadas**. Cada estímulo reproducido a los sujetos que realizaban el test venía determinado por las diferentes características descritas en el apartado anterior (4.4), que eran: la distancia a sintetizar, el tipo de sonido atendiendo a su naturaleza, el uso de reverberación, el ángulo de escucha. A estos descriptores hay que añadir el sistema de reproducción, que podía ser WFS, VBAP o también el altavoz real de referencia (denominado REF) en el caso de que se reprodujera una de las referencias ocultas. Además también se caracterizaron numéricamente las respuestas que daba cada sujeto asociadas a cada uno de los estímulos. Según el diseño del test descrito en el apartado anterior para cada estímulo se planteaban dos cuestiones: “Indica a qué distancia te parece que está el sonido”, cuya respuesta se denominó “distancia respuesta”, y la pregunta “¿Qué seguridad tienes?”, cuya respuesta se denominó “seguridad”.

Para simplificar la caracterización numérica de las distancias reproducidas se optó por asociar un número de posición a cada una de ellas, según la tabla 4.5-1.

	distancia/ posición de Referencia						
Distancias reales (m)	1,74	2,46	3,47	4,9	6,92	9,77	13,81
Posición	0	1	2	3	4	5	6

Tabla 4.5-1 Número de posición asignado a cada distancia

De la misma manera se caracterizaron numéricamente todas las variaciones posibles de las distintas características de cada estímulo y de sus respuestas, y a cada característica se le asoció un dígito. Así el primer dígito caracterizaba el tipo de sonido, el segundo la distancia sintetizada, el tercero la reverberación, el cuarto el sistema de reproducción, el quinto el ángulo de escucha, el sexto la distancia respondida por el sujeto y el séptimo el valor de seguridad asociado a la distancia respondida. Esta asignación, con todos los valores posibles, se puede ver resumida en la tabla 4.5-2.

Estímulo					Respuestas	
Tipo	Distancia sintetizada	Reverb	Sistema	Ángulo	Distancia	Seguridad
0 R Rosa	0	0 sin	0 WFS	0 - 0°	0	1 (poco)
1 Voz	1	1 con	1 VBAP	1 - 90°	1	2
2 Guitarra	2		2 REF		2	3
3 Puerta	3				3	4
	4				4	5 (mucho)
	5				5	
	6				6	

Tabla 4.5-2 Caracterización numérica de cada estímulo y sus respuestas asociadas

Por ejemplo, 0 2 0 0 1 3 5 correspondería a un estímulo de tipo ruido rosa, posicionado a la distancia 2 (3,47m), sin reverberación, reproducido por el sistema WFS, en el ángulo de escucha 90° y al que se ha respondido con una posición de distancia 3 (4,9m), con una seguridad de 5 (mucho).

Gracias a esta asignación de valores se podían tratar todos los datos de manera rápida y unívoca, utilizando un formato común de archivo CSV. Así, esta caracterización numérica fue muy útil para el posterior análisis y tratamiento estadístico de los resultados.

La reproducción de los distintos sonidos y la interfaz gráfica para manejar el test fue realizada mediante el **entorno de programación SuperCollider**. Toda la programación necesaria para implementar este sistema fue realizada por Andrés Pérez López, que era personal del iTeam durante la realización de este TFG, siguiendo el diseño realizado por el autor de este TFG.

El entorno SuperCollider era idóneo para llevar a cabo la tarea debido a la existencia de la librería WFSCollider (Negrao, 2011), un conjunto de código que implementa la espacialización de sonido en tiempo real mediante WFS.

Para el desarrollo de este trabajo, se tomaron partes del código de la librería WFSCollider, que está licenciada mediante GPL, y se adaptó a las necesidades específicas del desarrollo. Por otra parte, la estructura de la sala donde se realizó el test hace que el ordenador conectado al sistema de audio se encontrara en un lado, mientras que la persona sujeta al test debía estar en el centro de la sala. La arquitectura de SuperCollider hace posible conectar varios ordenadores de una forma muy sencilla.

También se deseaba que la persona que realizara el test interactuara de una forma cómoda. Para ello, se escogió realizar la interacción mediante un *joystick* o mando de videojuegos. SuperCollider provee de librerías específicas a tal efecto.

La estructura de la implementación en SuperCollider se puede resumir con la figura 4.5-1. La parte superior del esquema corresponde a la implementación propiamente, mientras que la parte inferior corresponde con el sistema de audio montado en la sala. El sujeto que realizaba el test se comunicaba con el ordenador A mediante un *joystick*, todo ello situado junto a él en el centro del octógono descrito por los altavoces. A su vez, el sujeto recibía el *feedback* del desarrollo de la prueba a través del monitor del ordenador A, mediante una Interfaz Gráfica de Usuario (GUI) creada a tal efecto y, por supuesto, también recibía como *feedback* el audio que percibía. El ordenador A recibía las órdenes del sujeto bajo estudio, y las transmitía mediante el protocolo OSC (The Center for New Music and Audio Technology (CNMAT)) al ordenador B. El ordenador B se encargaba de sintetizar el sonido deseado de acuerdo a las órdenes, y lo emitía a través de su interfaz de audio, llegando así al sujeto, que podría entonces emitir un juicio al respecto y responder a las cuestiones planteadas.

En realidad se implementaron hasta tres programas diferentes para poner en marcha el sistema. Dos que trabajaban conjuntamente para realizar el test (aplicaciones “test” y “render”) y un tercero (aplicación “monitor”) que se utilizaba para probar y ajustar el sistema.

Los dos programas que se utilizaban para realizar la prueba se ejecutaban simultáneamente, cada uno en uno de los ordenadores:

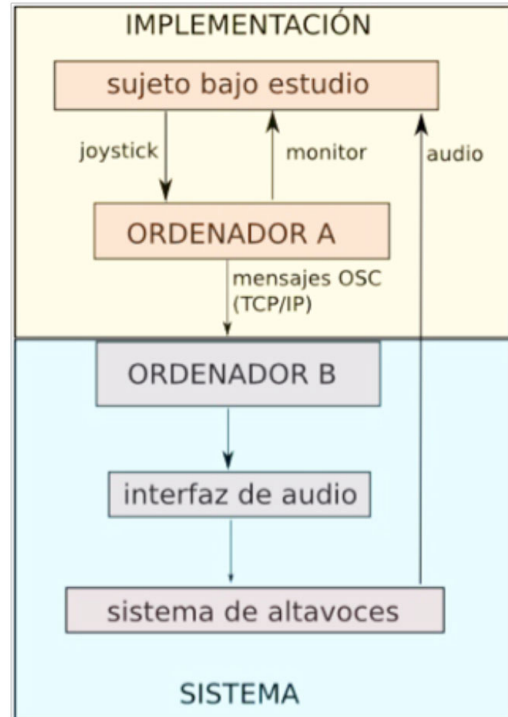
Ordenador A – aplicación “test”. Este programa implementaba cuatro funciones:

- . Interfaz gráfica para el sujeto bajo test (figura 4.5-2)
- . Envío de mensajes con la configuración del sonido al Ordenador B
- . Control del joystick

- Recolección de los resultados para el posterior análisis

Para cada test realizado por un sujeto se creaba un fichero de texto, en el cual se iban escribiendo los resultados del test para su posterior estudio. Los datos se guardaban en formato CSV para facilitar su tratamiento.

Figura 4.5-1 Estructura de la implementación en SuperCollider



indica a qué DISTANCIA te parece que está el sonido

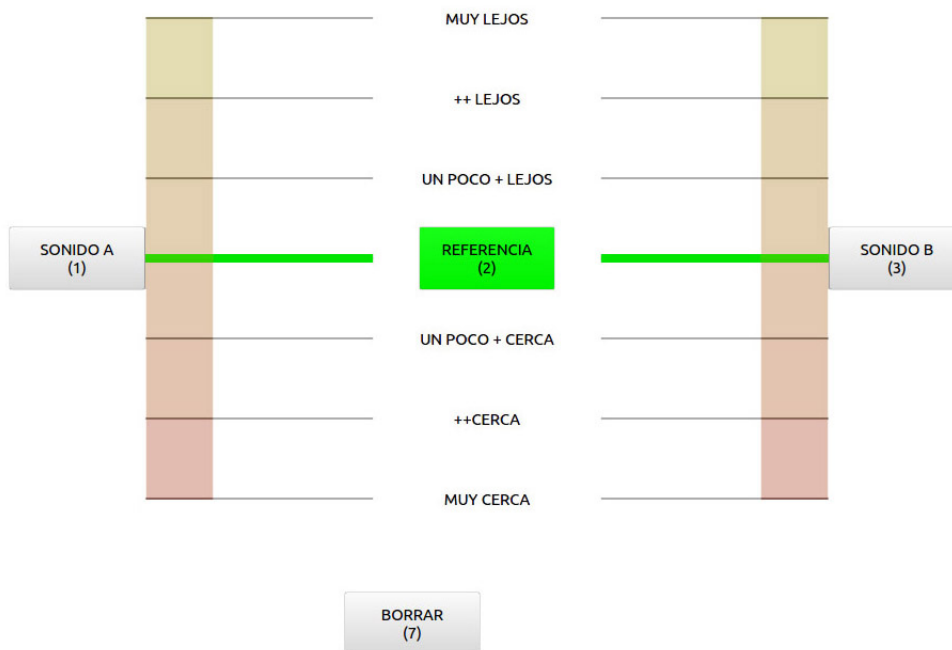


Figura 4.5-2 Captura de pantalla de la interfaz gráfica del programa “test” que manejan los sujetos durante la prueba

Ordenador B – aplicación “render”. Este programa implementa dos funciones:

- Recibir los mensajes desde el Ordenador A
- Interpretar los mensajes para su transformación en audio

Los mensajes consistían en la caracterización numérica de cada estímulo descrita antes, así se le indicaba al programa “render” qué y cómo tenía que reproducir en cada momento. Además, con el programa “render” se podía hacer un seguimiento de la prueba y comprobar que el sistema funcionaba correctamente. (Ver figura 4.5-3)

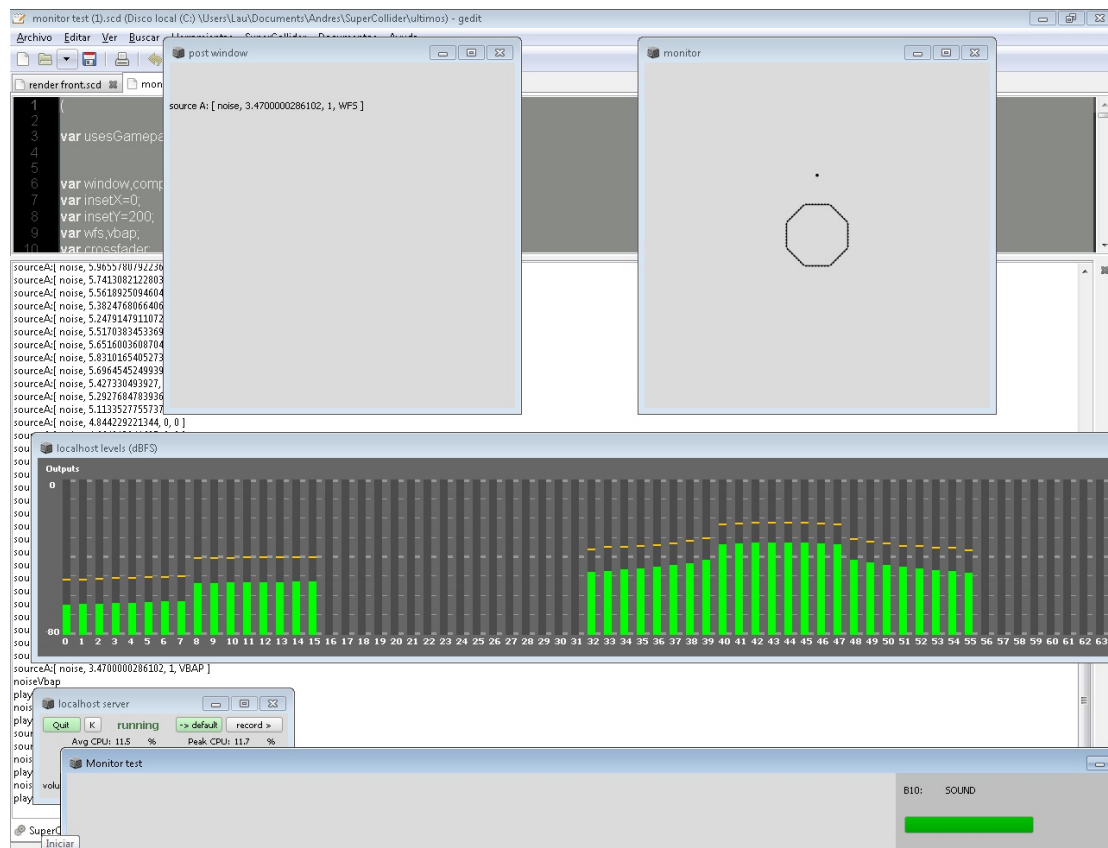


Figura 4.5-3 Captura de pantalla del programa “render” funcionando

El tercer programa, llamado “monitor”, se utilizó para las pruebas del sistema. El programa “monitor” era complementario del programa “test” y, al igual que éste, para su funcionamiento debía estar ejecutándose el programa “render” en el Ordenador B. Básicamente era igual que “test”, y compartía el mismo tipo de mensajes a enviar al “render”. La diferencia radica en que en el programa “monitor” el usuario puede controlar y elegir los parámetros del sistema, mientras que en “test” estos parámetros están prefijados y, por supuesto, son desconocidos para el sujeto bajo test. “Monitor” ha sido de gran ayuda para realizar los ajustes necesarios para el correcto funcionamiento del test psicoacústico. (Ver figura 4.5-4)

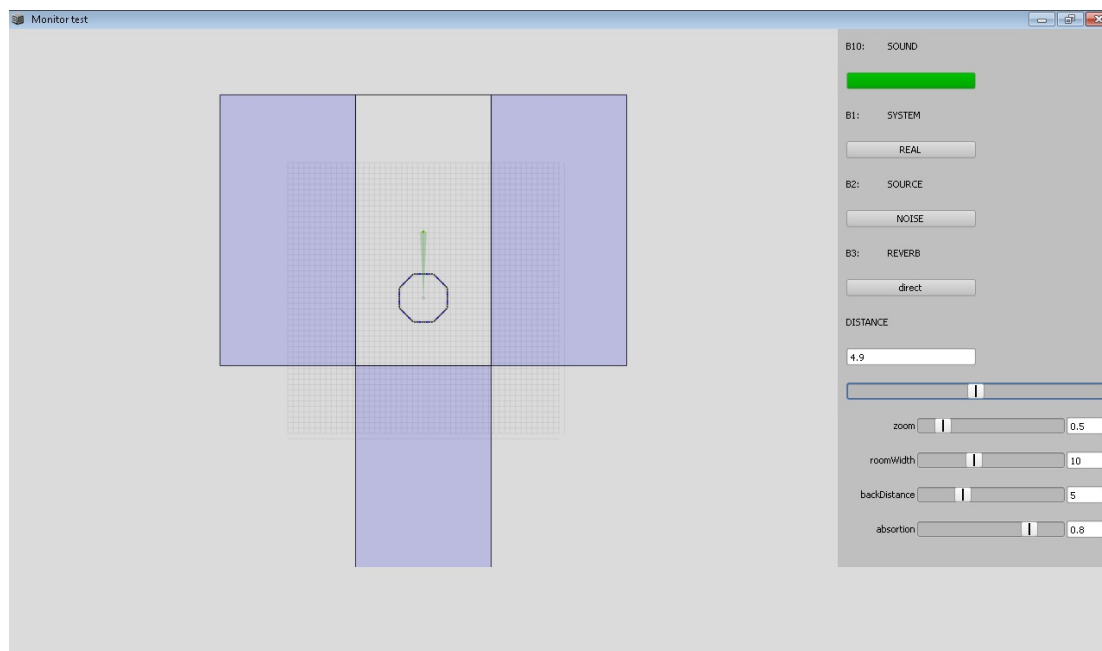
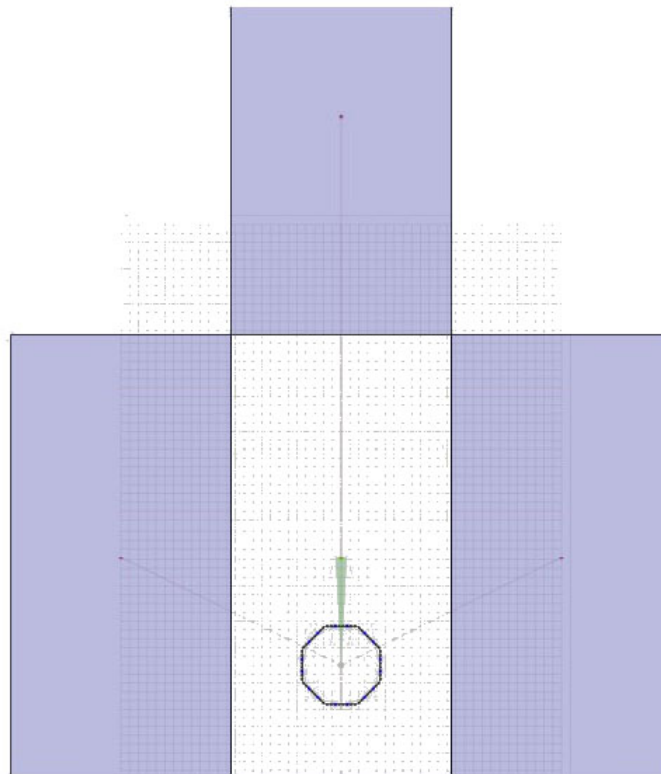


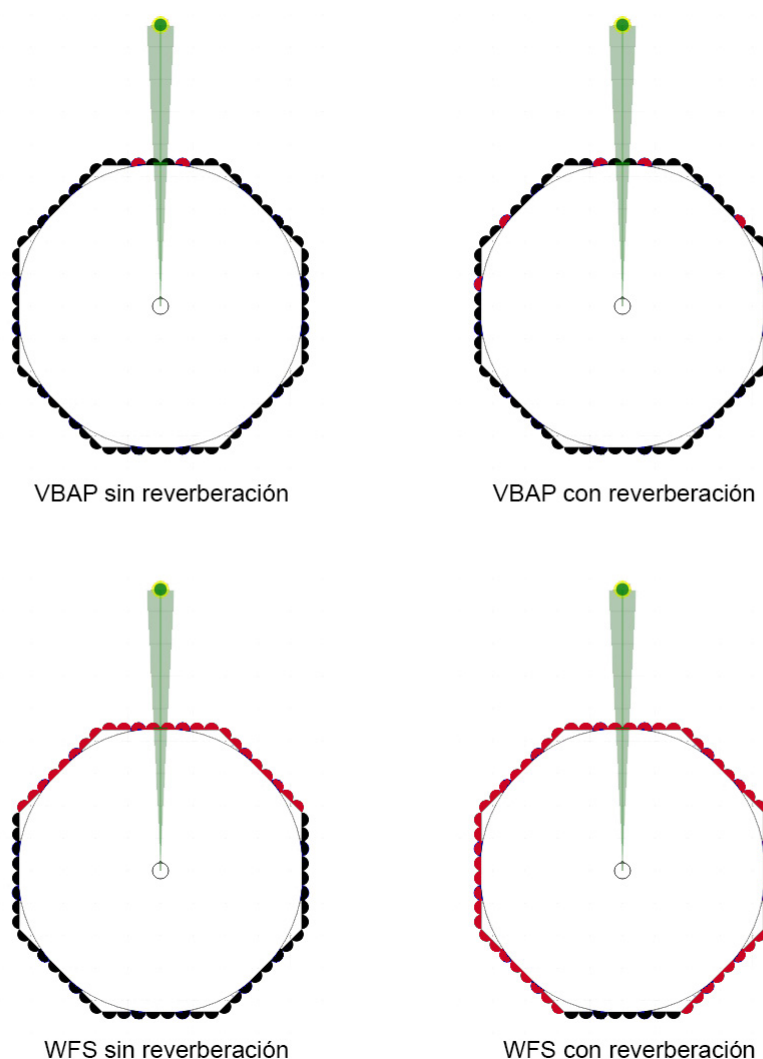
Figura 4.5-4 Captura de pantalla del programa “monitor” funcionando

La **reverberación** empleada sobre ciertos de los estímulos presentados a los sujetos que realizaban el test, la sintetizaba también el *software* SuperCollider. Esta reverberación se basaba en el método de las imágenes, técnica ya empleada antes en sistemas WFS (Juan & Yonghong, 2011). Según este método las reflexiones se simulan como fuentes secundarias situadas en posiciones simétricas con respecto a cada uno de las paredes de la sala, atenuadas según el coeficiente de estas paredes. En nuestro caso la potencia del ordenador de síntesis limitaba el número de reflexiones a tres, con lo que se sintetizaron las reflexiones provenientes del frente y de los lados del oyente (ver figura 4.5-5). En estas fuentes secundarias se introdujo una atenuación del 20%, es decir, cada pared atenuaba un 20% de la señal en la primera reflexión, con un tamaño de sala virtual de 20 x 10 m y el oyente situado a 5 m de las paredes. Además, para conseguir una reverberación más natural y simular las últimas reflexiones se añadió un efecto de reverberación mono sobre estas fuentes secundarias sintetizadas. Para ello se empleó un efecto de reverberación incluido en las librerías de SuperCollider llamado *Freeverb*, con los siguientes parámetros: *room* 0,5, *mix* 0,5, *hf damp* 0,7. Se eligió que la reverberación global resultante, aunque distinguible con facilidad, no destacara demasiado para no introducir mucha diferencia entre los estímulos sin y con reverberación. El resultado global de la reverberación añadía una cola de 20ms a los sonidos reproducidos.



**Figura 4.5-5 Primeras reflexiones sintetizadas por SuperCollider con una atenuación del 20%**

A la hora de realizar el **montaje de los arrays** de altavoces que reproducirían los sistemas WFS y VBAP, hubo que tener en cuenta la altura de los mismos. Tras disponerlos en forma de octógono en su posición de la sala según el diseño previo, se bajaron de altura hasta 1,40 m. Aunque esta altura pueda parecer alta para el oyente, se utilizó una banqueta regulable para adaptar la altura de cada oyente a la altura del montaje, alineando los centros acústicos de los altavoces con la posición de los oídos de los sujetos. En el montaje de *arrays* en forma de octógono no todos los altavoces funcionaban para reproducir todos los estímulos. El número de altavoces que intervenían para reproducir cada estímulo dependía del sistema de reproducción, según fuese WFS o VBAP, y también de si el estímulo se reproducía con o sin reverberación. Si el sistema en funcionamiento era VBAP sin reverberación el número de altavoces que intervenían se reducía al mínimo, tan solo dos. Si era VBAP con reverberación intervenían seis altavoces. Si se reproducía WFS sin reverberación funcionaban veinticuatro altavoces. Y si era WFS con reverberación funcionaban simultáneamente cincuenta y seis altavoces, el máximo (ver figura 4.5-6).



**Figura 4.5-6** Altavoces que intervienen en la reproducción de los estímulos (en rojo) según se empleara WFS o VBAP, con o sin reverberación

Además de los *arrays* de altavoces, también se utilizó para reproducir sonidos el **altavoz de referencia**. Este altavoz, que se montó expresamente para realizar este test, se situó a la distancia de referencia de 4,9 m (posición 3). Al situarse detrás de los *arrays*, este altavoz podía sufrir un efecto sombra en el punto de escucha que podría notarse en las altas frecuencias, afectando a la percepción de la distancia. Para evitar este posible problema el altavoz de referencia se colocó en altura, calculando el ángulo necesario para que se respetase la distancia específica de referencia y la inclinación necesaria para que el centro acústico del altavoz estuviese dirigido al oyente (ver figura 4.5-7). Este procedimiento utilizado en otros tests de percepción subjetiva sobre distancia sonora (Nielsen, 1991)(Wittek, Kerber, Rumsey, & Theile, 2004) es posible gracias a que la capacidad humana para percibir sonidos en altura es muy limitada (Davis & Stephens, 1974)(Butler & Belenduij, 1977).



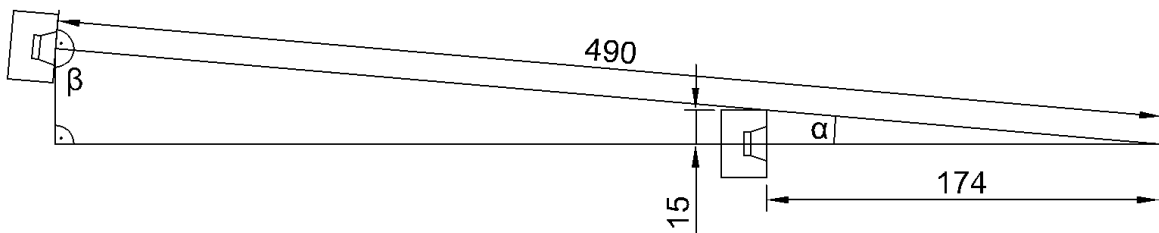
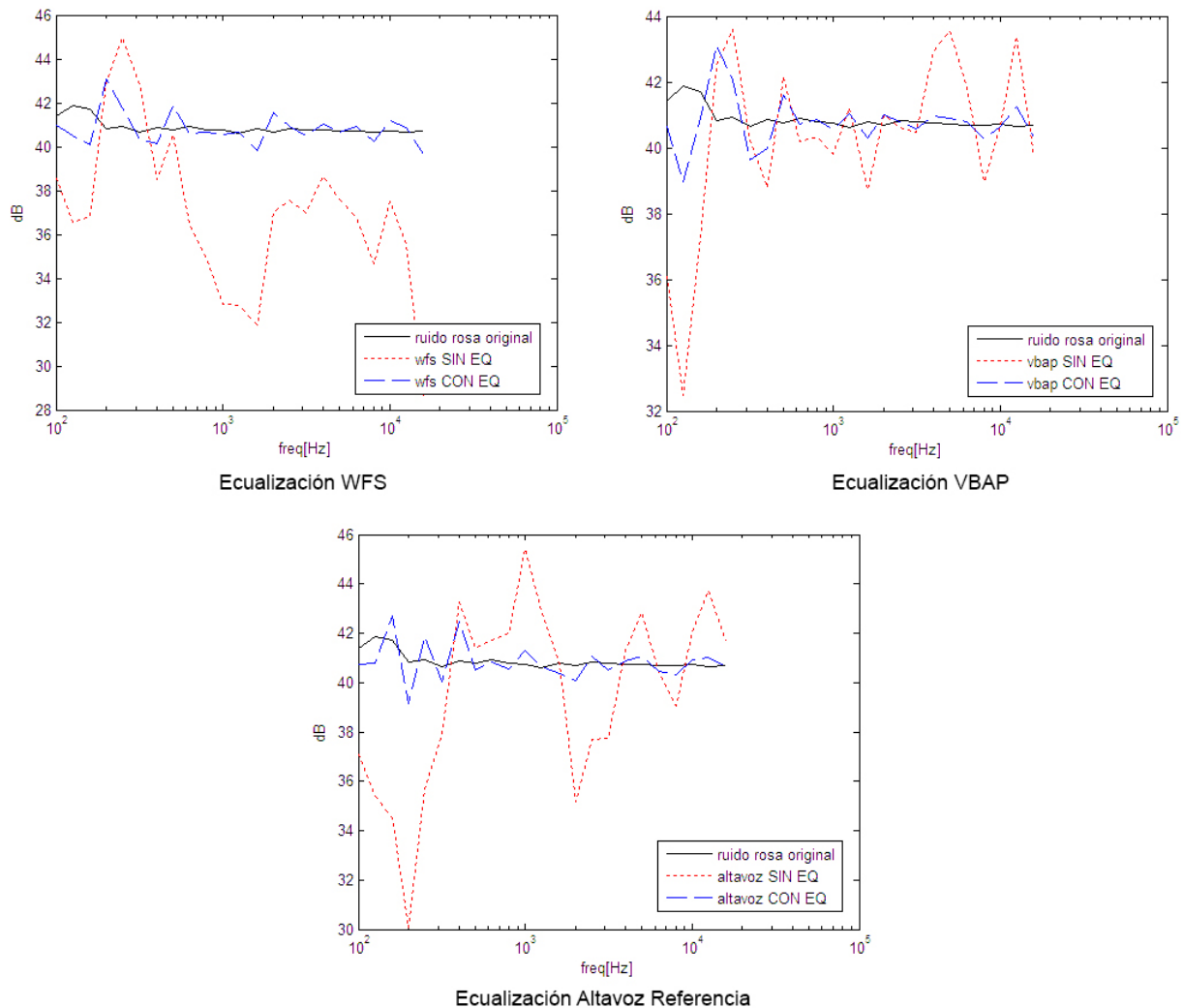


Figura 4.5-7 Fotografía del montaje del altavoz de referencia y diagrama del perfil de su posición con respecto a los arrays (Altavoz referencia a la izquierda del diagrama.  $\alpha=5^\circ$   $\beta=85^\circ$ )

Aunque todos los altavoces empleados en el montaje final eran del mismo tipo, el hecho de emplear diferentes sistemas de sonido (WFS, VBAP y el altavoz de referencia) introducía coloraciones en el sonido final. Para solucionar este problema se realizó una **ecualización de los sistemas y el altavoz de referencia**. Se emitió la misma señal de ruido rosa de 10 segundos de duración por cada uno de los sistemas y el altavoz de referencia y se tomó como distancia de ecualización la posición de referencia (4,9 m, posición 3). El sonido emitido por cada sistema fue grabado con un micrófono Earthworks M30 y una tarjeta de sonido externa MOTU Traveler conectada a un ordenador portátil. Las grabaciones obtenidas se analizaron mediante el *software* Matlab y para realizar los ajustes sobre las señales se empleó un ecualizador gráfico de 20 bandas incluido en el *software* Adobe

Audition CS6. Las señales resultantes se aplicaron al sistema correspondiente y se pudo comprobar como las coloraciones se redujeron drásticamente (ver figura 4.5-8).



**Figura 4.5-8 Gráficas de las equalizaciones de los sistemas WFS, VBAP y el altavoz de referencia**

Los mismos parámetros de equalización se aplicaron sobre todas las señales a emplear en el test, teniendo en cuenta los diferentes sistemas. El resultado obtenido fue una mejora e igualación sensible de los sonidos reproducidos por los distintos sistemas.

Para acabar de ajustar los sistemas se procedió a comprobar y **calibrar los niveles de reproducción**. Para ello se empleó un sonómetro Brüel & Kjaer 2250 colocado en la posición de escucha. El ajuste de nivel de los sistemas y el altavoz de referencia se hizo para la posición de referencia (4,9 m, posición 3). Para el ruido rosa se fijó un nivel de 69dBA (SPL). Los demás tipos de sonido elegidos eran sonidos existentes en la naturaleza (voz, guitarra y puerta) por tanto se quiso que tuvieran un nivel posible de encontrar en la realidad, para que sonaran de manera natural. Además se eligieron niveles que no difirieran

mucho y que fuesen coherentes entre sí. Los niveles ajustados para la posición de referencia fueron: voz-58dBA, guitarra-58dBA, puerta-62dBA. También se midió el nivel de todos los tipos de sonido en todas las posiciones sintetizadas, comprobando que se conseguía con bastante precisión el criterio de diseño de 3dB de diferencia entre una posición y otra. Los niveles del ruido rosa en las distintas posiciones sintetizadas por cada sistema puede verse en la tabla 4.5-3.

Distancia (m)	1,74	2,46	3,47	4,9	6,92	9,77	13,81
WFS (dBA)	81,2	78,2	72,5	68,8	65,1	62,1	59,1
VBAP (dBA)	78,2	76,8	72,2	69,4	66,4	63,4	60,4
Altavoz REF (dBA)				69,2			

Tabla 4.5-3 Niveles en dBA SPL del ruido rosa en todas las posiciones sintetizadas (WFS y VBAP) y en el altavoz de referencia

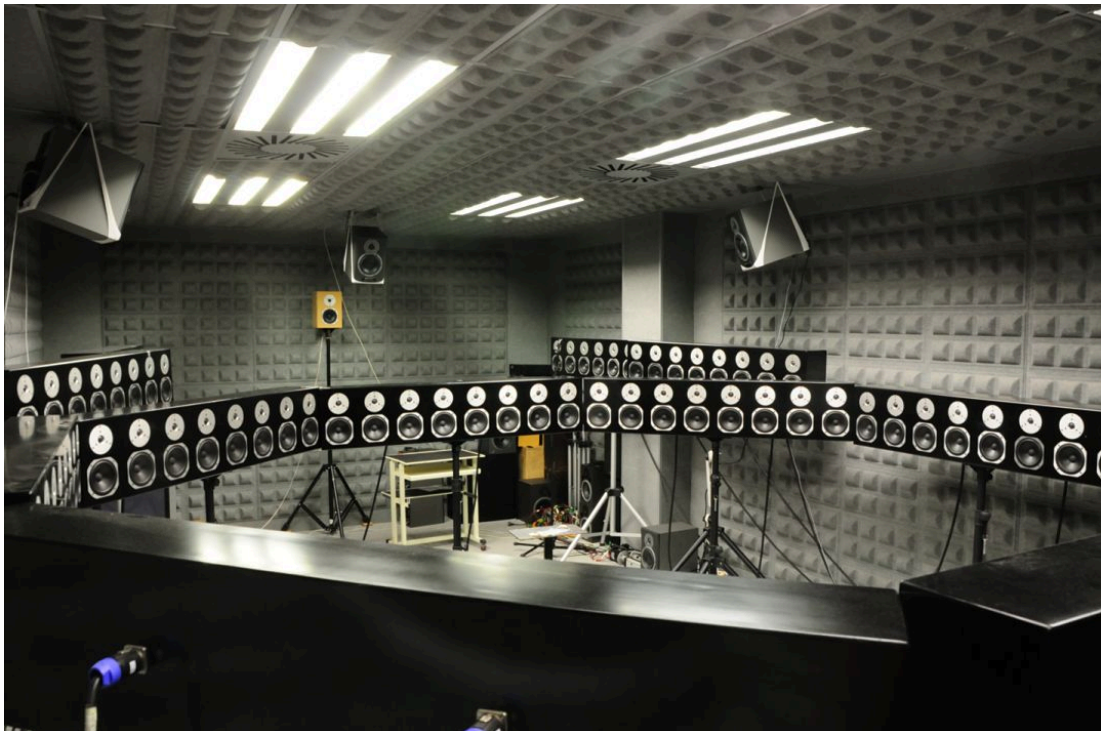


Figura 4.5-9 Fotografía del montaje de todos los altavoces, donde se pueden ver los *arrays* formando un octógono y el altavoz de referencia

La influencia de la visión sobre la escucha es capaz de alterar la votación sobre estímulos sonoros (Ehrenstein & Ehrenstein, 1999). Por este motivo se dispuso una **cortina acústica** delante del sujeto, de manera que éste no pudiera ver dónde se situaban los altavoces, ni siquiera cuántos de ellos había o dónde acababa la sala. Una cortina acústica

es una tela con la propiedad de ser transparente acústicamente pero opaca a la vista (Bech & Zacharov, 2006) (IEC. 60268-13, 1998). Gracias a este simple mecanismo y a un ligero control de la iluminación de la sala, se aseguró que los sujetos que realizaron el test no tuvieran influencias visuales que pudieran alterar sus respuestas (ver figura 4.5-11).



Figura 4.5-10 Fotografía del montaje del sistema durante la realización de unas pruebas

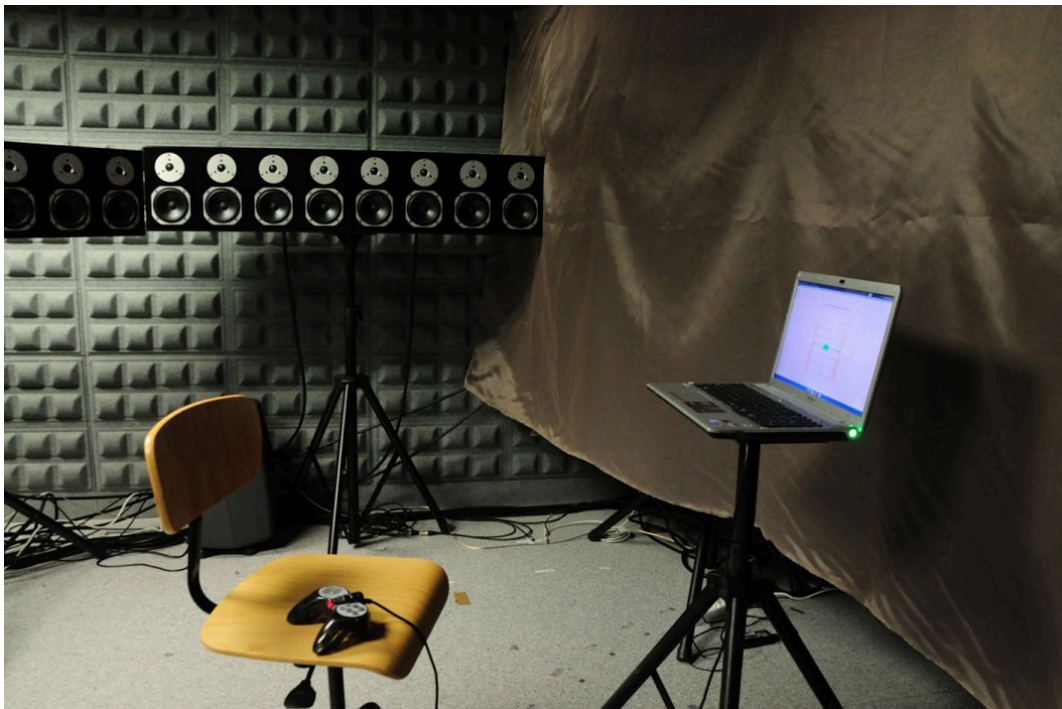


Figura 4.5-11 Puesto listo para realizar el test, donde se puede ver el ordenador portátil con la interfaz gráfica, el joystick y la cortina acústica.

## 4.6 – Procedimiento

Siguiendo el criterio establecido en el diseño del test, al hacer la **selección de los sujetos** se buscó a personas que tuvieran experiencia en la audición del sonido por sí mismo. Finalmente se seleccionaron **25 personas** entre los que había trabajadores y alumnos del Gtac (Grupo de Tratamiento de Audio y Comunicaciones del iTeam), músicos, ingenieros de sonido y titulados del Máster de Postproducción Digital de la UPV. Todos ellos con una audición normal, con edades comprendidas entre los 24 y 35 años, de los cuales 15 eran hombres y 10 mujeres.

El primer paso antes de realizar el test era dar **información previa a los sujetos**, explicando cuál era su tarea en el test y cómo era el procedimiento que debían seguir. Para facilitar este proceso y para asegurar que la información dada a todos los sujetos era la misma, se hicieron una hojas de “explicación del test” (ver Anexo – Hojas de explicación del test) en las que además se resumía brevemente la mecánica a seguir y se daban algunas indicaciones prácticas. Estas hojas eran comentadas con cada participante para resolver pequeñas dudas del texto.

Inmediatamente después de recibir esta información los sujetos eran introducidos en la sala, previamente acondicionada y con todos los sistemas del test preparados para realizar la **fase de adiestramiento**. Se ha visto que este adiestramiento es muy recomendable para conseguir resultados más fiables (Rec UIT-R BS.1534-1, 2003) (Bech & Zacharov, 2006). El objetivo de este procedimiento es que el sujeto se familiarice con el entorno del test (en este caso la interfaz gráfica y el control mediante el joystick) así como con la tarea encomendada (evaluar la distancia a la que le parece que está un sonido, comparándolo con un sonido de referencia). Absolutamente todos los sujetos que realizaron el test mostraron una rápida comprensión y adaptación al manejo de la interfaz. Además, se les permitió escuchar los distintos tipos de señal (ruido, voz, guitarra, puerta) y se les indicó que podrían escuchar los límites de distancia, es decir, lo que era muy cerca y muy lejos. Dado que la escala presentada a los sujetos era completamente subjetiva, era necesario establecer los límites mediante la experiencia de esta fase de adiestramiento, así los participantes podían saber que todos los estímulos del test estarían comprendidos en un rango concreto. Estos límites se escuchaban dos veces para cada tipo de sonido, en una primera vuelta se empleaba VBAP y en la segunda WFS, aunque esta información obviamente era desconocida para los sujetos. Después se presentaban estímulos aleatorios exactamente de la misma manera que se presentarían después durante el test real. Se instaba a todos los participantes a

evaluar unos diez o quince estímulos y se les informaba de que la prueba real se desarrollaría de la misma forma.

Como ya se comentó, el test tuvo que ser dividido por motivos prácticos en dos partes, una primera para los estímulos a 0° y otra segunda para los estímulos a 90°.

Tras la fase de adiestramiento se comenzaba con la **realización de la primera parte del test (0°)** (ver figura 4.6-1). La mecánica de ejecución era muy sencilla: Se presentaban cada vez dos sonidos a evaluar más el sonido de referencia y se debía evaluar cada uno de los estímulos con respecto a la referencia. Los sonidos podían escucharse en el orden que se deseara todas la veces que se quisiera.

Para que la comparación de sonidos tuviera sentido, el tipo de sonido (ruido rosa, voz, guitarra, puerta) presentado en cada par más su referencia era siempre el mismo. Es decir, siempre se comparaba un estímulo de voz con la referencia de voz, un estímulo de guitarra con la referencia de guitarra, etc.

Tras finalizar la primera parte del test, a todos los sujetos se les daba un pequeño premio de agradecimiento por su colaboración, y se tomaban un **descanso** de unos 15 minutos. Después comenzaba la **realización de la segunda parte del test (90°)** (ver figura 4.6-2), cuyo funcionamiento y desarrollo era exactamente igual a la primera parte, salvo por el hecho de que escuchaban los estímulos proviniendo de su lado derecho (90° a la derecha en el plano horizontal).

El tiempo estimado de realización de cada parte del test era de unos 20 minutos (120 estímulos en cada parte). En la práctica la duración media de la primera parte fue de 27 minutos y de la segunda parte 22 minutos. La diferencia de tiempos en la realización de las dos partes se debe sin duda al aprendizaje en la mecánica del test.

Al finalizar la segunda parte se les daba a todos los voluntarios otro pequeño premio de agradecimiento y se procedía a **comentar la prueba** con los sujetos. Hablar sobre el test una vez finalizado éste, puede proporcionar información útil para interpretar los resultados que se obtengan con el análisis de los datos. También suele aportar información adicional no esperada que puede ser relevante, o información para detectar posibles errores (Bech & Zacharov, 2006). Lo primero que se les pedía era que dieran su opinión sobre la prueba y también que comentasen libremente la experiencia. Después se les preguntaban ciertas cuestiones concretas:



Figura 4.6-1 Realización de la primera parte del test (0°)



Figura 4.6-2 Realización de la segunda parte del test (90°)

- ¿Qué parte del test te ha parecido más difícil?, (o lo que es lo mismo, ¿qué les había parecido más difícil, la escucha a 0° o a 90°?)
- Señala de dónde te parecía que provenían los sonidos. ¿Te ha parecido que alguno de ellos viniese de otra dirección, lateral o elevada?
- ¿Qué tipo de sonido (ruido rosa, voz, guitarra, puerta) te ha parecido más difícil?, ¿y cuál más fácil?
- ¿Quieres comentar algo más?

indica a qué DISTANCIA te parece que está el sonido

SONIDO A (1)

MUY LEJOS

++ LEJOS

UN POCO + LEJOS

REFERENCIA (2)

UN POCO + CERCA

++CERCA

MUY CERCA

SONIDO B (3)

¿Qué seguridad tienes?  
1-NADA seguro 5-MUY seguro

1

OK

¿Qué seguridad tienes?  
1-NADA seguro 5-MUY seguro

5

OK

BORRAR (7) SIGUIENTE (8)

Figura 4.6-3 Captura de pantalla de la interfaz con las respuestas marcadas



## 4.7 – Resultados y análisis

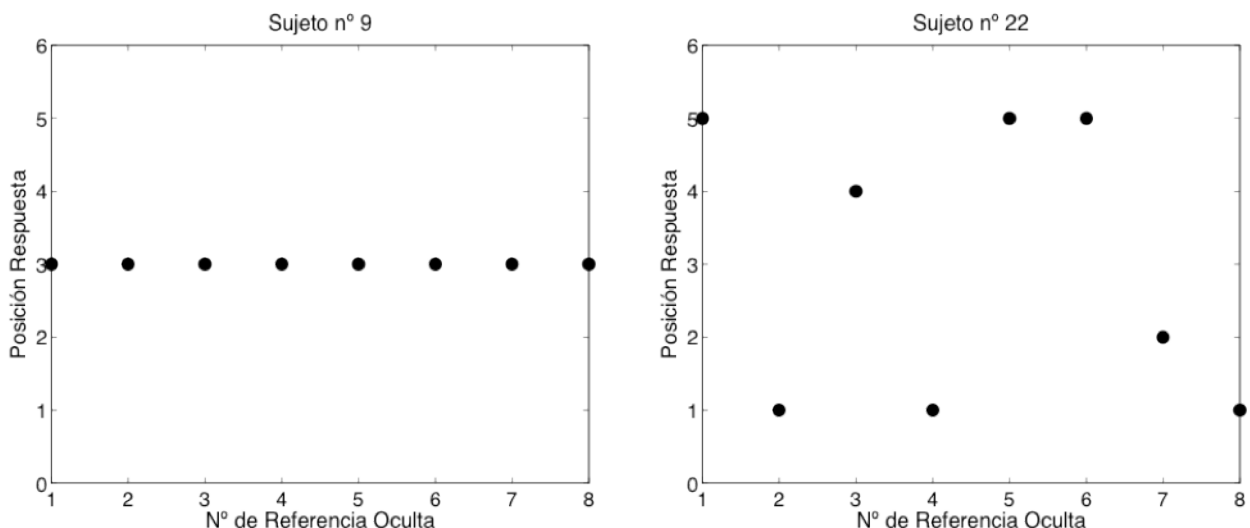
El tratamiento de los datos obtenidos con el test de percepción subjetiva se realizó mediante los programas Matlab y SPSS.

### Medida de la fiabilidad de los resultados y postselección de sujetos

Antes de realizar un análisis pormenorizado de los datos conviene estudiar la fiabilidad de los resultados obtenidos. Para ello se realizó un análisis del alfa de Cronbach de las respuestas entre los diferentes sujetos. Así se puede obtener una medida de la concordancia de las respuestas obtenidas, lo que ofrece una referencia sobre la fiabilidad de los datos.

La fiabilidad alfa de las respuestas de distancia es de  $\alpha=0,992$  (para  $N=25$  sujetos) lo que indica una muy buena fiabilidad. El valor alfa para las respuestas de seguridad es de  $\alpha=0,842$  (para  $N=25$  sujetos) que indica una buena fiabilidad. (Ver tablas B-1, B-2 anexo III)

El proceso de postselección de los sujetos se emplea para descartar a aquellos que tengan resultados muy incoherentes, para que sus valores no desvíen los análisis posteriores. Para ayudar en este proceso se introdujeron en el diseño del test varios estímulos que funcionaban como referencias ocultas. Tras analizar las respuestas de estas referencias ocultas se observó que prácticamente todos los sujetos eran capaces de localizar con mucha precisión (uno o dos fallos, como máximo) las referencias ocultas, excepto el sujeto número 22 que daba respuestas muy dispares (ver figura 4.7-1).



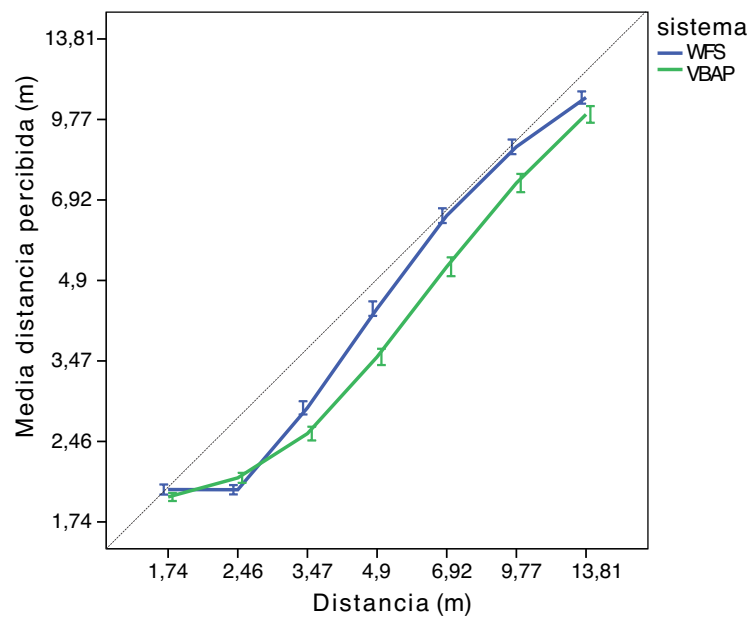
**Figura 4.7-1 Respuestas sobre referencias ocultas.**

**Izquierda sujeto 9, todas acertadas (caso común). Derecha sujeto 22, ningún acierto (caso único)**

Dados estos resultados tan llamativos para el sujeto 22 se comprobó el valor de correlación de todas sus respuestas con respecto al resto de sujetos, obteniéndose un valor de correlación de 0,822. Aunque este valor resulta un poco inferior a la media no es lo suficientemente bajo como para justificar su eliminación del grupo. De hecho el valor de alfa corregido si se elimina el sujeto 22 tan solo asciende a 0,993, por lo que no es útil prescindir de este sujeto (ver tabla B-1 en anexo III). Se hizo una inspección de sus resultados generales y se descubrió que el resto de sus respuestas eran coherentes. Por tanto finalmente no se prescindió de ningún sujeto en la postselección.

### Aciertos por sistema

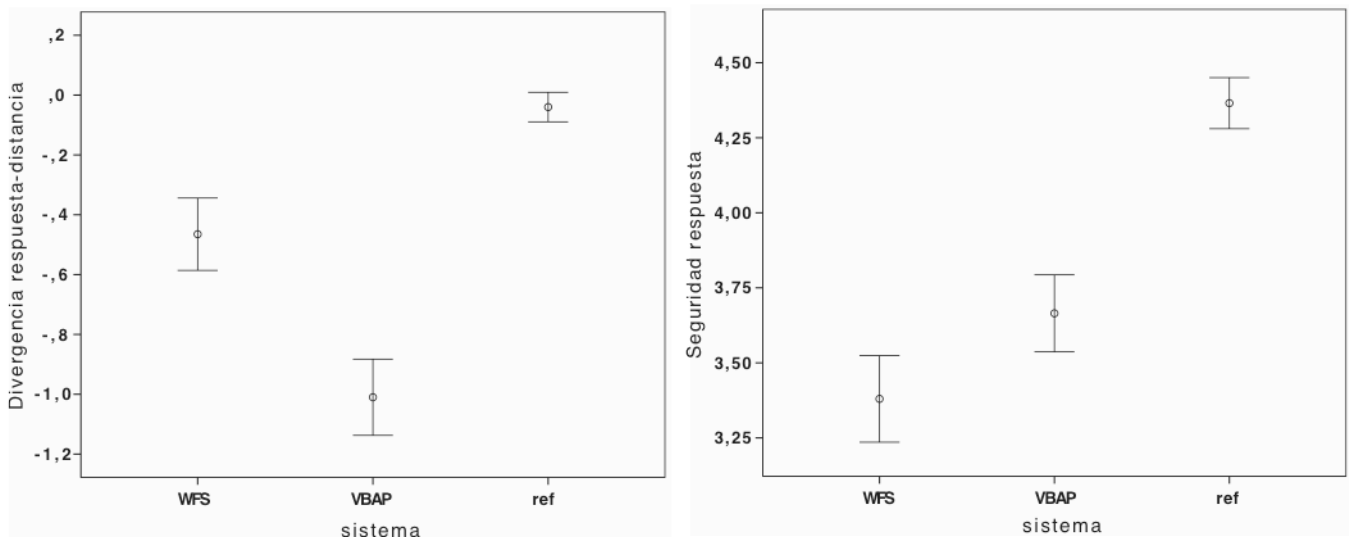
En la figura 4.7-2 se ha representado la media de las respuestas de distancia para cada sistema. Como se puede ver ambos sistemas consiguen resultados coherentes. Además se aprecia como los dos tienden a acercarse las distancias sintetizadas. WFS parece acercarse más al comportamiento ideal (diagonal de la gráfica).



**Figura 4.7-2 Media de las respuestas de distancia de cada sistema (WFS y VBAP). Intervalos de confianza (IC) del 95%**

Dado que ambos sistemas tienden a acercarse las distancias, al representar la divergencia entre la respuesta de distancia y la distancia sintetizada los valores obtenidos serán negativos. Cuanto más cercana a cero sea la divergencia mejor comportamiento tendrá el sistema. En la figura 4.7-3 (izquierda) podemos ver la divergencia de las respuestas de WFS y VBAP para la posición de referencia (posición 3, a 4,9m) y también para las referencias ocultas (ref). Se observa como las referencias ocultas tienen una divergencia prácticamente inexistente, indicando un comportamiento casi ideal, mientras que WFS y VBAP tienen una

mayor divergencia, apreciándose una diferencia entre los dos sistemas. En la figura 4.7-3 (derecha) se ha representado la media de las respuestas de seguridad de WFS, VBAP y las referencias ocultas para la posición de referencia (posición 3, a 4,9m). Se observa como ambos sistemas tienen una media de seguridad inferior a la de las referencias ocultas. El hecho de que las referencias ocultas tengan tan buen comportamiento es un indicador de buena fiabilidad, además permite comparar el rango de valores con respecto a los sistemas bajo estudio.



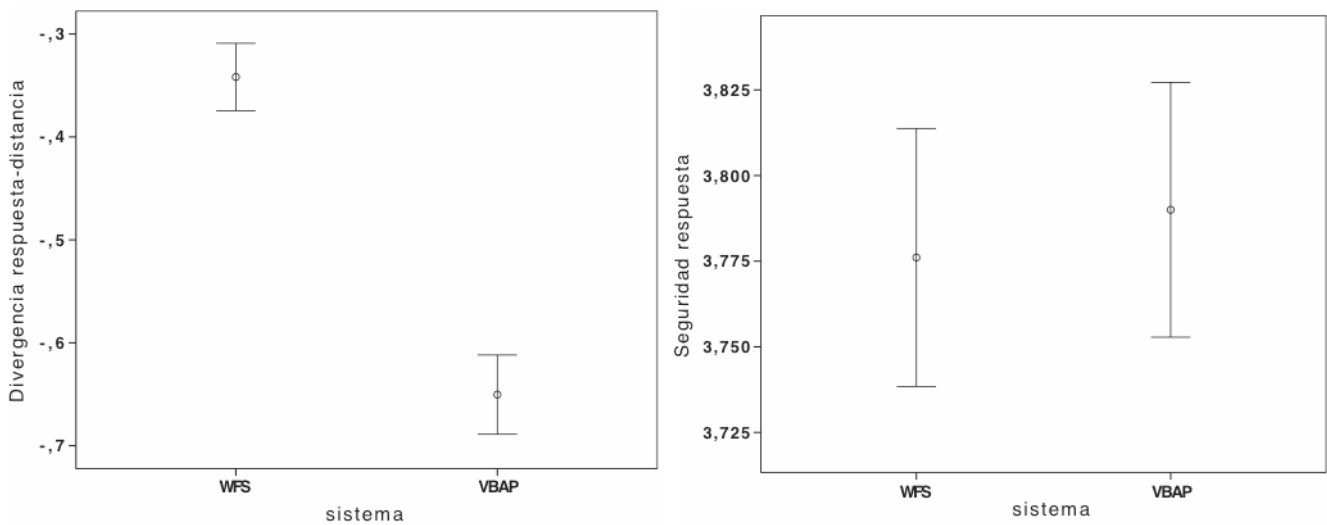
**Figura 4.7-3 Divergencia de las respuestas de distancia (izquierda) y media de las respuestas de seguridad (derecha) en la posición de referencia (posición 3, a 4,9m), para WFS, VBAP y referencias ocultas. IC 95%**

La comparación entre sistemas para todas las posiciones se puede ver en la figura 4.7-4, donde se ha representado la divergencia total en las respuestas de distancia de WFS y VBAP a la izquierda, y la media total de las respuestas de seguridad a la derecha. En estas gráficas se puede ver el comportamiento general de un sistema con respecto al otro, observándose una menor divergencia para WFS y por tanto un mejor comportamiento que VBAP. Sin embargo, la gráfica sobre respuestas de seguridad nos muestra que ambos sistemas ofrecen una seguridad media, sin que ninguno de ellos destaque sobre el otro. Aunque VBAP tiene una valoración levemente mejor que WFS, los intervalos de confianza en seguridad abarcan aproximadamente el mismo rango, indicando que no hay prácticamente diferencia de seguridad entre ambos sistemas.

Para estudiar en detalle la influencia del sistema de reproducción en el test, se realizó una prueba-t de muestras pareadas para WFS y VBAP. La prueba-t permite comparar medias de manera muy fina, para lo cual necesita de unas condiciones de experimentación

concretas en las que comparar medias que provengan de los mismos sujetos frente a los mismos estímulos.

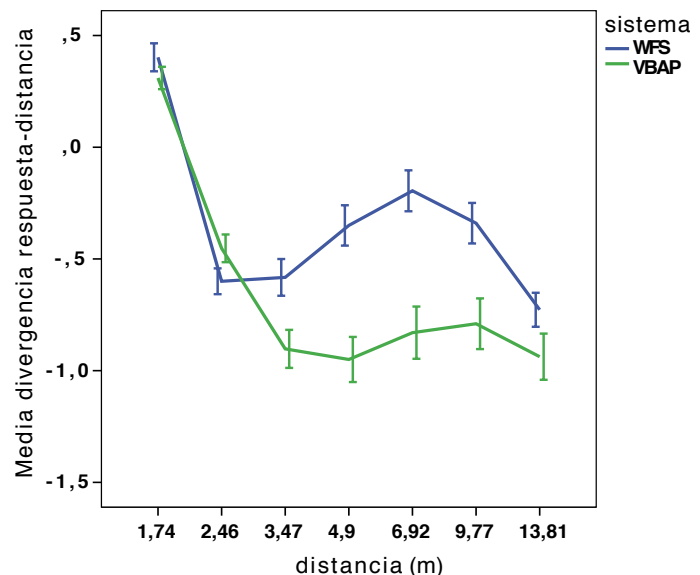
La prueba-T pareada muestra que el sistema utilizado (WFS vs VBAP) tiene una influencia altamente significativa en las respuestas de posición ( $t=16$ ,  $gl=2799$ ,  $p<0.001$ ), pero no es significativa para la seguridad ( $p=0.52$ ). Para estimar el tamaño del efecto, realizamos un análisis de varianza de medidas repetidas que muestra además que el porcentaje de variabilidad explicado por el sistema para las respuestas de posición es  $\eta^2=0,085$  (8,5%) y no es significativo para la seguridad. (Ver tablas B-3 y B-4 en anexo III)



**Figura 4.7-4 Divergencia de las respuestas de distancia totales (izquierda) y media total de las respuestas de seguridad (derecha), para WFS y VBAP. IC 95%**

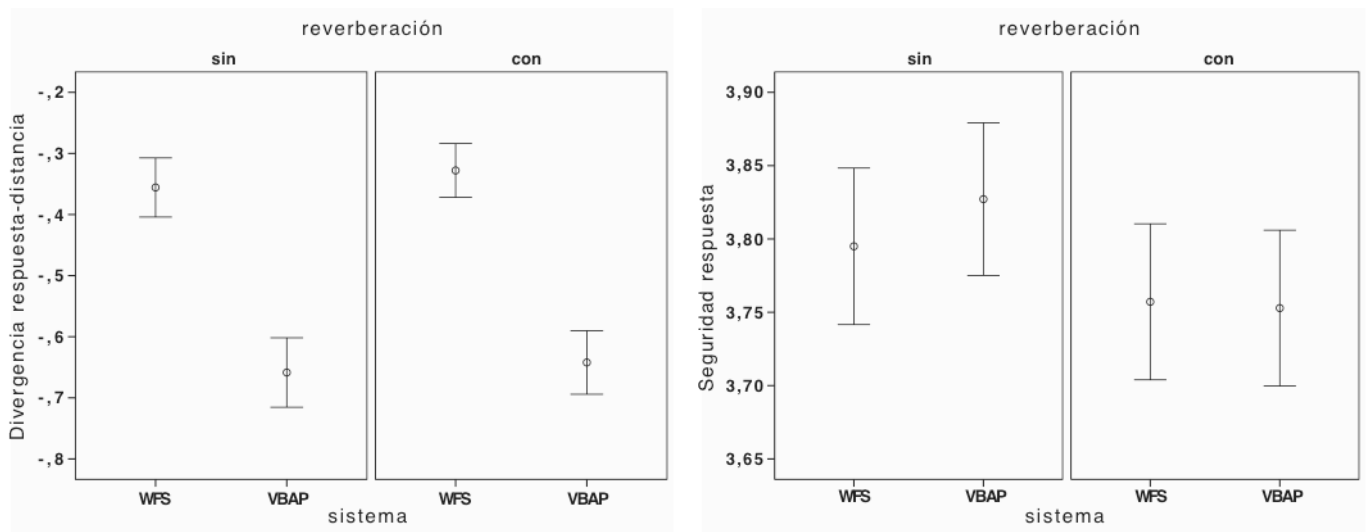
La divergencia en el comportamiento de los sistemas para cada una de las distancias está reflejada en la figura 4.7-5, donde se puede ver que aunque los dos sistemas tienen una distribución de la divergencia similar, WFS tiene un mejor comportamiento que VBAP, especialmente para las distancias situadas en torno a la referencia y por detrás de ésta.

**Figura 4.7-5 Divergencia de las respuestas para cada distancia con WFS y VBAP. IC 95%**



### Influencia de la reverberación

La reverberación añadida no introduce ningún efecto destacable, a la vista de la figura 4.7-6. En la gráfica de la izquierda se representa la divergencia en las respuestas de distancia para los dos sistemas sin y con reverberación, ofreciendo resultados prácticamente similares. En la gráfica de la derecha se puede ver la media de seguridad para ambos sistemas con y sin reverberación. Aquí se observa como la reverberación introduce un efecto muy pequeño de inseguridad.



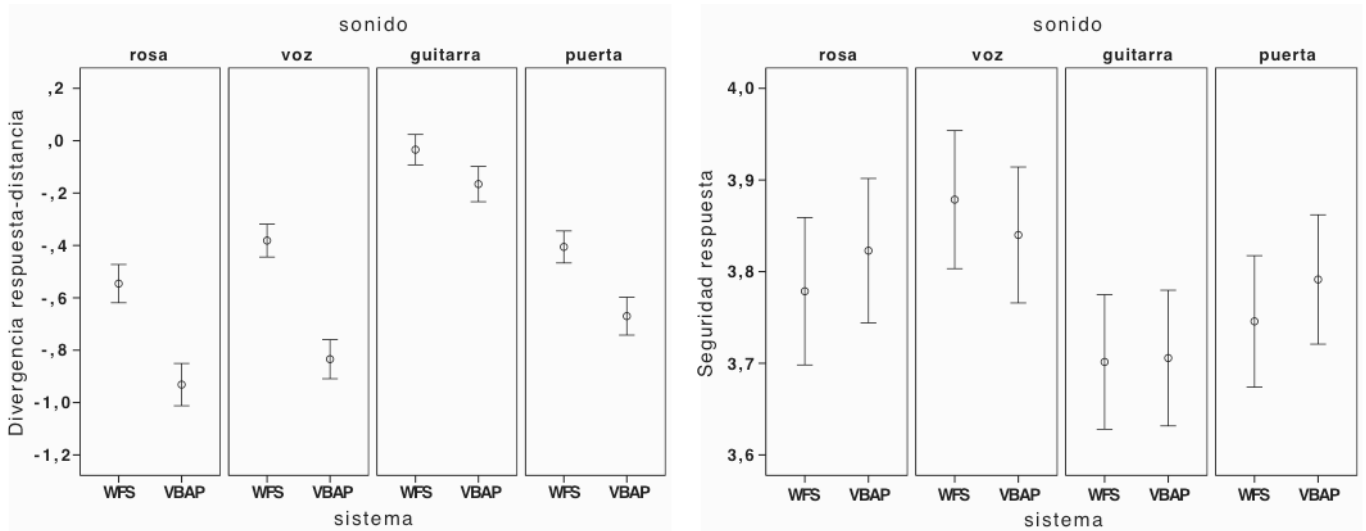
**Figura 4.7-6 Divergencia de respuestas de distancia (izquierda) y media de respuestas de seguridad (derecha) para WFS y VBAP, con y sin reverberación. IC 95%**

Para medir este efecto se realizó un análisis de varianza (ANOVA), obteniéndose que la influencia de la reverberación no es significativa ( $p=0,388$ ) sobre las respuestas de distancia. (Ver tabla B-5 en anexo III). Es decir, se confirma que la reverberación no tiene efectos relevantes en el experimento.

### Influencia del tipo de sonido

La representación de la divergencia de las respuestas de distancia para cada tipo de sonido ofrece resultados llamativos (ver figura 4.7-7 izquierda). En esta gráfica se aprecia que el sonido de guitarra es el que mejores resultados tiene, seguido del sonido de la puerta. La voz sería el siguiente tipo de sonido por acierto de resultados, seguido muy de cerca por el ruido rosa. Destaca la diferencia en la divergencia de distancia del ruido rosa y la voz según el sistema, teniendo claramente mejores resultados WFS.

Los resultados de seguridad por tipo de sonido se mueven de nuevo en márgenes similares, aunque curiosamente la guitarra tiene resultados ligeramente peores y la voz ligeramente mejores (ver figura 4.7-7 derecha).



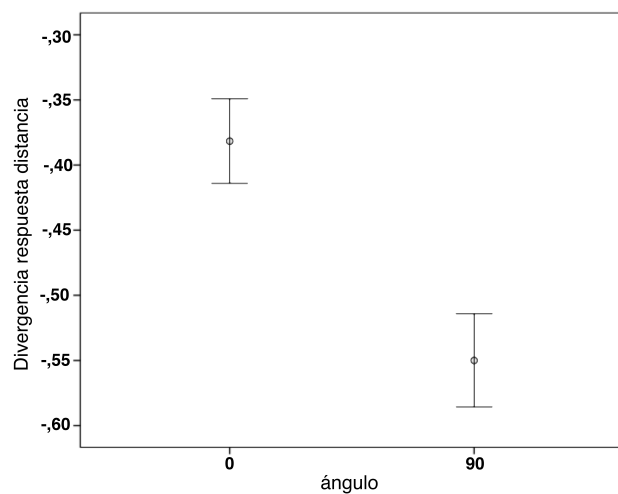
**Figura 4.7-7 Divergencia de respuestas de distancia (izquierda) y media de respuestas de seguridad (derecha) para WFS y VBAP, según el tipo de sonido. IC 95%**

Con un ANOVA de un factor se comprueba que el tipo de sonido tiene una influencia muy significativa ( $F=131.62$ ,  $gl=3$ ,  $p<0.001$ ) sobre los aciertos de distancia. También se muestra una relación cruzada entre el tipo de sonido y el sistema con significatividad muy alta ( $F=8.6$ ,  $gl=3$ ,  $p<0.001$ ), que corresponde a los mejores resultados de WFS en general y especialmente para ruido rosa y voz. (Ver tabla B-6 en anexo III)

#### Influencia del ángulo de escucha

El test se realizó a  $0^\circ$  y a  $90^\circ$  con el mismo número de estímulos en cada ángulo de escucha. En la figura 4.7-8 se puede ver cómo el número de aciertos a  $0^\circ$  es superior a los aciertos a  $90^\circ$ . Los resultados del ANOVA de un factor demuestran que el ángulo de escucha tiene una influencia muy significativa ( $F=56.08$ ,  $gl=1$ ,  $p<0.001$ ), pero que su relación cruzada con el sistema no es significativa ( $F=2.55$ ,  $gl=1$ ,  $p=0.110$ ). (Ver tabla B-6, anexo III)

**Figura 4.7-8 Divergencia en las respuestas de distancia para los ángulos de escucha  $0^\circ$  y  $90^\circ$ . IC 95%**

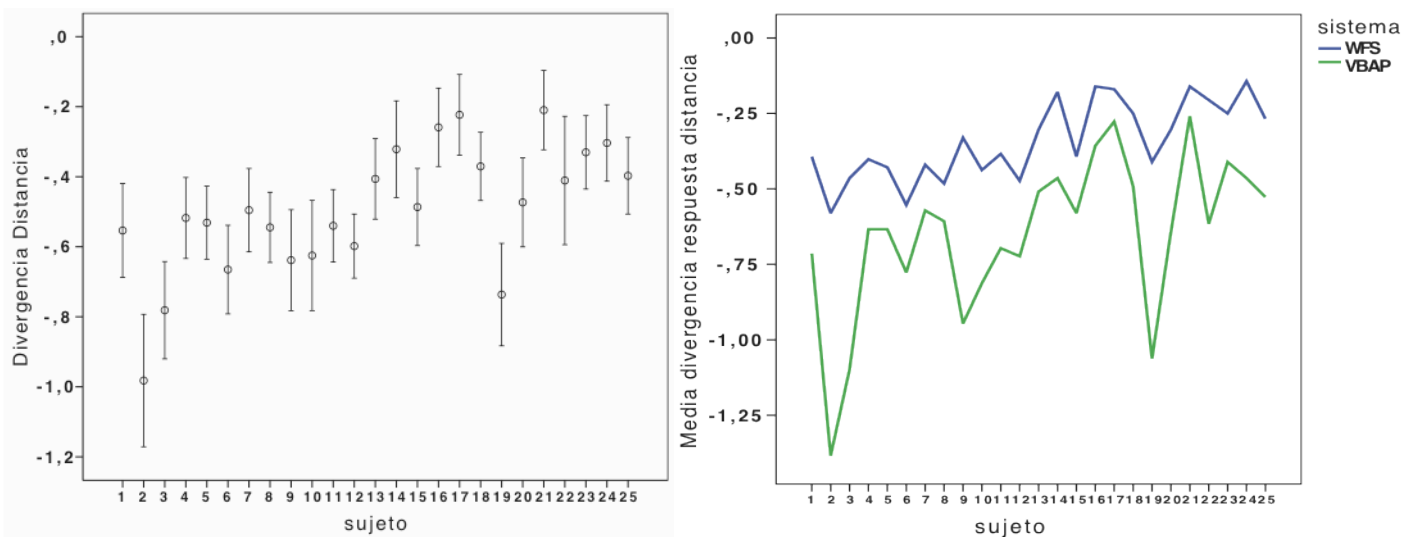


Es decir, aunque el ángulo de escucha es importante a la hora de acertar con la distancia correcta, no influye en los resultados para un sistema u otro, sin afectar a ninguno de manera apreciable.

### Influencia del sujeto

Para evaluar la influencia de cada sujeto en los aciertos de distancia se representó la divergencia media de las respuestas de cada uno de ellos (ver figura 4.7-9 izquierda). Como se puede ver hay algunos sujetos que destacan en su número de aciertos (sujetos 16, 17 y 21) y algún caso de malos resultados (sujetos 2 y 19).

Con un ANOVA de un factor se puede ver una relación cruzada muy significativa entre el sujeto y el sistema ( $F=2.33$ ,  $gl=24$ ,  $p<0.001$ ) (Ver tabla B-6 en anexo III). Al representar gráficamente la divergencia en las respuestas de distancia para cada sujeto, teniendo en cuenta WFS y VBAP por separado, destaca cómo algunos sujetos consiguen buenos resultados para WFS mientras que obtienen malos resultados para VBAP con mucha diferencia (ver figura 4.7-9 derecha). Llama la atención ver que algunos de estos sujetos son los mismos que tienen malos resultados en el caso general sin diferenciar sistema, lo que hace ver que esos peores resultados se deben al sistema VBAP.



**Figura 4.7-9 Divergencia de distancia media para cada sujeto (izquierda). IC 95% Divergencia de distancia media para cada sujeto teniendo en cuenta el sistema, WFS o VBAP (derecha)**

### Discusión

A la vista del análisis anterior se aprecia que los dos sistemas, WFS y VBAP, son capaces de reproducir distancia de manera bastante convincente. Dado que para la localización espacial VBAP emplea únicamente variaciones de nivel de presión sonora, se

confirma que el nivel de presión es el principal factor responsable de la detección de distancia con el sonido (Mershon & King, 1975) (Flanagan & Taylor, 1999).

Sin embargo, WFS obtiene mayor media de aciertos de distancia que VBAP, haciendo ver que hay factores adicionales en este sistema que mejoran la percepción de distancia. La principal característica de WFS es su capacidad para sintetizar frentes de onda, por tanto es de suponer que es justamente este frente de onda el que mejora la percepción de la distancia, probablemente debido a las diferencias interaurales que genera en el oyente.

Respecto al caso del sujeto 22 que no acertaba sobre las referencias ocultas, podemos suponer que esta persona era capaz de detectar las posibles coloraciones de los sonidos reproducidos por el altavoz de referencia, a pesar de la ecualización de los sistemas. Por algún motivo desconocido la percepción de estas coloraciones debía confundir a esta persona, respondiendo distancias dispersas, y sin embargo respondiendo de manera coherente sobre los sonidos reproducidos por los otros sistemas.

Las respuestas de seguridad no reflejan diferencias relevantes entre WFS y VBAP, y además los valores de sus medias rondan la puntuación de 3.7, un valor medio en la escala (1 poco seguro – 5 muy seguro). Esto hace ver que no hay una robustez aparente mayor por parte de ninguno de los sistemas. Tampoco se aprecia una robustez destacable por ninguna de las distancias. A la vista de estos resultados y de su grado de dispersión podemos pensar que pudo haber alguna confusión respecto al modo de responder sobre la cuestión “¿Qué seguridad tienes?”. En el pequeño cuestionario realizado tras finalizar el test, al comentar la cuestión de la seguridad, algunos sujetos manifestaron que habían entendido esta pregunta de manera muy personal, haciendo referencia a su seguridad como estado de ánimo o disposición, cuando lo que se pretendía que caracterizaran era la seguridad que habían obtenido en la escucha de cada sonido. Es decir, se quería obtener información relativa al sonido, y no tanto sobre su estado de ánimo personal. Este hecho justificaría que los valores estadísticos de este parámetro no sean muy relevantes. Quizá con una modificación sobre esta pregunta, o con una aclaración más profunda sobre la misma, se hubieran obtenido resultados más concluyentes. En cualquier caso los resultados nos dicen que ninguno de los sistemas muestra una robustez en seguridad mayor que el otro.

El uso de reverberación no introduce diferencias apreciables entre un sistema u otro, ni tampoco parece que la reverberación misma favorezca la localización en distancia de los sonidos. De hecho, las respuestas de seguridad sobre estos estímulos indican que la reverberación producía una ligera inseguridad a la hora de situarlos. En los comentarios posteriores a la realización del test hubo cinco personas que manifestaron que la



reverberación (o lo que se identificó por su descripción como la reverberación) les resultaba confusa. Como ya se comentó, la reverberación se generó a partir de la síntesis de fuentes secundarias atenuadas para las primeras reflexiones. Lo ideal hubiese sido poder sintetizar fuentes de tercer, cuarto orden e incluso muchas más, pero la potencia del ordenador empleado para la síntesis no permitió generar más que la primera reflexión. Quizá con la síntesis de muchas reflexiones se puedan obtener resultados más concluyentes.

Se ha visto que el ángulo de escucha tenía una gran influencia en la percepción de la distancia, pero no tenía relación aparente con el sistema empleado. A estos resultados habría que añadir la información obtenida en los comentarios posteriores al test. Diecisiete personas manifestaron que les pareció más difícil responder en el ángulo de escucha de 90°, a cuatro les pareció más difícil a 0° y a otros cuatro les pareció que ambos ángulos tenían la misma dificultad. Estas afirmaciones se corresponden con los resultados obtenidos, donde a 0° se obtiene mayor número de aciertos. Además de comentar sobre la dificultad de cada ángulo de escucha, a los sujetos también se les preguntó por la dirección aparente de la fuente de los estímulos. Cinco personas afirmaron haber percibido ciertas lateralizaciones a izquierda o derecha con algunos estímulos y uno de ellos afirmó escuchar algunos sonidos más elevados. Estas personas afirmaban haber escuchado estas lateralizaciones sobre todo en el ángulo de escucha de 90°. Probablemente estas personas estaban detectando la diferencia de altura entre los *arrays* de altavoces y el altavoz de referencia, pero lo interpretaban por comparación como pequeños desplazamientos a izquierda o derecha. Tan solo una persona estaba detectando correctamente la diferencia de altura y sólo algunas veces, esto confirma la suposición inicial de que la escucha humana en altura es muy deficiente, especialmente en el plano sagital medio (escucha a 0°).

Tras realizar el test, a los sujetos se les preguntó por el tipo de sonido que les había parecido más fácil o más difícil de situar. Los voluntarios ordenaron los tipos de estímulo según dificultad creciente, en general siguiendo dos patrones: el primero ruido rosa, puerta, guitarra y voz, y el segundo guitarra, puerta, voz y ruido. El segundo patrón se corresponde directamente con el orden por número de aciertos de los resultados.

El tipo de sonido influye mucho en la localización espacial de distancia teniendo mejores resultados aquellos sonidos con características impulsivas como la guitarra o la puerta. Este hecho es conocido en la escucha espacial, dadas las características espectrales planas de este tipo de sonidos y sus cortos transitorios (Blauert, 1996). Lo que llama la atención son los mejores resultados que obtienen el resto de sonidos para WFS. El ruido rosa y la voz tienen peores resultados que los sonidos impulsivos, pero en WFS obtienen un aumento de

aciertos destacable con respecto a VBAP. La media de aciertos de la voz llega incluso a igualar la de la puerta en WFS. Resulta muy relevante comprobar que para sonidos más difíciles de localizar en distancia, el sistema WFS consigue mejorar los resultados.

También se ha podido comprobar al analizar la relación de la influencia del sujeto con el sistema, que todas las personas que realizaron el test obtuvieron mayor número de aciertos con el sistema WFS. Es destacable además cómo algunos de ellos tienen una diferencia muy grande en el número de aciertos entre cada sistema, obteniendo malos resultados para VBAP pero bastante buenos para WFS. Este hecho hace pensar nuevamente en que, en general, la síntesis del frente de ondas mejora la percepción de la distancia, siendo algunos sujetos muy sensibles a este hecho.

#### 4.8 – Posibles investigaciones futuras

Con la experiencia de haber realizado este test de percepción subjetiva y a la vista de los resultados obtenidos, se pueden plantear varias líneas de trabajo para continuar investigando la percepción de la distancia con el sonido y los sistemas WFS y VBAP.

Sería interesante plantear otras pruebas de percepción que incluyeran la síntesis de reverberación mediante el método de las imágenes de una manera más perfecta. Añadiendo un gran número de reflexiones se podría estudiar la influencia del recinto en la percepción de la distancia, así como la capacidad de los sistemas WFS y VBAP para generar estas reflexiones. También se podría evaluar si hay un número de reflexiones necesarias para que la percepción de distancia mejore y qué sistema es capaz de reproducirlas más eficientemente. Además se podrían incluir como variable de estudio la proporción de sonido directo y sonido reverberante, para buscar cuál es la relación más óptima con cada sistema.

Otro estudio podría incluir mayor número de ángulos y posiciones de escucha, para intentar establecer mejor la influencia del sistema de sonido sobre este parámetro.

Una variable muy interesante de estudiar en la percepción de la distancia sería el movimiento del sonido. En concreto los sistemas de sonido WFS y VBAP ofrecen muchas posibilidades en este sentido. Este experimento podría revelar muchas relaciones cruzadas entre la velocidad del sonido, el ángulo que abarque el movimiento, la dirección, el tiempo de duración del sonido, etc. de posible interés en las síntesis de sonido espacial.

En general, la percepción de la distancia sonora es un campo con muchas posibilidades de estudio en la síntesis de sonido espacial y, en concreto, los sistemas de sonido WFS y VBAP aún tienen muchos abiertos a la experimentación y el desarrollo.

## **5 – Conclusiones**

En base a los objetivos planteados con el diseño del test de percepción subjetiva, podemos enumerar las siguientes conclusiones del experimento:

- Ambos sistemas de sonido espacial (WFS y VBAP) se han demostrado capaces de simular una cierta sensación sonora de distancia, que creemos mayor a la simplemente basada en la disminución de nivel de la señal.
- El tipo de sonido es determinante en la percepción de la distancia, obteniendo mejores resultados los sonidos impulsivos. Además WFS es capaz de reproducir mejor que VBAP la distancia sonora para otro tipo de sonidos que no sean impulsivos.
- El ángulo de escucha influye mucho en la percepción de la distancia sonora, pero no es concluyente su relación con los sistemas WFS o VBAP.
- La introducción de reflexiones de primer orden no ha mostrado una mejora apreciable en la percepción de la distancia. Queda pendiente verificar hasta que punto la introducción de múltiples órdenes de reflexión mejoraría esta sensación.
- De los resultados de este test se puede concluir que WFS ha mostrado una capacidad global mejor que VBAP para reproducir distancia sonora, al menos para fuentes situadas delante del oyente.

Estos resultados nos llevan a pensar que el nivel de presión sonora es el principal factor responsable de la percepción de distancia. Además los resultados parecen mostrar que la capacidad del sistema de sonido WFS de sintetizar frentes de onda interviene especialmente en la localización de fuentes sonoras en distancia.

La perspectiva tecnológica que ofrece haber realizado este trabajo final de grado permite plantearse cuestiones sobre el futuro del sonido espacial. El creciente interés de la industria audiovisual y de los medios sobre estas tecnologías augura muchos años de desarrollo y trabajo continuo en este tipo de sistemas de sonido. En general, la evolución lógica de la mezcla y escucha de audio se orienta hacia un aumento de calidad en la generación y reproducción del sonido, buscando cada vez mayor fidelidad y realismo. Esto se ha ido consiguiendo incrementando la calidad y también la espacialidad de los sistemas de sonido.

Existen nuevas aproximaciones a los procesos de mezcla de sonido que tienen en cuenta sistemas de reproducción escalables, en donde la mezcla sonora es capaz de

adaptarse por sí misma al sistema donde está siendo reproducida (SAOC, 2012). Esta característica de adaptabilidad de la mezcla sonora al sistema de reproducción, o incluso al lugar de reproducción, será especialmente útil para los sistemas de sonido espacial, permitiendo así que una misma mezcla sonora mantenga su coherencia espacial en ambientes, salas o lugares muy diferentes. El papel que pueden jugar los sistemas WFS o VBAP en este tipo de tecnologías puede ser muy importante y es uno de sus mayores puntos de interés.

Por otro lado, el mercado audiovisual está cada vez más interesado en ofrecer medios inmersivos en donde el contenido envuelva literalmente al espectador. El auge del cine 3D en los últimos años es un ejemplo vivo de esa tendencia, pero no es el único. El mercado doméstico ya dispone de televisores o consolas de videojuegos 3D. La industria de contenidos audiovisuales está ansiosa por ofrecer experiencias más inmersivas. En este contexto, un mayor entendimiento y desarrollo de los sistemas de sonido espacial en general, y en concreto WFS o VBAP, puede ser algo muy valioso, dado que los sistemas audiovisuales inmersivos están buscando una espacialización sonora más desarrollada y precisa.

## 6 - Bibliografía

- Bech, S; & Zacharov, N. (2006). *Perceptual Audio Evaluation-Theory, Method and Application*. Sussex, Inglaterra: John Wiley & Sons Ltd.
- Begault, D. R. (2000). *3-D Sound for Virtual Reality and Multimedia*. Boston, Estados Unidos: Academic Press.
- Berkhout, A. (1988). A holographic approach to acoustic control. *Journal of the Audio Engineering Society* , 36 (12), 977-995.
- Berkhout, A. (1987). *Applied Seismic Wave Theory*. Elsevier Science.
- Berkhout, A. V; & de Vries, D. (1993). Acoustic control by wave field synthesis. *Journal of the Acoustical Society of America (JASA)* , 93 (5), 2764-2778.
- Berkhout, A; Vogel, P; & de Vries, D. (1992). Use of wave field synthesis for natural reinforcement sound. *92th AES convention*. Viena: AES.
- Blauert, J. (2005). *Communication acoustics*. Berlin: Springer-Verlag.
- Blauert, J. (1996). *Spatial Hearing: The psychophysics of human sound localization*. Cambridge: MIT Press.
- Bleda, S. (2009). *Contribuciones a la implementación de sistemas Wavefield Synthesis*. Valencia: Universidad Politécnica de Valencia.
- Butler, R; & Belenduij, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. *Journal of the Acoustic Society of America (JASA)* , 61, 1264-1269.
- Davis, R; & Stephens, S. (1974). The effect of intensity on localization of sound sources on the median plane. *Journal Sound and Vibration* , 35, 223-229.
- EBU Tech 3276-2nd Edition (1998) Listening conditions for the assessment of sound programme material: monophonic and two-channel stereophonic. Ginebra: European Broadcasting Union
- EBU Tech 3276-E (2004) Listening conditions for the assessment of sound programme material. Supplement 1: Multichannel sound. Ginebra: European Broadcasting Union

- Ehrenstein, W; & Ehrenstein, A. (1999). Psychophysical Methods. In U. Windhorst, & H. Johansson, *Modern Techniques in Neuroscience Research* (pp. 1211-1241). Berlin: Springer-Verlag Berlin Heidelberg.
- Flanagan, S; & Taylor, V. (1999). Investigation into the Relationship between Subjective Loudness and Auditory Distance Perception. *107th AES convention*. Nueva York: AES.
- Gescheider, G. (1997). *Psychophysics: The fundamentals*. Mahwah, Nueva Jersey: Lawrence Erlbaum Associates Inc.
- Gutiérrez, P. (2008). *Mezcla y escucha de sonido multicanal*. Madrid: Universidad Politécnica de Madrid (UPM).
- Holbrook, K. (1988) Multichannel and Surround Sound Reinforcement past, present and future. *AES 6th international conference*. Arlington, Estados Unidos.
- Holman, T. (1996). The number of audio channels. *AES 100th convention*.
- Huygens, C. (1678). *Traité de la Lumière*.
- IBM SPSS Statistic. (Sitio web) [Consultado en Enero de 2013] Disponible en: <http://www-01.ibm.com/software/es/analytics/spss/products/statistics/>
- IEC. 60268-13. (1998). Sound System Equipment - Part 13: Listening Tests on Loudspeakers. International Electrotechnical Commission.
- Ingard, U. (1953). A review of the influence of meteorological conditions on sound propagation. *Journal of the Acoustic Society of America (JASA)* , 25, 405-411.
- ISO 9613-1:1993(E). (1993). Atenuación del sonido en propagación al aire libre. Cálculo de la absorción del sonido por la atmósfera. Organización Internacional de Normalización.
- Jackson, B. (2010) Beyond 5.1 *Mix magazine*, Septiembre
- Juan, L; & Yonghong, Y. (2011). Distance perception synthesis in 3D audio rendering using loudspeaker array. *Journal of Institute of Electrical and Electronics Engineers* , 290-293.

- Kerber, S; Wittek, H., Fastl, H., & Theile, G. (2004). Experimental investigations into the distance perception of nearby sound sources: Real vs WFS virtual nearby sources. *CFA/DAGA'04*. Estrasburgo.
- Llinares, J; Llopis, A., & Sancho, J. (1996). *Acústica arquitectónica y urbanística*. Valencia: Universidad Politécnica de Valencia.
- Mariño, R. (2011). Técnicas de sonido binaural en la postproducción audiovisual. Valencia: Universidad Politécnica de Valencia.
- Matlab (Sitio web). [Consultado en Diciembre de 2012] Disponible en: <http://www.mathworks.es/products/matlab/>
- Mershon, D. H; & King, E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics* , 18 (6), 409-415.
- Middlebrooks, J; & Green, D. (1989). Directional dependence of interaural envelope delays. *Journal of the Acoustical Society of America (JASA)* , 87, 2149-2162.
- Nagel, D. (2009, Octubre). *Iosono, Todd-AO Launch First 3-D Sound Mixing Facility*. [Consultado el 1 de Febrero de 2013] Disponible en: <http://mediadesigner.digitalmedianet.com/articles/viewarticle.jsp?id=881390&afterinter=true>
- Nakahara, M. (2005). *Multichannel Monitoring Tutorial Booklet 2nd edition*. Yamaha - Sona. SONA Corporation.
- Negrao, M. (2011, Marzo 24). *Presentation of WFSCollider*. [Consultado el 20 de Enero de 2013] Disponible en: [http://prezi.com/wxnd\\_cohbnum/steim-gol-wfs-and-wfscollider-overview/](http://prezi.com/wxnd_cohbnum/steim-gol-wfs-and-wfscollider-overview/)
- Neher, T; Rumsey, F., & Brookes, T. (2002). Training of listeners for the evaluation of spatial sound reproduction. *AES 112th convention*. Munich.
- Nielsen, S. (1991). Depth perception - finding a design goal for sound reproduction systems. *AES 90th convention*. Paris.
- Nunnally, J; & Bernstein, I. (1994). *Psychometric Theory*. Nueva York: McGraw-Hill.
- Packer, R; Jordan, K. (2002). *Multimedia: From Wagner to Virtual Reality*. Norton.

- Pulkki, V. (1997). Virtual sound source positioning using vector base amplitude panning. *J. Audio Engineering Society*, vol.45, nº6, pp. 456-466.
- Pulkki, V. (2001a). Spatial sound generation and perception by amplitude panning techniques. Helsinki, Finland: Helsinki University of Technology.
- Pulkki, V. (2001b). Localization of amplitude-panned virtual sources II: Two- and three-dimensional panning. *Journal of the Audio Engineering Society* , 49 (9), 753-767.
- Pulkki, V; & Karjalainen, M. (2001). Localization of amplitude-panned virtual sources I: Stereophonic panning. *Journal of the Audio Engineering Society* , 49 (9), 739-752.
- Rec UIT-R BS. 1284. (2003). Métodos generales para la evaluación subjetiva de la calidad de sonido. Ginebra: Union Internacional de Telecomunicaciones.
- Rec UIT-R BS.1116-1 (1997) Métodos para la evaluación subjetiva de pequeñas degradaciones en los sistemas de audio incluyendo los sistemas de sonido multicanal. Ginebra: Union Internacional de Telecomunicaciones
- Rec UIT-R BS.1534-1. (2003). Método para la evaluación subjetiva del nivel de calidad intermedia de los sistemas de codificación. Ginebra: Unión Internacional de Telecomunicaciones.
- Rec UIT-R BS.775-2. (2006). Sistema de sonido estereofónico multicanal con y sin acompañamiento de imagen. Ginebra: Union Internacional de Telecomunicaciones.
- Rumsey, F; Griesinger, D; Holman, T., Sawaguchi, M., Steinke, G., Theile, G., et al. (2001). *AESTD1001 - Multichannel surround sound systems and operations*. Audio Engineering Society .
- Rumsey, F. (2002). Spatial Quality Evaluation for reproduced Sound: Terminology, Meaning and a Scened-based paradigm. *Journal of AES*, 50 (9), 651-666
- SAOC. (2012). *White Paper. MPEG Spatial Audio Object Coding*. Fraunhofer Institute for Integrated Clrcuits. Erlangen: Fraunhofer IIS.
- Slater, J. (2008). The Dolby solution to Digital 3D. *Cinema Technology*, Marzo.
- Snow, W. (1953). Basic principles of stereophonic sound. *Journal of Motion Oicture and Television Engineers* , 61 (11), 922-940.



- Spors, S; Rabenstein, R; & Ahrens, J. (2008). The theory of wave field synthesis revisited. *124th AES convention*. Amsterdam: AES.
- SuperCollider.net*. [Consultado en Noviembre 2012] Disponible en: <http://supercollider.sourceforge.net/>
- The Center for New Music and Audio Technology (CNMAT), UC Berkeley. [Consultado el 20 de Enero de 2013] Disponible en: <http://opensoundcontrol.org/>
- Theile, G. (1990). On the performance of two-channel and multi-channel stereophony. *88th AES convention*. Montreux: AES.
- Theile, G; & Wittek, H. (2004). Wavv Field Synthesis: A promising spatial audio rendering concept. *Acoustical Science and Technology* , 25 (6), 393-399.
- Verheijen, E. (1997). Sound reproduction by Wave Field Synthesis. Delft, Países Bajos: Delft University of Technology.
- von Békésy, G. (1960). *Experiments in hearing*. Nueva York: McGraw-Hill.
- Wightman, F; & Kistler, D. (1992). The dominant role of low-frequency interaural time differences in sound localization. *Journal of the Acoustical Society of America (JASA)*, 19 (3), 1648-1661.
- Wittek, H. (2007). Perceptual differences between Wave Field Synthesis and stereophony. Surrey, Reino Unido: University of Surrey.
- Wittek, H; Kerber, S; Rumsey, F; & Theile, G. (2004). Spatial perception in WFS rendered sound fields: Distance of real and virtual nearby sources. *AES 116th convention*. Berlin.
- Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *Journal of the Acoustical Society of America (JASA)* , 111 (4), 1832-1846.

## Anexo I - Figuras adicionales (medios empleados)



Figura A-1 Amplificadores YAMAHA RX-V650 (14 unidades)

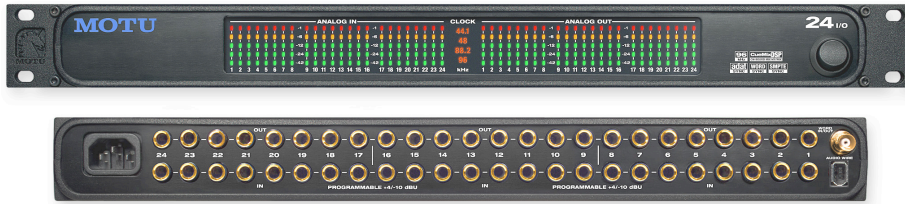


Figura A-2 Tarjetas de sonido MOTU 24IO (4 unidades)

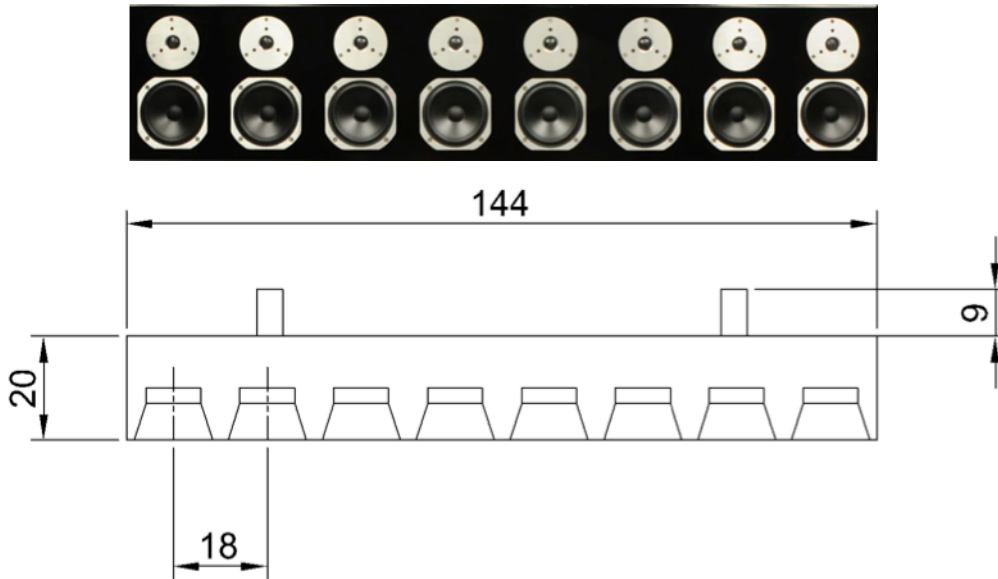


Figura A-3 Arrays de 8 altavoces (8 unidades), con conos Beyma T2010 y 5MP60/N

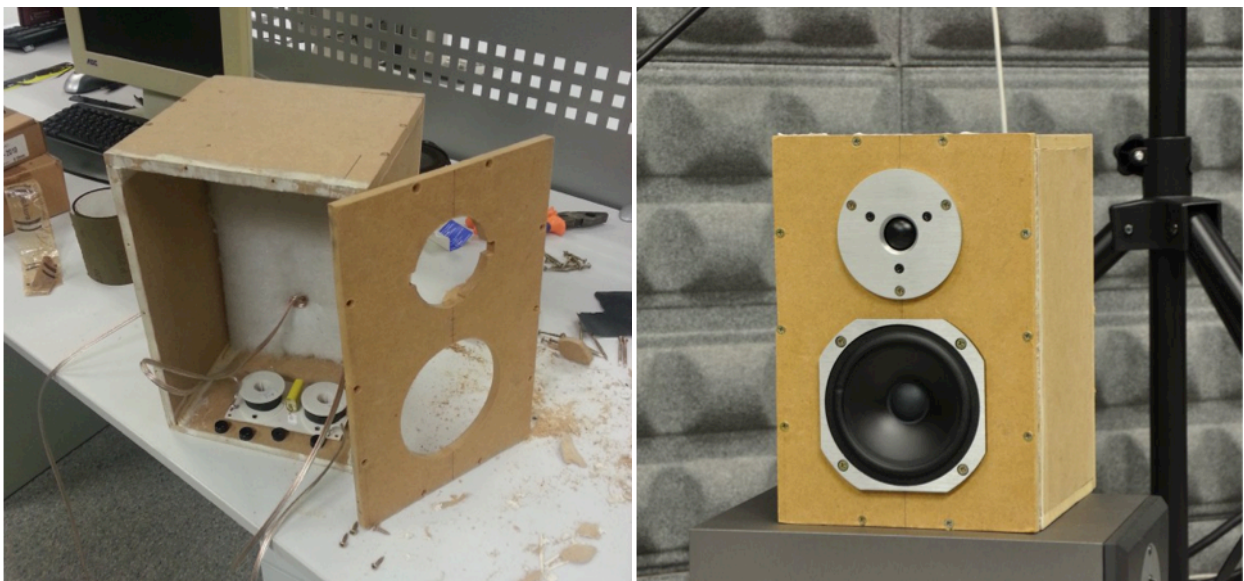


Figura A-4 Altavoz de referencia, con conos Beyma T2010 y 5MP60/N

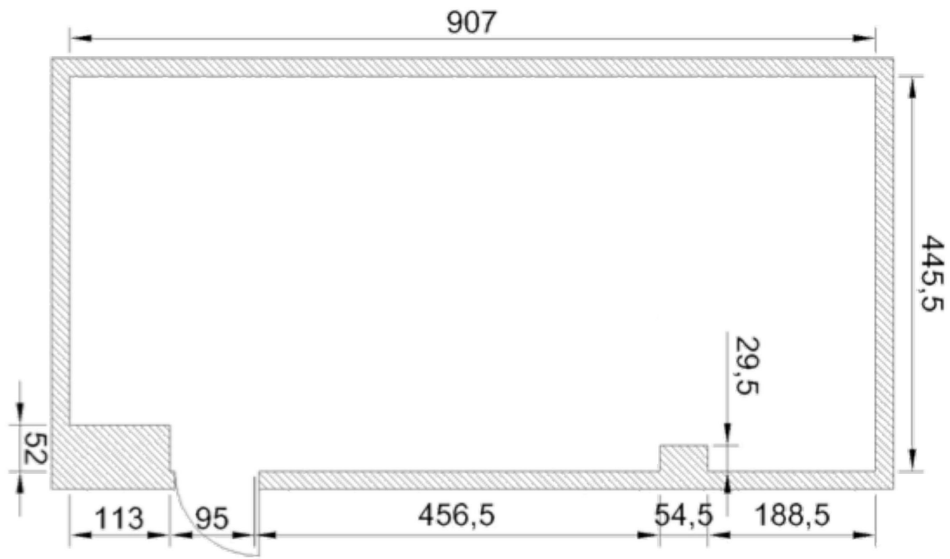


Figura A-5 Planta de la sala



Figura A-6 Micrófono omnidireccional Earthworks M30



Figura A-7 Tarjeta de sonido MOTU Traveler



Figura A-8 Sonómetro Brüel & Kjaer 2250

## Anexo II - Hojas de explicación del test

### EXPLICACIÓN DEL TEST-ENTRENAMIENTO DE LOS VOLUNTARIOS

#### OBJETIVO DEL TEST

Evaluar a qué **distancia** te parece que están los sonidos, comparándolos con un sonido de referencia.

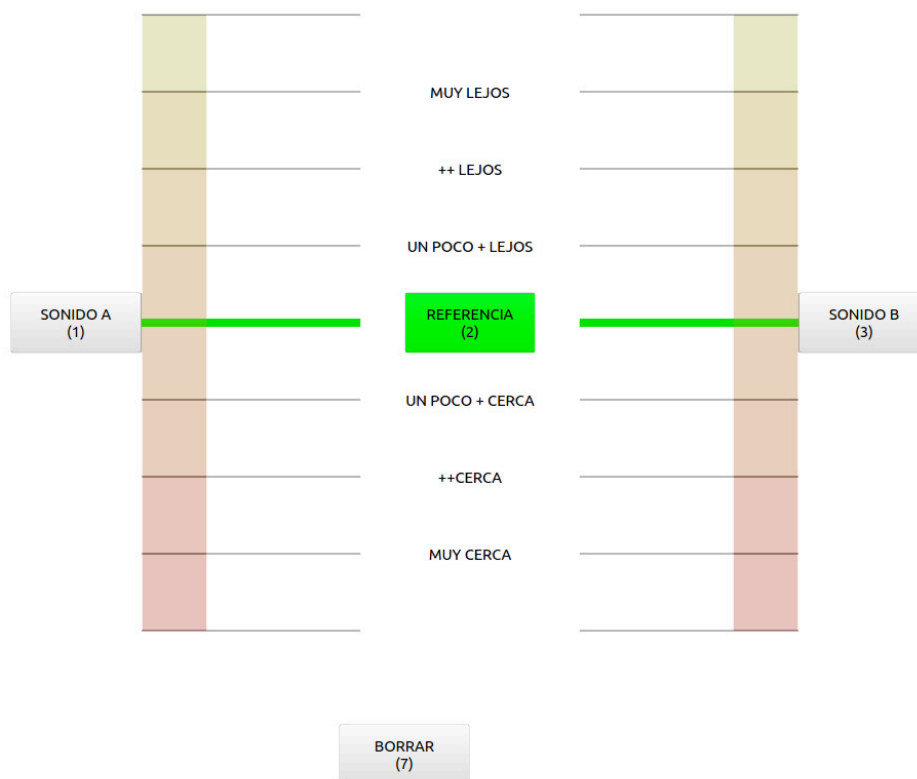
#### PROCEDIMIENTO

Tendrás que responder a dos cuestiones para cada sonido:

- Indica a qué distancia te parece que está el sonido (cuánto de lejos o de cerca)
- ¿Qué seguridad tienes? (1 nada – 5 mucho)

Usarás un mando de videojuegos, y una pantalla te guiará en el proceso.

indica a qué DISTANCIA te parece que está el sonido



Se presentan dos sonidos (A y B) cada vez. Habrá que determinar la DISTANCIA CON RESPECTO AL SONIDO DE REFERENCIA para CADA UNO de ellos.

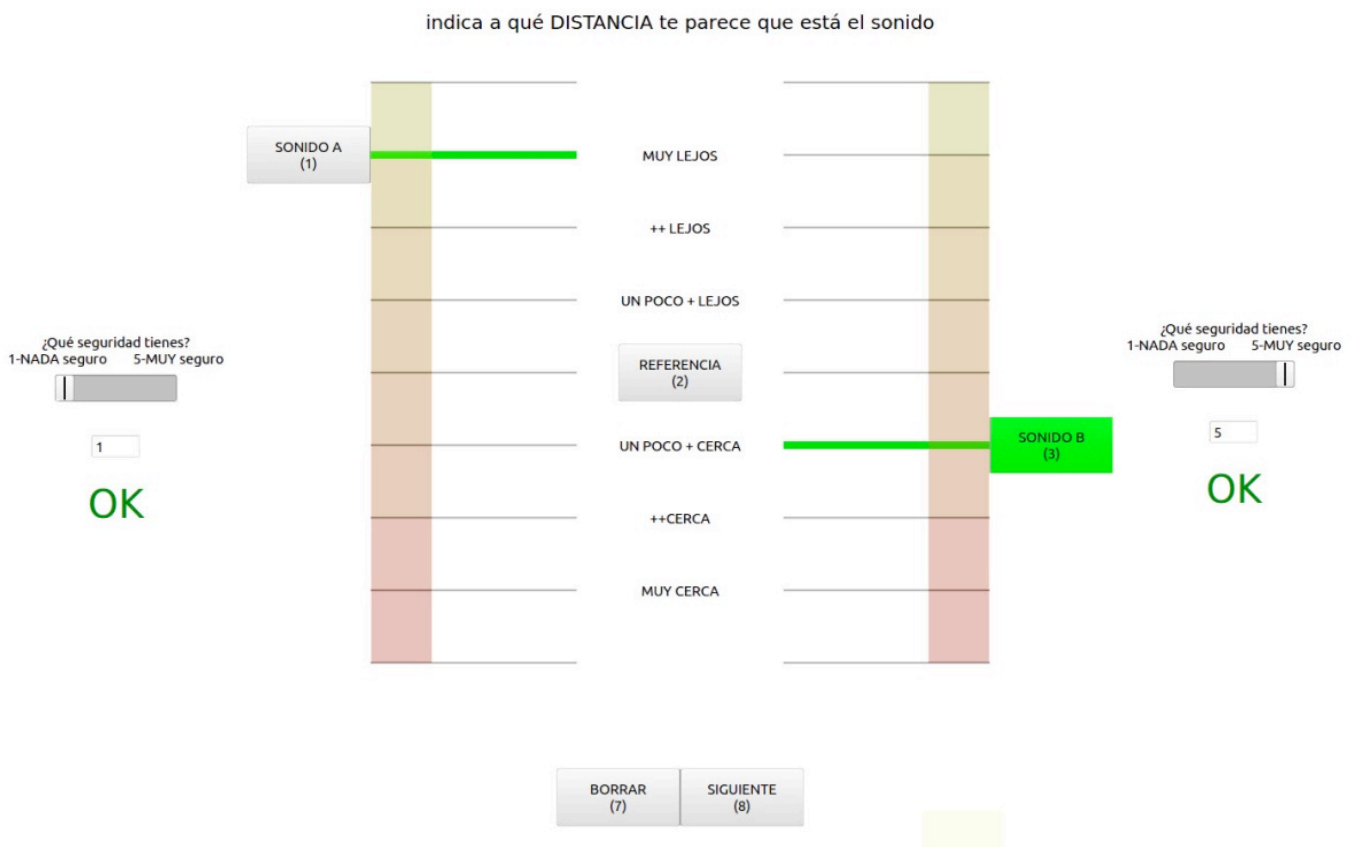
Escucharemos un sonido (A ó B) pulsando el botón correspondiente del mando. Después moviendo el cursor arriba o abajo seleccionaremos dónde nos parece que se sitúa el sonido con respecto a la referencia (el sonido de referencia se puede escuchar en cualquier momento con solo pulsar su botón).

Al confirmar la respuesta aparecerá la siguiente pregunta: “¿Qué seguridad tienes?” que se deberá evaluar de 1 (nada) a 5 (mucho).

Una vez respondidas las dos cuestiones para cada sonido aparecerá un nuevo botón en la pantalla (SIGUIENTE), que se activará con el correspondiente botón del mando. Al pulsarlo se nos presentará un nuevo par de sonidos para evaluar.



En la siguiente imagen se puede ver una pantalla de ejemplo completada. Se ha determinado que el sonido A parece estar “muy lejos” del de referencia, con una seguridad de 1 (nada). Mientras que el sonido B parece estar “un poco más cerca” que la referencia con una seguridad de 5 (muy seguro).



DOS TESTS SIMILARES: Uno de frente (0°) y otro girado 90°

TIEMPO ESTIMADO: 20 minutos, descanso, 20 minutos

RELÁJATE Y DISFRUTA, RECUERDA:

Esto no es un concurso. No te preocupes si no estás del todo seguro, eso también se evalúa.

No hay respuesta buena o mala. Lo que se quiere estudiar no es tu capacidad de escucha, sino la capacidad de reproducción de ciertos sistemas de sonido.

RECOMENDACIONES

No muevas, gires o bajes la cabeza durante la escucha.

Escucha los sonidos las veces que quieras, pero no demasiadas. Se alargaría mucho la prueba.

### Anexo III - Tablas de análisis estadísticos

**Estadísticos de fiabilidad**

Alfa de Cronbach	N de elementos
,992	25

**Tabla B-1 Análisis de fiabilidad de las respuestas de distancia**

**Estadísticos total-elemento**

	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Alfa de Cronbach si se elimina el elemento
RespDistancia1	60,88	1807,800	,914	,992
RespDistancia2	61,28	1849,665	,790	,993
RespDistancia3	61,08	1827,483	,890	,992
RespDistancia4	60,85	1789,174	,948	,992
RespDistancia5	60,85	1807,015	,934	,992
RespDistancia6	60,97	1819,083	,901	,992
RespDistancia7	60,83	1830,126	,926	,992
RespDistancia8	60,87	1799,656	,958	,992
RespDistancia9	60,95	1859,203	,875	,992
RespDistancia10	60,94	1795,038	,898	,992
RespDistancia11	60,86	1807,084	,948	,992
RespDistancia12	60,91	1797,704	,958	,992
RespDistancia13	60,73	1819,042	,933	,992
RespDistancia14	60,65	1780,497	,942	,992
RespDistancia15	60,81	1814,733	,949	,992
RespDistancia16	60,58	1843,266	,917	,992
RespDistancia17	60,58	1824,095	,919	,992
RespDistancia18	60,70	1798,395	,956	,992
RespDistancia19	61,04	1835,889	,884	,992
RespDistancia20	60,80	1801,803	,928	,992
RespDistancia21	60,55	1830,784	,927	,992
RespDistancia22	60,78	1798,349	,822	,993
RespDistancia23	60,67	1822,550	,940	,992
RespDistancia24	60,64	1786,331	,958	,992
RespDistancia25	60,73	1826,175	,943	,992

**Estadísticos de fiabilidad**

Alfa de Cronbach	N de elementos
,842	25

**Tabla B-2 Análisis de fiabilidad de las respuestas de seguridad**

**Estadísticos total-elemento**

	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Alfa de Cronbach si se elimina el elemento
RespSeguridad1	91,93	105,137	,425	,835
RespSeguridad2	91,53	108,384	,287	,840
RespSeguridad3	91,70	104,678	,540	,832
RespSeguridad4	91,51	103,088	,536	,831
RespSeguridad5	91,69	104,550	,495	,833
RespSeguridad6	92,05	109,190	,278	,840
RespSeguridad7	92,28	107,549	,275	,841
RespSeguridad8	91,47	105,715	,520	,833
RespSeguridad9	92,10	107,073	,394	,837
RespSeguridad10	91,62	105,567	,402	,836
RespSeguridad11	90,79	110,141	,335	,839
RespSeguridad12	91,52	100,560	,546	,830
RespSeguridad13	92,00	109,602	,199	,844
RespSeguridad14	91,59	107,540	,374	,837
RespSeguridad15	91,47	107,162	,492	,835
RespSeguridad16	92,24	106,475	,378	,837
RespSeguridad17	91,33	107,844	,266	,842
RespSeguridad18	91,53	105,229	,496	,833
RespSeguridad19	92,45	103,353	,383	,838
RespSeguridad20	91,53	105,882	,395	,836
RespSeguridad21	92,55	107,964	,394	,837
RespSeguridad22	91,46	106,718	,291	,841
RespSeguridad23	91,57	106,012	,293	,841
RespSeguridad24	91,97	103,781	,369	,838
RespSeguridad25	91,54	105,212	,497	,833



**Prueba T**

Estadísticos de muestras relacionadas					
		Media	N	Desviación típ.	Error tip. de la media
Par 1	DivergenciaDistanciaWFS	-,34	2800	,883	,017
	DivergenciaDistanciaVBAP	-,65	2800	1,039	,020
Par 2	RespSeguridadWFS	3,78	2800	1,016	,019
	RespSeguridadVBAP	3,79	2800	1,004	,019

Correlaciones de muestras relacionadas				
		N	Correlación	Sig.
Par 1	DivergenciaDistanciaWFS y DivergenciaDistanciaVBAP	2800	,452	,000
	RespSeguridadWFS y RespSeguridadVBAP	2800	,347	,000

Prueba de muestras relacionadas									
		Diferencias relacionadas				t	gl	Sig. (bilateral)	
		Media	Desviación típ.	Error tip. de la media	95% Intervalo de confianza para la diferencia				
					Inferior				Superior
Par 1	DivergenciaDistanciaWFS - DivergenciaDistanciaVBAP	-,309	1,015	,019	-,271	,346	16,087	2799	,000
	RespSeguridadWFS - RespSeguridadVBAP	-,014	1,154	,022	-,057	,029	-,639	2799	,523

**Tabla B-3 Prueba-T de muestras pareadas entre los sistemas WFS y VBAP**

Estadísticos descriptivos			
	Media	Desviación típica	N
DivergenciaDistancia0	-,34	,883	2800
DivergenciaDistancia1	-,65	1,039	2800

Contrastes multivariados <sup>a</sup>							
Efecto		Valor	F	Gl de la hipótesis	Gl del error	Sig.	Eta al cuadrado parcial
DivergenciaDistancia	Traza de Pillai	,085	258,802 <sup>b</sup>	1,000	2799,000	,000	,085
	Lambda de Wilks	,915	258,802 <sup>b</sup>	1,000	2799,000	,000	,085
	Traza de Hotelling	,092	258,802 <sup>b</sup>	1,000	2799,000	,000	,085
	Raíz mayor de Roy	,092	258,802 <sup>b</sup>	1,000	2799,000	,000	,085

a. Diseño: Intersección  
 Diseño intra-sujetos: DivergenciaDistancia  
 b. Estadístico exacto

**Prueba de esfericidad de Mauchly<sup>a</sup>**

Efecto intra-sujetos	W de Mauchly	Chi-cuadrado aprox.	gl	Sig.	Epsilon <sup>b</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Límite-inferior
DivergenciaDistancia	1,000	0,000	0		1,000	1,000	1,000

Contrasta la hipótesis nula de que la matriz de covarianza error de las variables dependientes transformadas es proporcional a una matriz identidad.

a. Diseño: Intersección  
 Diseño intra-sujetos: DivergenciaDistancia  
 b. Puede usarse para corregir los grados de libertad en las pruebas de significación promediadas. Las pruebas corregidas se muestran en la tabla Pruebas de los efectos inter-sujetos.

**Pruebas de efectos Intra-sujetos.**

Origen		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial
DivergenciaDistancia	Esfericidad asumida	133,303	1	133,303	258,802	,000	,085
	Greenhouse-Geisser	133,303	1,000	133,303	258,802	,000	,085
	Huynh-Feldt	133,303	1,000	133,303	258,802	,000	,085
	Límite-inferior	133,303	1,000	133,303	258,802	,000	,085
Error(DivergenciaDistancia)	Esfericidad asumida	1441,697	2799	,515			
	Greenhouse-Geisser	1441,697	2799,000	,515			
	Huynh-Feldt	1441,697	2799,000	,515			
	Límite-inferior	1441,697	2799,000	,515			

**Pruebas de contrastes Intra-sujetos**

Origen		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial
DivergenciaDistancia	Lineal	133,303	1	133,303	258,802	,000	,085
Error(DivergenciaDistancia)	Lineal	1441,697	2799	,515			

**Tabla B-4 Análisis de varianza de medidas repetidas para los sistemas WFS y VBAP**

Pruebas de los efectos inter-sujetos

Variable dependiente: DivergenciaDistancia

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial
Modelo corregido	190,499 <sup>a</sup>	7	27,214	29,576	,000	,036
Intersección	1378,086	1	1378,086	1497,695	,000	,211
angulo	45,721	1	45,721	49,689	,000	,009
reverb	,686	1	,686	,746	,388	,000
sistema	133,303	1	133,303	144,873	,000	,025
angulo * reverb	8,178	1	8,178	8,888	,003	,002
angulo * sistema	2,083	1	2,083	2,264	,132	,000
reverb * sistema	,046	1	,046	,050	,824	,000
angulo * reverb * sistema	,483	1	,483	,525	,469	,000
Error	5145,414	5592	,920			
Total	6714,000	5600				
Total corregida	5335,914	5599				

a. R cuadrado = ,036 (R cuadrado corregida = ,034)

Tabla B-5 ANOVA de un factor donde se observa la baja significatividad (p) de la reverberación

Pruebas de los efectos inter-sujetos

Variable dependiente: DivergenciaDistancia

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta al cuadrado parcial
Modelo corregido	1096,914 <sup>a</sup>	399	2,749	3,372	,000	,206
Intersección	1378,086	1	1378,086	1690,505	,000	,245
angulo	45,721	1	45,721	56,086	,000	,011
sistema	133,303	1	133,303	163,523	,000	,030
sonido	321,889	3	107,296	131,621	,000	,071
sujeto	181,396	24	7,558	9,272	,000	,041
angulo * sistema	2,083	1	2,083	2,555	,110	,000
angulo * sonido	11,638	3	3,879	4,759	,003	,003
angulo * sujeto	53,494	24	2,229	2,734	,000	,012
sistema * sonido	21,039	3	7,013	8,603	,000	,005
sistema * sujeto	45,751	24	1,906	2,338	,000	,011
sonido * sujeto	139,057	72	1,931	2,369	,000	,032
angulo * sistema * sonido	3,001	3	1,000	1,227	,298	,001
angulo * sistema * sujeto	16,774	24	,699	,857	,663	,004
angulo * sonido * sujeto	62,576	72	,869	1,066	,330	,015
sistema * sonido * sujeto	33,622	72	,467	,573	,999	,008
angulo * sistema * sonido * sujeto	25,570	72	,355	,436	1,000	,006
Error	4239,000	5200	,815			
Total	6714,000	5600				
Total corregida	5335,914	5599				

a. R cuadrado = ,206 (R cuadrado corregida = ,145)

Tabla B-6 ANOVA para medir los efectos del sonido, ángulo y sujeto sobre las respuestas de distancia