



FACULTAD DE INGENIERÍA Y CIENCIAS AGROPECUARIAS

EVALUACIÓN PSICOACÚSTICA DEL EFECTO DE PRECEDENCIA EN
TRES TÉRMINOS DE: AZIMUT, ELEVACIÓN Y DISTANCIA CON
DISTINTOS NIVELES DE REVERBERACIÓN A TRAVÉS DE
AURALIZACIÓN BINAURAL

AUTORES

Andrés Aguaisa y Karla Morán

AÑO

2018



FACULTAD DE INGENIERÍA Y CIENCIAS AGROPECUARIAS

EVALUACIÓN PSICOACÚSTICA DEL EFECTO DE PRECEDENCIA EN TRES
TÉRMINOS DE: AZIMUT, ELEVACIÓN Y DISTANCIA CON DISTINTOS
NIVELES DE REVERBERACIÓN A TRAVÉS DE AURALIZACIÓN BINAURAL.

Trabajo de Titulación presentado en conformidad con los requisitos
establecidos para optar por el título Ingenieros en Sonido y Acústica.

Profesor Guía

MSc. Jorge Páez Rodríguez

Autores

Víctor Andrés Aguaisa Moreno

Karla Estefanía Morán Vargas

Año

2018

DECLARACIÓN DEL PROFESOR GUÍA

“Declaro haber dirigido este trabajo, EVALUACIÓN PSICOACÚSTICA DEL EFECTO DE PRECEDENCIA EN TRES TÉRMINOS DE: AZIMUT, ELEVACIÓN Y DISTANCIA CON DISTINTOS NIVELES DE REVERBERACIÓN A TRAVÉS DE AURALIZACIÓN BINAURAL, a través de reuniones periódicas con los estudiantes Víctor Andrés Aguaisa Moreno y Karla Estefanía Morán Vargas, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación”

Jorge Páez Rodríguez

Máster en Gestión y Evaluación de la Contaminación Acústica

C.I. 175677315-4

DECLARACIÓN DEL PROFESOR CORRECTOR

“Declaro haber revisado este trabajo, EVALUACIÓN PSICOACÚSTICA DEL EFECTO DE PRECEDENCIA EN TRES TÉRMINOS DE: AZIMUT, ELEVACIÓN Y DISTANCIA CON DISTINTOS NIVELES DE REVERBERACIÓN A TRAVÉS DE AURALIZACIÓN BINAURAL, de los estudiantes Víctor Andrés Aguaisa Moreno y Karla Estefanía Morán Vargas, dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación”.

Carlos Andrés Ramos Romero
Magíster en Acústica y Vibraciones
C.I. 171404496-1

DECLARACIÓN DE AUTORÍA DE LOS ESTUDIANTES

“Declaramos que este trabajo es original, de nuestra autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes.”

Víctor Andrés Aguaisa Moreno
C.I. 172716167-9

Karla Estefanía Morán Vargas
C.I. 171959465-5

AGRADECIMIENTOS

Agradezco a Dios por otorgarnos la capacidad y el talento para superar cada una de las dificultades que se presentaron a lo largo de esta ingeniería.

A mis padres y hermanos quienes me ofrecieron su apoyo incondicional y sin ningún tipo de cuestionamiento.

A mis amigos por ser un importante sostén anímico en los momentos de duda y dificultad.

Andrés Aguaisa

AGRADECIMIENTOS

Agradezco a mis padres por haberme inculcado grandes valores. Agradezco a mi hermano por ser mi compañero fiel. También a todas las personas, que han sido parte de mi vida universitaria.

Karla Morán

DEDICATORIA

Dedico este trabajo de titulación a todas las personas quienes se interesan por la psicoacústica y dedican su vida al estudio y desarrollo de soluciones para quienes poseen problemas auditivos. A todos los colegas de esta profesión dentro del Ecuador para que podamos seguir por la senda de la investigación y la innovación acústica desde este pequeño y bello rincón del mundo.

Andrés Aguaisa

DEDICATORIA

Quiero dedicar este trabajo de tesis a mi abuelita, Piedad, quien fue y sigue siendo un pilar fundamental en mi vida, quien me alivia y me da fortaleza en los momentos difíciles.

Karla Morán

RESUMEN

El presente trabajo de titulación es una investigación de la localización del sonido en tres términos de: azimut, distancia y elevación. Se realizaron grabaciones binaurales de impulsos sonoros en distintas salas. Posteriormente se convolucionaron a través el software MatLab junto con dos estímulos: música y ruido rosa, recreando de esta manera los sonidos en cada una de las salas.

Se evaluó a una población de hombres y mujeres que han aprobado la materia de entrenamiento auditivo de la carrera Ingeniería en Sonido y Acústica. A cada participante se le hizo escuchar mediante audífonos varios sonidos auralizados y se les pidió que seleccionen, en un mapa simulado de cada sala, el punto de origen del sonido.

Los resultados arrojan que los encuestados presenta dificultades al determinar la localización exacta de un estímulo, independientemente de la reverberación de la sala. Sin embargo, un tiempo de reverberación aproximado de 1 segundo ayuda a reconocer con facilidad la ubicación del sonido. Se evidenció que los participantes tiene mayor porcentaje de aciertos en el eje de azimut, en comparación a la distancia y elevación.

ABSTRACT

This degree work is an investigation of the identification of sound origin in three terms: azimuth, distance and elevation. Binaural recordings of sound impulses were performed in different rooms. Subsequently, recreating the sounds in each of the rooms using the MatLab software, by convolution along with two stimuli: music and pink noise.

It evaluates a population of men and women who have approved the ear training subject of the major in sound and acoustic engineering. Each participant heard several auralized sounds by headphones, then asked to select the point where they think the sound originated on a simulated map of each room.

The results show that the participants being has difficulties in determining the exact location of a stimulus, regardless of the reverberation of the room. But, a reverb time of approximately 1 second helps to recognize the location of the sound easier. It was evident that the participants being has a higher percentage of successes in the azimuth axis, compared to the distance and elevation.

ÍNDICE

1	Introducción	1
1.1	Antecedentes	1
1.2	Justificación.....	4
1.3	Objetivos	5
1.3.1	General	5
1.3.2	Específicos.....	5
1.4	Alcance	6
2	Fundamento teórico	7
2.1	Localización de una fuente sonora	7
2.1.1	ITD	7
2.1.2	ILD	7
2.2	Efecto de precedencia	8
2.3	Tiempo de reverberación	9
2.4	Procesamiento de señales.....	10
2.4.1	Auralización y convolución	11
3	Metodología	12
3.1	Variables	12
3.2	Descripción de las salas	14
3.2.1	Sala 1: Sala comunal - Edificio Aranjuez.....	14
3.2.2	Sala 2: Estudio de grabación - Magic Estudio	15
3.2.3	Sala 3: Auditorio Queri - Universidad de las Américas	16

3.2.4	Sala 4: Estudio de grabación - CR3	17
3.3	Mediciones del tiempo de reverberación	19
3.3.1	Instrumentación	19
3.3.2	Descripción del proceso	19
3.3.3	Resultados de los tiempos de reverberación.....	20
3.4	Grabación de respuesta impulsiva.....	22
3.4.1	Instrumentación	22
3.4.2	Proceso de grabación	23
3.4.2.1	Ubicación del dummy head	23
3.4.2.2	Puntos de grabación.....	24
3.4.2.3	Grabación in situ.....	27
3.5	Selección de salas	28
3.6	Simulaciones en Sketch Up de las salas escogidas	28
3.6.1	Sala 1: Sala comunal - Edificio Aranjuez.....	30
3.6.2	Sala 3: Auditorio Queri - Universidad de las Américas	31
3.6.3	Sala 4: Estudio de grabación - CR3	31
3.7	Procesamiento de señales.....	31
3.7.1	Instrumentación	31
3.7.2	Selección de estímulos	32
3.7.3	Corrección de nivel del dummy head	33
3.7.3.1	Instrumentación	33
3.7.3.2	Metodología y resultados	33

3.7.4	Proceso de convolución.....	35
3.7.4.1	Entrada	35
3.7.4.2	Procesamiento.....	36
3.7.4.3	Salida.....	37
3.8	Pruebas psicoacústicas	38
3.8.1	Selección de audios convolucionados.....	38
3.8.2	Instrumentación	41
3.8.3	Cálculo de la población.....	41
3.8.4	Evaluación psicoacústica.....	42
4	Resultados y análisis	44
4.1	Audiometrías	44
4.2	Aciertos en los tres términos.....	45
4.3	Aciertos en azimut.....	53
4.4	Aciertos en elevación	57
4.5	Aciertos en distancia	63
4.6	Por elección de puntos.....	67
4.6.1	Mujeres	68
4.6.2	Hombres	71
4.7	Por punto más seleccionado.....	74
4.7.1	Mujeres	74
4.7.2	Hombres	75
5	Conclusiones y recomendaciones	76

5.1	Conclusiones.....	76
5.2	Recomendaciones	80
	REFERENCIAS	82
	ANEXOS	85

1 Introducción

1.1 Antecedentes

El ser humano tiene la capacidad para localizar fuentes sonoras puntuales en tres ejes: horizontal (azimut), vertical (elevación) y profundidad (distancia). En azimut la posición está definida por las diferencias interaurales, diferencia entre ambos oídos, de tiempo (ITD) y de nivel (ILD). La información espectral tiene gran influencia en el eje vertical debido a la pérdida de contenido en ciertas frecuencias por la ubicación del receptor con respecto a la fuente. En profundidad se puede encontrar el mayor número de indicadores como: la intensidad, reverberación e información espectral. Los nuevos hallazgos experimentales en neurofisiología de la audición indican que la corteza auditiva no es necesaria para discriminar los parámetros físicos del sonido: intensidad, altura tonal, timbre, y duración. El oído es quien se encarga de analizarlos y codificarlos (Arias y Ramos, 2003, pp. 375-376).

La anatomía propia del oído humano, junto con procesos neuronales, permiten conocer la procedencia de un sonido mediante la discriminación de los parámetros mencionados anteriormente. Estos procesos han sido objeto de estudios cuyos resultados se presentan en diferentes metodologías que permiten recrearlos, con el fin de obtener la localización de una fuente sonora. Estos métodos están divididos en: DOA, ITD, ILD y HRTF (Pourmohammad y Ahadi, 2012, p. 307).

El tiempo de reverberación contribuye a que los seres humanos logren ubicar con mayor precisión la posición de una fuente sonora. Esto se debe a que el ser humano desde su niñez, realiza sus actividades cotidianas en ambientes reverberantes, por lo cual se adapta a la reverberación. Así lo ha demostrado un estudio que concluye que, en un ambiente reverberante, la percepción de

distancia de la fuente sonora es más precisa que en un ambiente anecoico (Paquier *et al.*, 2016, p. 198).

Las señales captadas por cada uno de los oídos son diferentes y dependen de factores referentes al tiempo de arribo, parámetros de la onda sonora y la anatomía del ser humano. Con respecto a este último factor, existe un método HRTF. Según Balan *et al.* (2014, pp. 34-35), Pourmohammad y Ahadi (2012, p. 307) HRTF se define como la señal modificada después de un recorrido hasta llegar al oído, siendo afectada por la difracción, reflexiones de la cabeza y la parte externa del oído (pinea). La influencia de la pinea es de suma importancia debido a que provoca pérdidas de nivel en un rango de frecuencias específico, simulando un filtro elimina banda o también conocido como filtro notch, en la respuesta de HRTF, y también picos que afectan a la localización de la fuente especialmente en la elevación.

La estimación de la ubicación de la fuente en el eje horizontal (azimut) depende del ángulo de procedencia, la acústica del recinto y el tipo de sonido emitido. De la Torre (2016, p. 71), concluyó que la localización de la fuente, respecto al ángulo de azimut se limita a ± 45 grados, por lo que dentro de esta angulación existe menos incertidumbre. A medida que se presenta mayor reverberación, la localización del ángulo puntual no es precisa. En azimut, tanto la energía directa como la reverberante, es distinta para ambos oídos.

Las recreaciones binaurales son un recurso preciso que permite el análisis tanto de parámetros acústicos como de percepción auditiva. Estas tienen por objeto obtener la respuesta impulsiva del recinto y, a partir de ahí, recrear las características acústicas de la sala, como el STI que describe la inteligibilidad de la palabra del espacio evaluado. Un estudio reciente revela que se pueden realizar predicciones precisas de STI a través de simulación acústica cuando se

obtiene correctamente la respuesta impulsiva de la sala en cuestión (Zhu *et al.*, 2015, p. 73).

Una reproducción binaural exitosa no depende únicamente del rendimiento adecuado del recinto, se evalúa de acuerdo con dos criterios: autenticidad y plausibilidad de la sala presentada. La autenticidad depende de una correcta obtención de la respuesta impulsiva de modo que la recreación del recinto tenga un rendimiento adecuado. La plausibilidad mide el nivel de credibilidad y aceptación que tendrán las personas al experimentar estímulos a través de estas recreaciones. Oberem *et al.* (2016, p. 78) realizó un testeo de plausibilidad con 80 sujetos de prueba en el cual se concluyó que la reproducción binaural a través de audífonos con métodos HRTF y HpTF fueron en general plausibles y, por lo tanto, pueden ser utilizados en experimentos psicoacústicos.

Sin embargo, es muy importante mencionar que la fidelidad de la recreación binaural dependerá del espectro en frecuencia correspondiente al sonido a recrearse. Las altas frecuencias son muy susceptibles a coloraciones causadas por reverberación (entorno real) o respuesta impulsiva (recreación binaural). En el mismo estudio, se obtuvo como resultado que el 40% de los sujetos de estudio lograron percibir la diferencia en la recreación binaural de ruido rosa debido a la coloración en altas frecuencias.

Se puede afirmar que no se han realizado estudios de forma equitativa en los tres ejes; Ghamdan *et al.* (2017, p. 2) expresa que varias investigaciones se basan en la estimación de percepción sonora en azimut y en la actualidad en elevación, mientras que para la distancia se realizan estudios con arreglos microfónicos.

La ausencia de una preparación previa para la percepción sonora puede generar una desviación importante en los resultados. Una persona promedio tiende a subestimar la distancia que existe entre fuente y receptor como lo demuestra Paquier *et al.* (2016, p. 186), que concluyó, que los sujetos de prueba no fueron precisos al determinar la distancia de la fuente en tres contextos: percepción auditiva, visual y una combinación de ambas. Nilsson y Schenkman (2016, p. 223), realizaron un estudio donde los resultados de discriminación de la fuente sonora son mejores para personas con ceguera de nacimiento debido a que la audición es su único medio de ubicación y se han acostumbrado al mismo.

Cabe mencionar que en áreas como la robótica estos estudios se vuelven fundamentales para su implementación en sistemas de ubicación como lo menciona Zhong *et al.* (2016, p. 84), quien plantea un algoritmo para detección de fuentes sonoras múltiples, el cual se basa en principios del oído humano. Con estos antecedentes, en este trabajo de titulación se plantea profundizar el análisis de percepción en los tres ejes de ubicación con distintos niveles de reverberación. La aplicación de una recreación virtual se realiza mediante una convolución que es la combinación de un sonido puntual con las características de un recinto, es decir la respuesta impulsiva del mismo.

1.2 Justificación

El aporte de esta investigación tiene dos partes importantes, la primera es el estudio de la ubicación de un estímulo sonoro en tres términos: azimut, elevación y distancia, y la segunda es un análisis del efecto que tiene la reverberación del recinto sobre la localización de un sonido determinado. En el Ecuador existe un estudio que se enfocó únicamente en la localización de un estímulo sonoro mediante auralización en el eje de azimut con niveles de reverberación bastante elevados, esta es una razón por la cual se ve la importancia de abordar este tema, pero con mejoras como, una extensión del estudio en los tres términos, ya mencionados, y con niveles de reverberación menores. Esto a su vez permite

una investigación en el campo de la psicoacústica del efecto de precedencia, que en el país es casi nula y a nivel global muy escasa.

A su vez el presente trabajo es una base para futuras investigaciones que pueden incluir tecnología más avanzada, es decir llevar la investigación a programación o procesamiento de señales para diferentes utilidades en el campo tecnológico como en videojuegos o robots, o reconocimiento de la ubicación de una fuente puntual mediante programación. Todo lo antes mencionado será realizado primordialmente con bases acústicas y promoviendo la importancia de esta área de investigación.

1.3 Objetivos

1.3.1 General

Estimar el efecto de precedencia de un sonido específico con diferentes niveles de reverberación, mediante auralización y estímulos fusionados, en tres términos: azimut, elevación y distancia.

1.3.2 Específicos

- Seleccionar tres recintos con distintos tiempos de reverberación que cumplan con condiciones óptimas de altura y simetría para el estudio.
- Obtener la respuesta impulsiva de cada sala mediante una dummy head, con variación en tres términos: azimut, elevación y distancia.
- Auralizar diferentes estímulos sonoros con los datos de los recintos obtenidos previamente mediante el proceso de convolución en la plataforma MatLab.

- Aplicar el entrenamiento auditivo que tienen los estudiantes de la carrera de Ingeniería en Sonido y Acústica para reducir la incertidumbre en los resultados.
- Comparar los resultados obtenidos, observando el número de aciertos y porcentajes de precisión considerando criterios como tiempo de reverberación, tipo de estímulo sonoro, sexo, etc.

1.4 Alcance

Esta investigación abarca una serie de aspectos dentro del área de la acústica. En primer lugar, la medición y análisis de tiempos de reverberación en distintas salas para su posterior elección; estas van a variar según su arquitectura y tratamiento acústico. Segundo, el proceso de auralización mediante la respuesta impulsiva de cada sala con base en el procesamiento de señales. Finalmente, la psicoacústica que se encarga de relacionar la parte psicológica del ser humano y la percepción de las ondas sonoras, en este caso, se analiza la localización de una fuente sonora.

El alcance del proyecto de tesis se rige a las personas que hayan cursado la materia de entrenamiento auditivo en la carrera de Ingeniería en Sonido y Acústica, comprendidas entre los 20 a 26 años. A cada una de las personas se les realizará un test con el fin de obtener resultados para la ubicación de un estímulo en los tres términos: azimut, elevación y distancia. Además, se necesita personas que tengan una audición normal, por lo cual previo a la encuesta se realizará un test auditivo virtual para validar la participación de las personas.

2 Fundamento teórico

2.1 Localización de una fuente sonora

La localización de una fuente sonora para un ser humano implica dar una dirección, una distancia y también una extensión espacial. Esta localización depende de factores externos como el ambiente en el que se reproduzca el sonido o de la anatomía de cada uno de los seres humanos. Se puede estudiar la localización de una fuente sonora con sonido monoaural o con un binaural, nos enfocaremos en este último (Pulkki y Karjalainen, 2014, pp. 219-220).

El sonido binaural hace referencia a un estímulo que se escucha por los dos oídos con cierta diferencia de tiempo, denominada diferencia interaural. Existen otros parámetros específicos que tienen base en la separación de los oídos del ser humano y son el ITD y el ILD, a continuación se describen cada término.

2.1.1 ITD

La diferencia de tiempo interaural es la diferencia de tiempo con la que llega un sonido a un oído y al otro, debido a la distancia entre los mismo y también a que el sonido se refracta en la cabeza. La máxima diferencia entre la llegada de un sonido a los oídos es de 700 μ s que se da a 90 grados en azimut (Bomhardt, 2017, p. 4).

2.1.2 ILD

La diferencia de nivel interaural es la diferencia de nivel de presión con la que llega un sonido a un oído y al otro, debido a la "Sombra del sonido". La sombra del sonido es la atenuación que produce la cabeza en el nivel de presión de un estímulo. El ILD depende de la frecuencia y es más notoria en frecuencias altas (Goldstein, 2008, p. 442).

2.2 Efecto de precedencia

El efecto de precedencia o también denominado efecto Hass es un mecanismo de ayuda para la localización del sonido en espacios reverberantes. Cuando se tienen dos sonidos iguales, pero uno tiene cierto retraso, el ser humano para encontrar la ubicación de la fuente se basa en el sonido original y no en el retardado, en otras palabras, suprime las primeras reflexiones basando su respuesta en sonido directo (Pulkki y Karjalainen, 2014, pp. 239-240).

El sonido directo y el retardado se percibirán como un solo estímulo si el retardo no pasa de 30 a 40 ms, caso contrario se escucharán como dos estímulos distintos (Pulkki y Karjalainen, 2014, pp. 239-240).

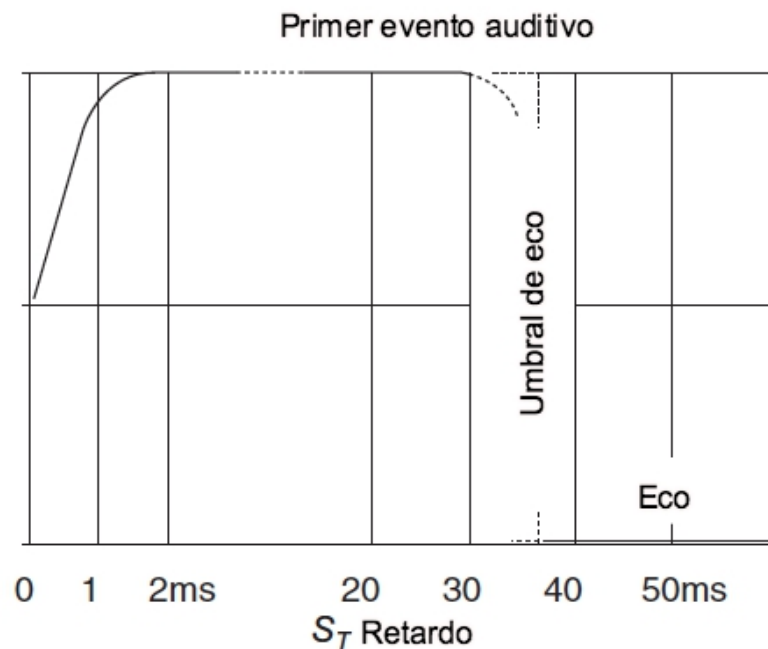


Figura 1. Umbral del eco.

Adaptada de Pulkki y Karjalainen, 2014, p. 240.

También se debe tomar en cuenta la diferencia de nivel entre el sonido directo y el sonido con retardo, es así que si el sonido con retardo está por debajo del

umbral, el timbre del mismo cambiará como se explica en la Figura 2 (Pulkki y Karjalainen, 2014, pp. 239-240).

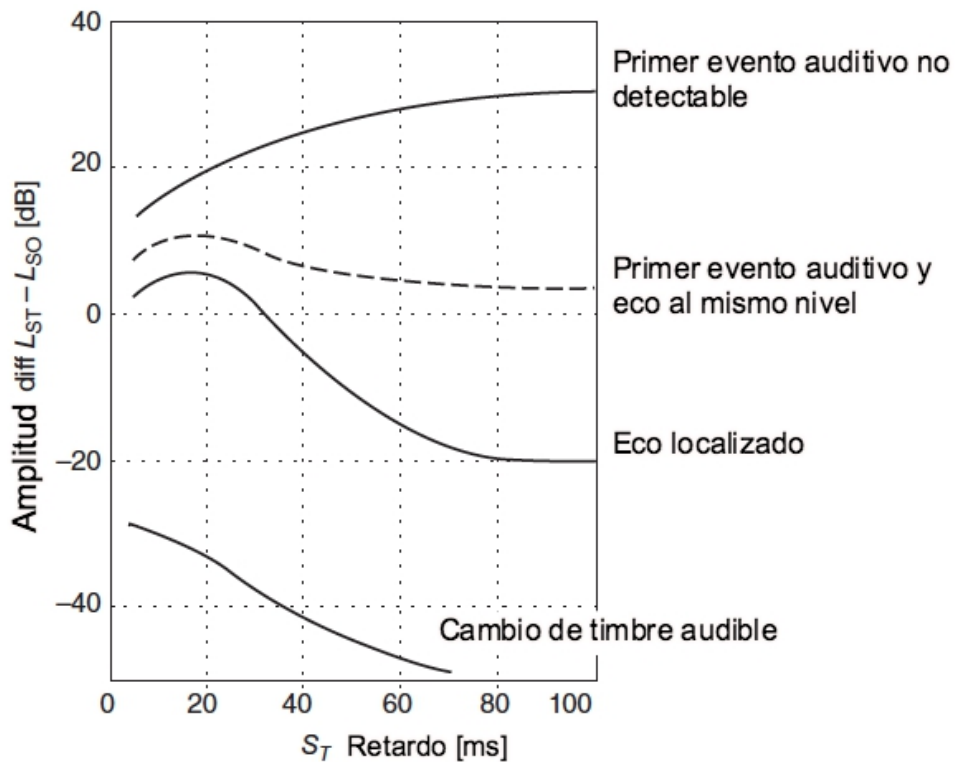


Figura 2. Umbral de amplitud.

Adaptada de Pulkki y Karjalainen, 2014, p. 240.

2.3 Tiempo de reverberación

El tiempo de reverberación es el tiempo que le toma a un sonido, después de detenerse, en decaer 60 dB por debajo de su valor inicial de intensidad. Este parámetro nos permite conocer la calidad acústica de la sala, que se ve determinada por factores como: materiales de las superficies, geometría de la sala, objetos reflejantes o absorbentes, entre otros (McMullan, 2012, p. 239).

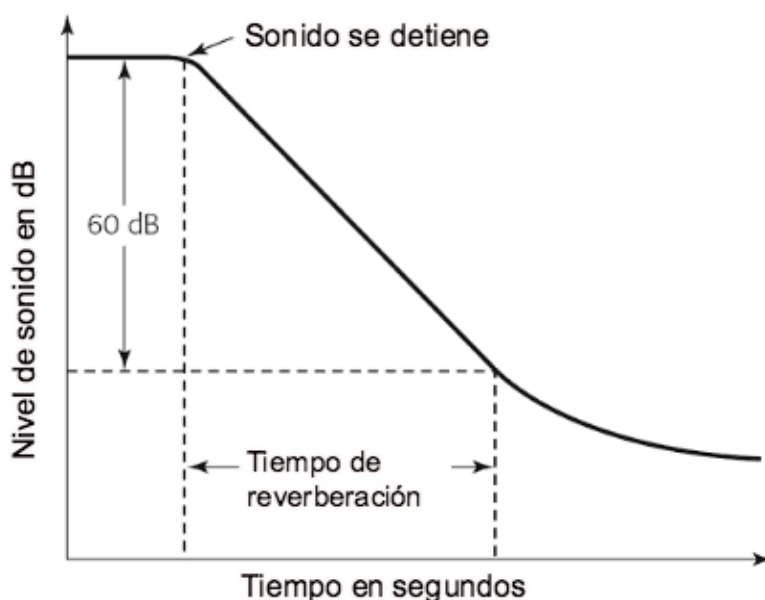


Figura 3. Tiempo de reverberación.

Adaptada de McMullan, 2012, p. 239.

2.4 Procesamiento de señales

Las señales al pasar del dominio físico al dominio digital pueden ser sometidas a una serie de procesamientos, siempre y cuando se realice la correcta discretización mediante softwares especializados.

Una señal continua es una sucesión de valores que corresponden a una variable independiente. La principal diferencia con una señal discreta es que la variable independiente se rige dentro de un conjunto definido y limitado de valores, contrario a la señal continua en donde el valor no se rige a un margen definido.

El muestreo juega un papel fundamental dentro de la discretización de la señal; se trata de tomar muestras de una señal continua dentro de una cierta frecuencia para poder recrearla en tiempo real. Los procesadores actuales permiten

manejar frecuencias de muestreo muy elevadas lo que garantiza una alta fidelidad con respecto a las señales analógicas.

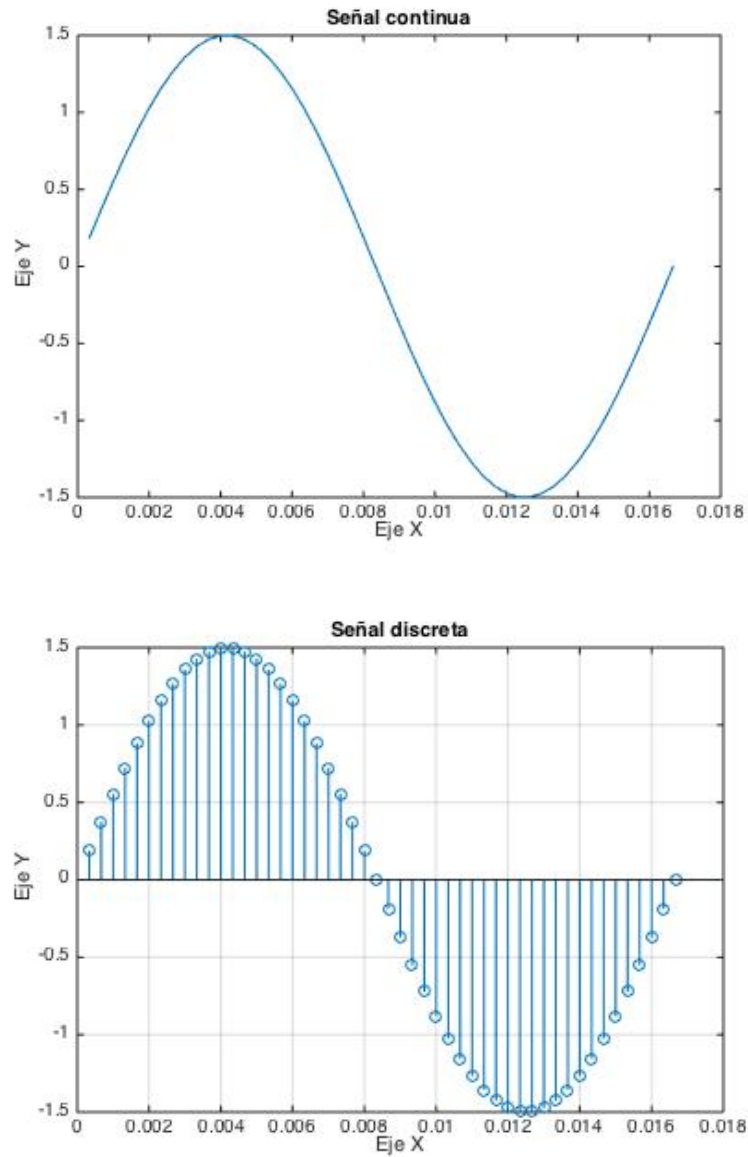


Figura 4. Comparación de señal continua y señal discreta.

2.4.1 Auralización y convolución

El proceso de convolución consiste en tomar dos señales y realizar una operación entre ambas de manera que las características, como las primeras

reflexiones de una señal, pasen a formar parte de la otra, creando así una señal completamente nueva. Para que este proceso sea posible ambas señales deben estar en el dominio discreto, de lo contrario no se puede realizar la operación entre ambas.

La auralización es un concepto que permite la extracción de las características espaciales de la señal analizada, para esto se ve necesaria la implementación de sistemas estereofónicos que capten con precisión todo el entorno reflectante de un recinto. Se denomina HRIR, a la descomposición en dominio de la frecuencia de la señal impulsiva captada mediante los sistemas anteriormente descritos.

3 Metodología

3.1 Variables

Tabla 1.

Descripción de variables a utilizarse en el estudio.

Variable	Tipo	Evaluación	Análisis	Unidad
Tiempo de reverberación	Numérico	Mediciones con base en el proceso descrito en la norma UNE EN ISO 3382.	Comparación entre los recintos medidos para la selección final.	Segundos
Porcentaje de aciertos y	Numérico	Testeo de opción múltiple	Comparación de los	Porcentaje

errores de los sujetos de experimentación		usando distintas muestras convolucionadas con diferentes niveles de reverberación.	resultados entre los tres niveles de reverberación a través de gráficos.	
Azimut	Numérico	Grabación de impulsos en diferentes grados en el eje horizontal, con pasos de 45 grados.	Número de aciertos y desaciertos.	Escalar
Distancia	Numérico	Grabación de impulsos con variaciones de azimut y a diferentes distancias en el eje de profundidad.	Número de aciertos y desaciertos.	Escalar
Elevación	Numérico	Grabación de impulsos con variaciones de azimut y a diferentes distancias en el eje de altura.	Número de aciertos y desaciertos.	Escalar

3.2 Descripción de las salas

3.2.1 Sala 1: Sala comunal - Edificio Aranjuez¹

El salón está ubicado en el onceavo piso del edificio Aranjuez ubicado en el sector Bellavista de la ciudad de Quito.



Figura 5. Ubicación geográfica del Edificio Aranjuez.

El recinto es una sala comunal, con fines recreativos, por lo cual cuenta con un espacio destinado a una cocina, accesos a sauna y turco por puertas. La sala es asimétrica y cuenta con paredes de concreto y ventanas de vidrio simple, mientras que el piso es totalmente de porcelanato. Además, tiene un mesón de mármol en uno de sus rincones. No tiene ningún tipo de amoblado ni tratamiento acústico.

¹ A partir de ahora se llamará al recinto "Aranjuez".



Figura 6. Foto de 180 grados correspondiente a la sala comunal del edificio Aran juez.

3.2.2 Sala 2: Estudio de grabación - Magic Estudio²

El recinto Magic Estudio se encuentra ubicado en el sector de la Mariana de Jesús en el norte de la ciudad de Quito.



Figura 7. Ubicación geográfica de Magic Estudio.

Al ser un estudio de grabación, el recinto cuenta con acondicionamiento acústico, como paneles acústicos variables (piedra, madera, material absorbente), zonas de difusión y vidrio doble. Sin embargo, al tener paneles de piedra en el sector

² A partir de ahora se llamará al recinto "Magic".

izquierdo, se optó por utilizar paneles absorbentes de manera que estas reflexiones no influyan en la grabación de respuesta impulsiva.

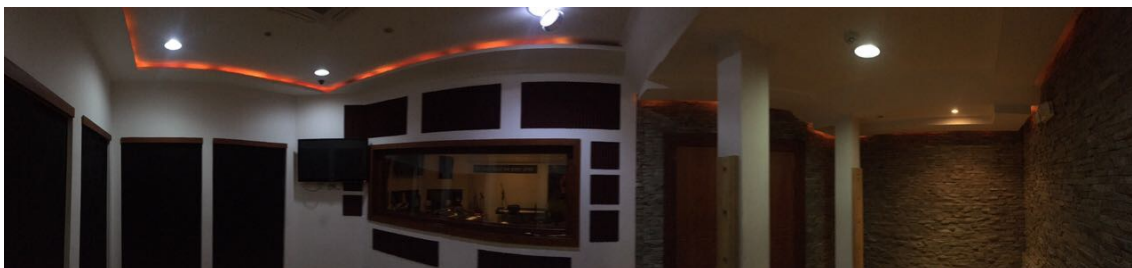


Figura 8. Foto de 180 grados correspondiente al estudio de grabación Magic.

3.2.3 Sala 3: Auditorio Queri - Universidad de las Américas³

Esta sede se encuentra ubicada en el sector de la Eloy Alfaro y Queri al norte de la ciudad de Quito.



Figura 9. Ubicación geográfica de la sede Queri.

³ A partir de ahora se llamará al recinto "Queri".

El recinto está destinado a conciertos de la Escuela de Música, charlas y conferencias. La sala no tiene ningún tipo de acondicionamiento acústico y cuenta con una altura bastante considerable. Su techo no es paralelo con el piso y es simétrico de manera lateral. El material de las paredes consiste, principalmente, en gypsum y techo flotante. Las bases y columnas son metálicas, pero no representan un porcentaje considerable de la estructura del recinto.

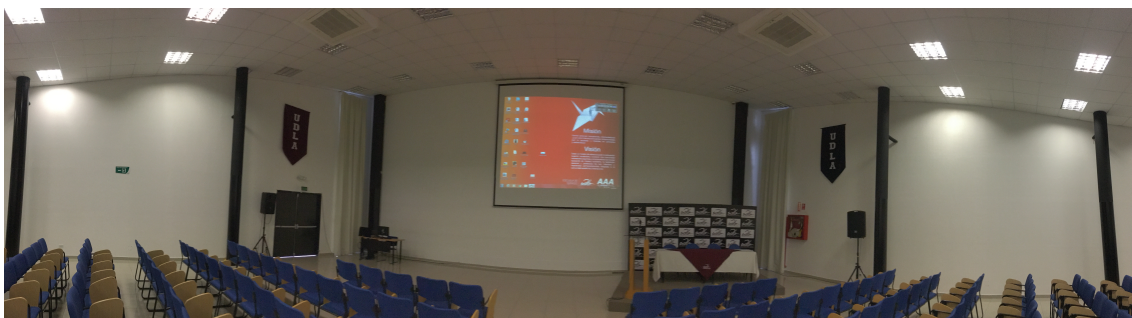


Figura 10. Foto de 180 grados correspondiente al auditorio de la UDLA sede Queri.

3.2.4 Sala 4: Estudio de grabación - CR3⁴

Este recinto se encuentra dentro de la sede Granados de la Universidad de las Américas, el cual está ubicado en la Avenida de los Granados y De los Colimes.

⁴ A partir de ahora se llamará al recinto “CR3”.



Figura 11. Ubicación geográfica de la sede Granados de la Universidad de las Américas.

Este estudio de grabación tiene acondicionamiento acústico como: difusores, paneles acústicos variables (madera, piedra y material absorbente), vidrios dobles y paneles móviles. El material del piso es madera y el techo cuenta con paneles absorbentes. También existen paneles móviles que los ubicamos en las ventanas para evitar reflexiones.

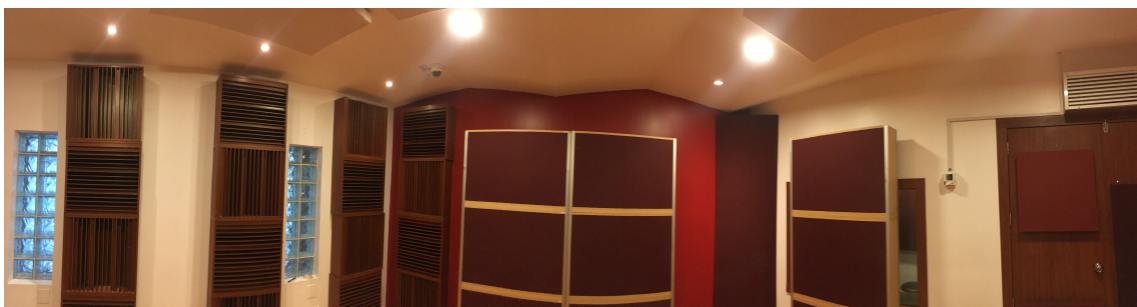


Figura 12. Foto de 180 grados correspondiente al estudio de grabación CR3 en la UDLA sede Granados.

3.3 Mediciones del tiempo de reverberación

3.3.1 Instrumentación

- Sonómetro NTi XL2.
- Micrófono NTi 2371.
- Globos No 5
- Flexómetro y metro

3.3.2 Descripción del proceso

Las mediciones del tiempo de reverberación se hicieron con base en la norma UNE EN ISO 3382 que se detalla en el Anexo 2. Dentro de las salas se seleccionaron cuatro puntos en total para la medición de los tiempos de reverberación, siendo tres de fuente y uno de micrófono. Cabe aclarar que luego se intercambié una posición de micrófono por una de fuente. La fuente de elección fue un globo con un diámetro aproximado de 15 cm y se utilizó el sonómetro NTi XL2 ya que permite medir la respuesta impulsiva.

Al momento de la medición se encontraron únicamente dos personas en la sala, como lo recomienda la norma. Se explotó cada globo con una aguja en los puntos marcados, variando la altura. Al momento de explotar los globos se alejó se tomó una distancia adecuada para no ser un obstáculo en las mediciones.



Figura 13. Proceso de medición de tiempo de reverberación en la sala Aranjuez.

3.3.3 Resultados de los tiempos de reverberación

En cuanto al cálculo se realizó un promedio lineal de las mediciones por tercio de octava desde 100 Hz hasta 5000 Hz y luego un promedio para ver el tiempo de reverberación total de la sala. Las tablas de los datos de medición de cada sala se encuentran en el Anexo 3.

La Tabla 2 y la Figura 14 corresponden a los valores promedio obtenidos en cada sala.

Tabla 2.

Tiempo de reverberación de todos los recintos en bandas de tercio de octava.

Frecuencia (Hz)	Tiempo de reverberación de salas (s)			
	CR3	Magic	Queri	Aranjuez
100	0,36	0,78	1,81	4,09
125	0,35	0,67	1,81	2,78
160	0,41	0,75	1,81	2,48
200	0,40	0,66	1,27	2,92
250	0,37	0,49	1,19	3,06
315	0,36	0,51	1,04	3,10
400	0,30	0,43	0,95	3,14
500	0,41	0,64	0,82	2,96
630	0,35	0,68	0,88	3,09
800	0,33	0,61	0,88	3,05
1000	0,34	0,55	0,93	3,01
1250	0,36	0,49	1,02	2,84
1600	0,34	0,41	1,00	2,54
2000	0,39	0,48	1,08	2,59
2500	0,36	0,49	1,17	2,25
3150	0,33	0,44	1,16	1,89
4000	0,34	0,40	1,08	1,86
5000	0,34	0,38	0,99	1,68
TR Global	0,36	0,55	1,16	2,74

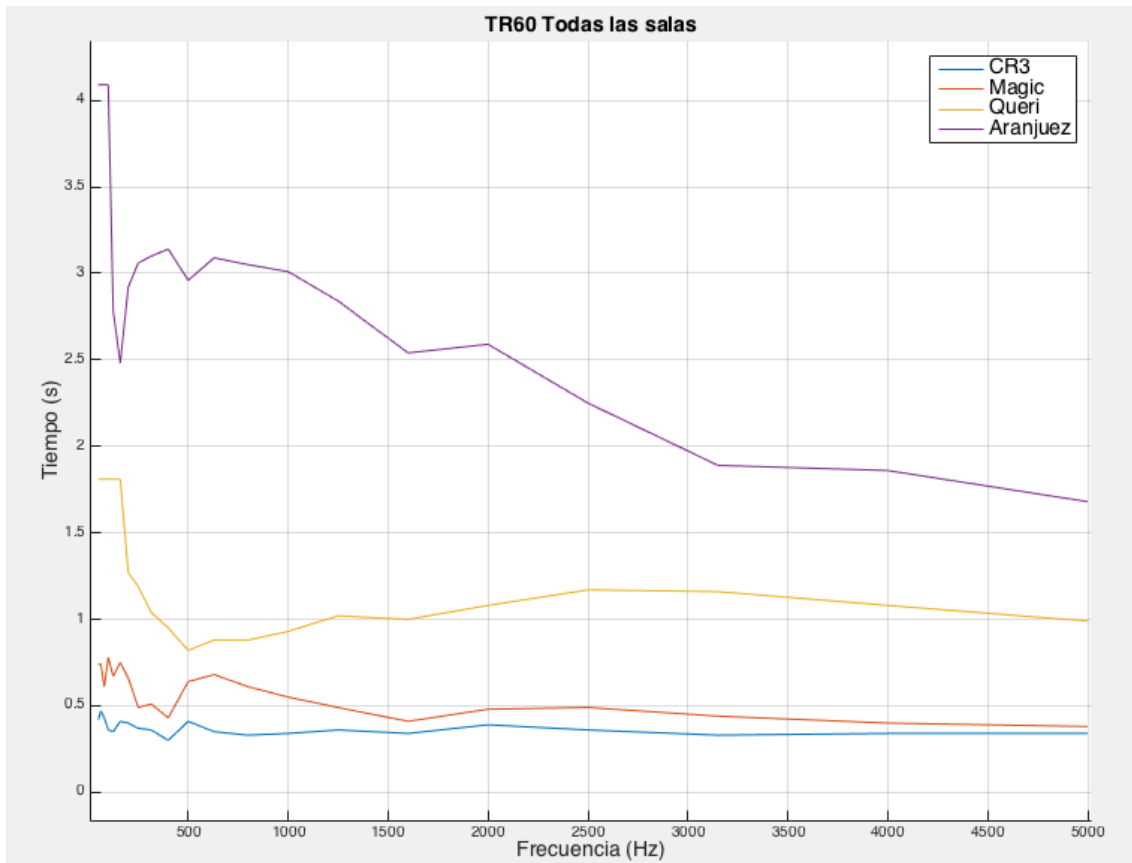


Figura 14. Gráfico de tiempos de reverberación de todos los recintos.

3.4 Grabación de respuesta impulsiva

3.4.1 Instrumentación

- Dummy head: Neumann KU 100⁵
- Computadora portátil MacBook Pro
- Interfaz PreSonus Audiobox - 1818VSL
- Software Pro Tools 10
- Globos No 5
- Flexómetro y metro

⁵ Las especificaciones del dummy head se encuentran en el Anexo 4.

3.4.2 Proceso de grabación

3.4.2.1 Ubicación del dummy head

El primer factor a considerar fue la distancia de separación con respecto a las paredes; se determina una separación de más de 0,5 m para evitar las primeras reflexiones.

Se buscó que la cabeza tenga una posición completamente céntrica, de manera que toda la carga energética del impulso a generarse se distribuya de igual forma para ambos lados. Debido a que muchas veces los puntos para la grabación quedaban a menos de 0,5 m, se decidió en algunos recintos ubicar la cabeza más apegada a la parte posterior.

En cuanto a la altura del dummy head se consideró la altura promedio de hombres de nacionalidad ecuatoriana, esta es de 1,67 m (Metro, s.f.).



Figura 15. Dummy head ubicado a una altura de 1,67 m del suelo, en la sala CR3.

3.4.2.2 Puntos de grabación

A partir de la investigación realizada por De la Torre (2016), se determina que los evaluados tienen mejor audición en un plano cerrado de azimut. Sin embargo, no existe un estudio con respecto a la distancia y elevación por lo cual se propuso los siguientes parámetros:

- 5 posiciones en azimut
- 2 variaciones de distancia
- 2 variaciones de elevación

Para marcar los puntos en azimut se utilizó un graduador que se ubicó en el centro del pedestal que sostenía la cabeza. Para la distancia se establecieron dos valores; 1 m y 2 m a partir del centro de la cabeza, esto tomando en cuenta que existía la posibilidad de realizar la grabación en estos puntos con las dimensiones de las salas. En cuanto a la elevación se usó como centro la nariz, a partir de la cual se tomó 22,5 grados para arriba y 22,5 grados para abajo, ya que las alturas de las salas permitían esta variación sin tener contacto directo con el techo o el piso al reventar el globo.



Figura 16. Identificación de puntos para la generación de impulsos en la sala Aranjuez.

Para establecer la distancia y la elevación se realizaron cálculos matemáticos con funciones trigonométricas, partiendo de los datos antes mencionados. La figura 17 muestra el esquema del triángulo rectángulo en el cual se basan los cálculos.

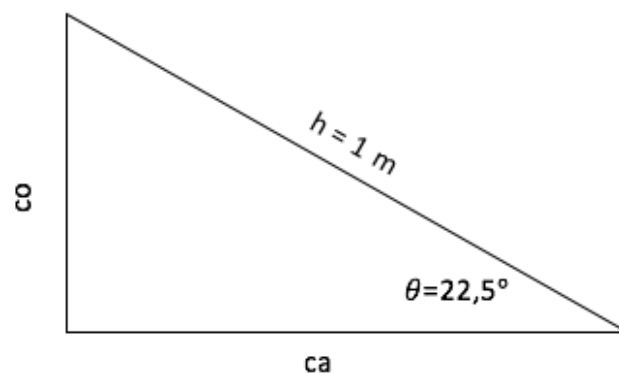


Figura 17. Triángulo equivalente para el cálculo de la altura de los impulsos.

Partiendo como altura base en 1,61 m, lo cual corresponde a la altura del oído a aplicarse en este estudio, se añade o se resta el valor de “co” que es calculado en base al ángulo θ . La variable “ca” corresponde a la distancia calculada en base al mismo ángulo. En la Tabla 3 se presentan las coordenadas en las cuales se generarán los impulsos.

Tabla 3.

Coordenadas en los 3 términos de los puntos para la generación de impulsos.

Azimut (grados)	Distancia (cm)	Elevación (cm)
90	184,77	237,6
45	184,77	237,6
0	184,77	237,6
-45	184,77	237,6
-90	184,77	237,6
90	92,38	199,3
45	92,38	199,3
0	92,38	199,3
-45	92,38	199,3
-90	92,38	199,3
90	92,38	122,7
45	92,38	122,7
0	92,38	122,7
-45	92,38	122,7
-90	92,38	122,7
90	184,77	84,4
45	184,77	84,4
0	184,77	84,4
-45	184,77	84,4
-90	184,77	84,4

3.4.2.3 Grabación in situ

Luego de colocar la cabeza en el lugar correspondiente, se procedió a reventar los globos en cada punto marcado. El nivel aproximado del impulso fue de 80 dB.



Figura 18. Generación de impulsos mediante la explosión de globos, en la sala Queri.

Para la grabación se aplicó un nivel óptimo de entrada, el cual fue aplicado en todas las salas. Es importante que este nivel sea correcto para que no exista saturaciones al momento de grabar la respuesta impulsiva.



Figura 19. Perillas de ganancia 1 y 2 para la grabación de canales L y R del dummy head .

La sesión se estableció con una frecuencia de muestreo de 96000 Hz, la cual asegura una mejor fidelidad en la grabación binaural especialmente en alta frecuencia. Se usó una profundidad de 24 bits, para tener mayor resolución de captura del audio. Para cada una de las salas se realizó el mismo procedimiento.

3.5 Selección de salas

Para realizar la evaluación psicoacústica se seleccionó tres salas:

- CR3: 0,36 s de TR60 global.
- Queri: 1,16 s de TR60 global.
- Aranjuez: 2,74 de TR60 global.

La elección se hizo con base en el tiempo de reverberación obteniendo, se tiene una sala seca (CR3), media reverberante (Queri) y muy reverberante (Aranjuez). Esta diferencia de tiempos de reverberación permite analizar en qué entornos la ubicación de fuente se ve favorecida o desfavorecida.

La sala Magic fue descartada debido a su asimetría acústica; en un lado se tenía paneles difusores que, a pesar de ser cubiertos con paneles móviles absorbentes, crearon mayor reverberación de un lado y por tanto la respuesta impulsiva se vio afectada.

3.6 Simulaciones en Sketch Up de las salas escogidas

Para poder orientar a los encuestados y darles una idea de cómo son cada una de las salas se crearon varios archivos de los recintos seleccionados en Sketch Up. Se realizó una aproximación de las medidas reales y también de la geometría de las salas, simulando un rectángulo.

Se colocó un ser humano de la altura promedio del ecuatoriano 1,67 m y se ubicaron los puntos donde se explotaron los globos. Para facilidad del encuestado y el encuestador se utilizó un código de colores y números para cada uno de los puntos y también un código para cada una de las salas. Esta codificación se presenta en la Tabla 4 y Tabla 5.

Tabla 4.

Codificación de puntos e identificación por colores.

Codificación de puntos					
Color	Número	Código	Azimut (grados)	Distancia (cm)	Elevación (cm)
Rojo (R)	1	<i>R1</i>	90	184,77	237,60
	2	<i>R2</i>	45	184,77	237,60
	3	<i>R3</i>	0	184,77	237,60
	4	<i>R4</i>	-45	184,77	237,60
	5	<i>R5</i>	-90	184,77	237,60
Verde (G)	1	<i>G1</i>	90	92,38	199,30
	2	<i>G2</i>	45	92,38	199,30
	3	<i>G3</i>	0	92,38	199,30
	4	<i>G4</i>	-45	92,38	199,30
	5	<i>G5</i>	-90	92,38	199,30
Azul (B)⁶	1	<i>B1</i>	90	92,38	122,70
	2	<i>B2</i>	45	92,38	122,70
	3	<i>B3</i>	0	92,38	122,70
	4	<i>B4</i>	-45	92,38	122,70
	5	<i>B5</i>	-90	92,38	122,70
Amarillo (A)	1	<i>Y1</i>	90	184,77	84,40
	2	<i>Y2</i>	45	184,77	84,40
	3	<i>Y3</i>	0	184,77	84,40
	4	<i>Y4</i>	-45	184,77	84,40
	5	<i>Y5</i>	-90	184,77	84,40

⁶ Para el color azul se usa la inicial "B" por su traducción en inglés "Blue".

Tabla 5.

Codificación de salas

Codificación de salas	
Nombre de Sala	Código
<i>Queri</i>	Q
<i>Aranjuez</i>	A
<i>CR3</i>	C

Se estableció un código de colores en inglés que distinga diferente elevación y distancia, luego se enumeró según el azimut comenzando de derecha hasta izquierda; a esto se le añadió la inicial de los nombres de cada una de las salas.

Por ejemplo, para referir el punto dentro del eje azul que está a 90 grados a la derecha se utilizará el código "B1". A partir de ahora se utilizará este código para referirnos a los puntos a evaluar en cada sala.

A continuación, se muestra una imagen de cada una de las salas simuladas.

3.6.1 Sala 1: Sala comunal - Edificio Aranjuez

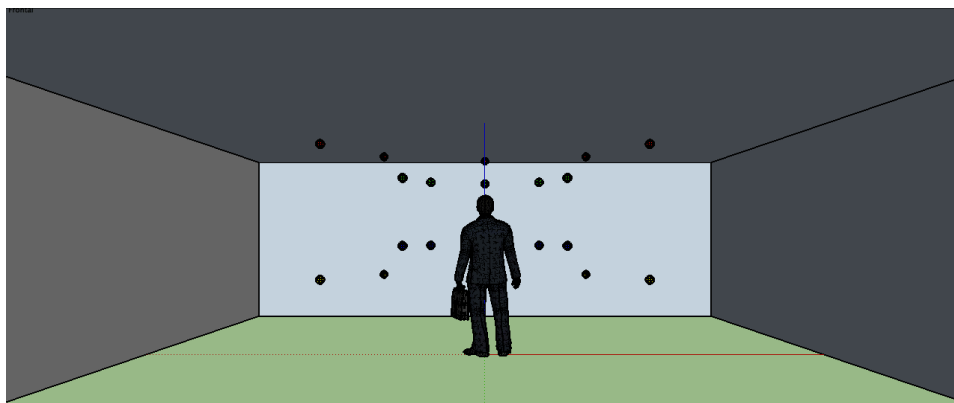


Figura 20. Sala equivalente a la sala comunal del edificio Aranjuez.

3.6.2 Sala 3: Auditorio Queri - Universidad de las Américas

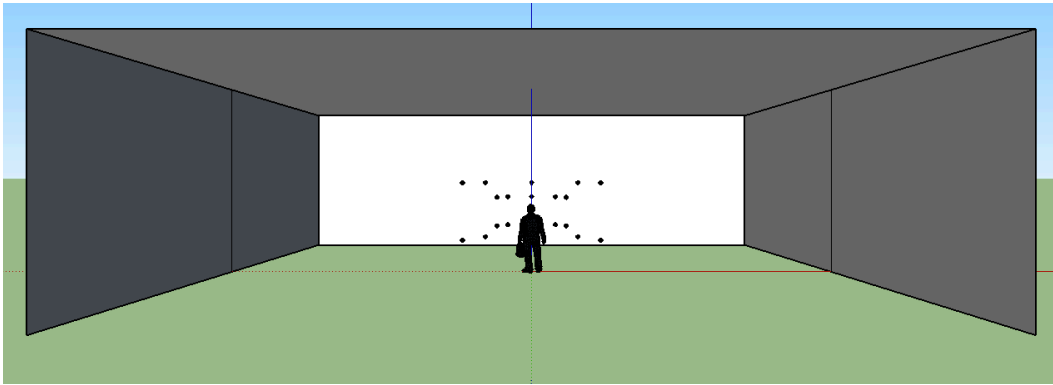


Figura 21. Sala equivalente al auditorio Queri de la Universidad de las Américas.

3.6.3 Sala 4: Estudio de grabación - CR3

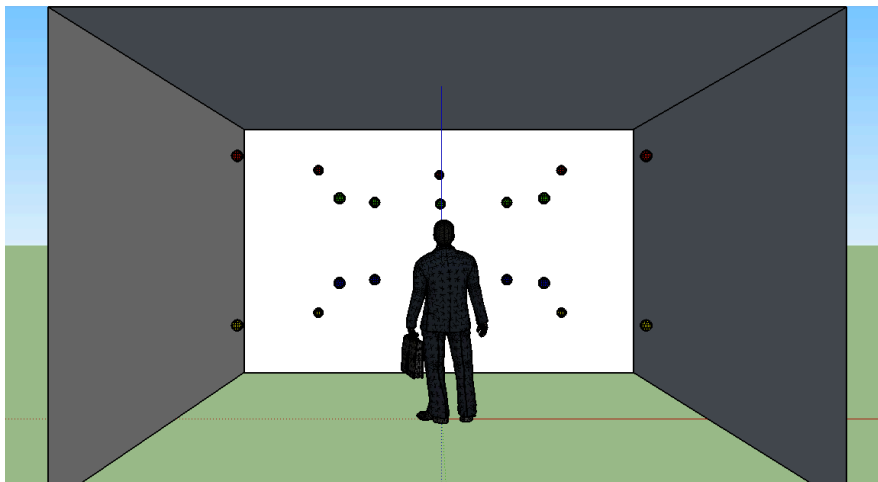


Figura 22. Sala equivalente al estudio de grabación CR3.

3.7 Procesamiento de señales

3.7.1 Instrumentación

- Software MATLAB 2013
- Computadora portátil MacBook Pro

3.7.2 Selección de estímulos

Para la selección de los audios a convolucionar se tomó en cuenta los resultados de De la Torre (2016, p. 58), en donde la música y el ruido obtuvieron mayor porcentaje de aciertos.

Se seleccionó un fragmento de la canción “We are the Champions” de Queen que, según estudios de la Universidad de Londres, Universidad de New York y la Universidad de Goldsmith, es considerada una de las canciones más pegadizas de la historia (Europa Press, 2011). Como segundo estímulo sonoro se escogió ruido rosa debido a que la densidad espectral es inversamente proporcional a la frecuencia (Domingo, p. 10), por lo que es constante por banda de tercio de octava. En la Figura 25 se observa la respuesta en frecuencia de ambos estímulos.

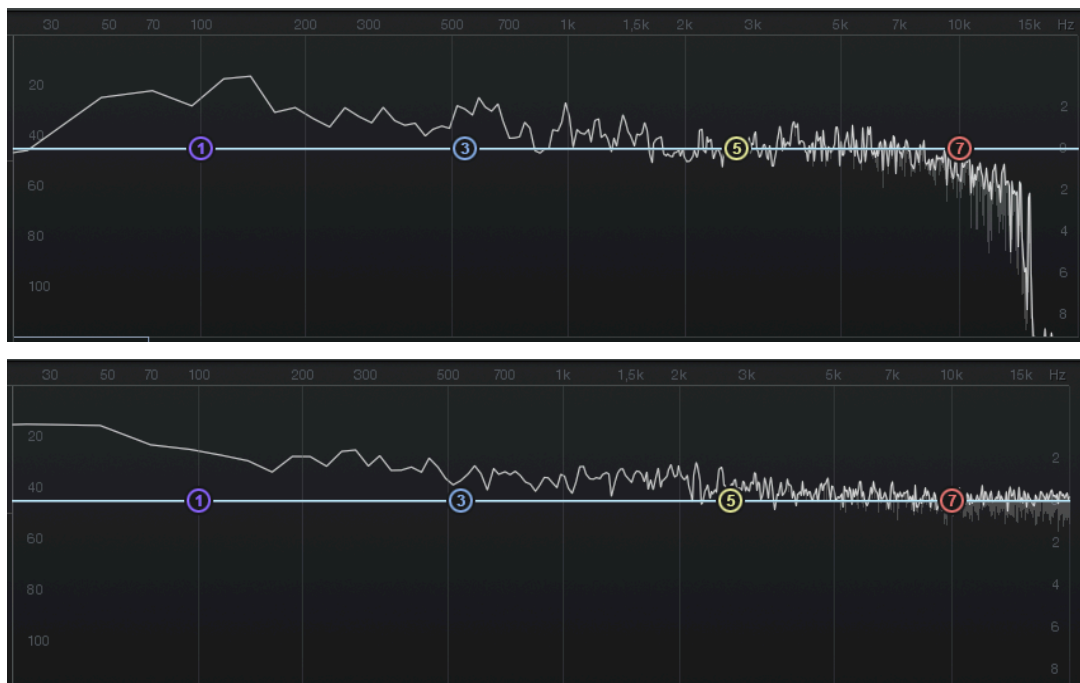


Figura 23. Respuesta en frecuencia de la canción “We are the Champions” (arriba) y del ruido rosa (abajo) tomada a través del plug in Ozone 6.

3.7.3 Corrección de nivel del dummy head

3.7.3.1 Instrumentación

- Dummy head: Neumann KU 100
- Computadora portátil MacBook Pro
- Interfaz PreSonus Audiobox - 1818VSL
- Software Pro Tools 10
- Altavoz
- Flexómetro y metro

3.7.3.2 Metodología y resultados

Al realizar la normalización en MatLab se observó que los valores de ganancia eran siempre mayores en el lado derecho independientemente de la localización de la fuente de procedencia, por lo cual se deduce que hay una descompensación en el canal izquierdo. Debido a esto, se realizó una prueba con ruido rosa, puesto que tiene la misma energía en bandas de octava.

Se colocó el altavoz a nivel de la oreja del dummy head y a 1 m de la misma. Se realizaron dos grabaciones de 10 s:

- Oreja izquierda apuntando al altavoz.
- Oreja derecha apuntando al altavoz.

Tabla 6.

Promedio logarítmico de las diferencias entre canales L y R.

	<i>Oreja izquierda</i>		<i>Oreja derecha</i>		ΔL	dBfs
Toma 1	Canal L	-21,7	Canal R	-16,2	5,5	
	Canal R	-21	Canal L	-25,7	4,7	
Toma 2	Canal L	-23	Canal R	-17,5	5,5	
	Canal R	-22,4	Canal L	-27	4,6	
Promedio					5,1	

En la Figura 24 se observa la respuesta en frecuencia de ambos canales. Al comparar las señales en cada canal del dummy head no se evidencia un espectro diferente, por tanto se descarta que el problema sea a nivel de frecuencia.



Figura 24. Rango en frecuencia de canal izquierdo (arriba) y canal derecho (abajo), tomada con el plug in Ozone 6.

Dado que un dummy head simula el comportamiento del oído humano es de esperarse que la captación sea igual en ambos micrófonos. Se observa que el micrófono correspondiente a la oreja izquierda tiene una descompensación de 5,1 dBfs con respecto al micrófono derecho.

3.7.4 Proceso de convolución

El proceso de convolución se realizó en el software Matlab a través de un código que permite la combinación entre un estímulo sonoro y un impulso, para la recreación de la reverberación de la sala específica.

3.7.4.1 Entrada

```
%Entrada y lectura de señales
fs = 96000; %Frecuencia de muestreo

%Señal monofónica
[senal, fs_senal] = audioread ('sonido.wav');

%Señal binaural
[impulso, fs_impulso] = audioread ('impulso.wav');
impulsoL = impulso (:,1); %Canal izquierdo
impulsoR = impulso (:,2); %Canal derecho
```

Figura 25. Código de entrada de señales para procesamiento.

La información correspondiente al audio se lee mediante el comando “*audioread*” y se almacena en arreglos. En el caso de la señal binaural se dispone de dos arreglos ya que se tendrán los canales izquierdo y derecho.

3.7.4.2 Procesamiento

```

%Proceso de convolución

%Canal izquierdo
%Longitud de la señal convolucionada
Salida_L = length(senal) + length(impulsoL)-1;
Salida_L2 = pow2(nextpow2(Salida_L));

%Transformación de Fourier de la señal de entrada
SENALFFTL = fft(senal, Salida_L2);

%Transformación de Fourier del impulso
IMPULSOL = fft(impulsoL, Salida_L2);

%Convolución entre la señal de entrada y la respuesta impulsiva
REVERBERACION1 = SENALFFTL .* IMPULSOL;

%Transformada inversa de Fourier del resultado de convolución
reverberacion1 = real(ifft(REVERBERACION1, Salida_L2));

```

Figura 26. Código de convolución de canal izquierdo.

```

%Canal derecho
%Longitud de la señal convolucionada
Salida_R = length(senal) + length(impulsoR)-1;
Salida_R2 = pow2(nextpow2(Salida_R));

%Transformación de Fourier de la señal de entrada
SENALFFTR = fft(senal, Salida_R2);

%Transformación de Fourier del impulso
IMPULSOR = fft(impulsoR, Salida_R2);

%Convolución entre la señal de entrada y la respuesta impulsiva
REVERBERACION2 = SENALFFTR .* IMPULSOR;

%Transformada inversa de Fourier del resultado de convolución
reverberacion2 = real(ifft(REVERBERACION2, Salida_R2));

```

Figura 27. Código de convolución de canal izquierdo.

Se establece la longitud que tendrá la señal resultante del proceso de convolución. La función “*pow2*” permite que únicamente se consideren las partes

reales de la variable de entrada y “*nextpow2*” es una función que se utiliza para encontrar el exponente más pequeño de la potencia en base dos dentro de los valores reales. Tal como afirma Carrizo (2014, p.2), para el proceso de convolución es necesario tener las señales dentro del dominio de frecuencia, lo cual se logra a través de la transformada de Fourier. En MatLab se aplica el comando “*fft*” que permite directamente obtener esta transformada tanto en la señal monofónica como en los canales L y R del impulso binaural.

Una vez obtenidas las transformadas se realiza la operación producto punto entre las mismas para obtener la señal resultante. Finalmente, se aplica la transformada inversa mediante el comando “*ifft*” que retornará la señal convolucionada al dominio del tiempo. Este proceso se aplica en ambos canales como se puede apreciar en las imágenes anteriores.

3.7.4.3 Salida

```
%Salida

%Obtención valores máximos para división en normalización
maxrev1=max(abs(reverberacion1));
maxrev2=max(abs(reverberacion2));

%Normalización
reverberacion1 = reverberacion1/24;
reverberacion2 = reverberacion2/24;

%Unión de ambos canales
ReverberacionEstereo(:,1) = reverberacion1;
ReverberacionEstereo(:,2) = reverberacion2;

%Escritura
audiowrite(ReverberacionEstereo,96000,24,'AudioConvolucionado.wav');
```

Figura 28. Código de normalización, ensamblaje y escritura del audio convolucionado.

Antes de realizar la escritura del archivo de audio se normaliza los valores para evitar las saturaciones de los archivos. Para este proceso se ha extraído los máximos absolutos mediante el comando “*max*”. En el Anexo 5 observa las tablas con los valores correspondientes. El valor más elevado es de 23,71, por lo que se ha determinado que el valor para la división sea 24 tanto para el canal derecho como izquierdo. Esta corrección se vuelve importante ya que la amplitud de salida de la señal puede distorsionar la percepción en el término de distancia. La división por el mismo término, tanto en los canales L y R, asegura que se mantenga la misma diferencia y, por tanto, no se creen errores en las encuestas. Finalmente se realiza la unión de ambos canales en la variable “*ReverberacionEstereo*” y se procede a la escritura a través del comando “*audiowrite*”.

3.8 Pruebas psicoacústicas

3.8.1 Selección de audios convolucionados

Por cada sala se grabaron 20 impulsos y se tienen 3 salas, dando un total de 60 impulsos, adicionalmente, se tienen dos sonidos con los que se realizará la convolución dando un resultado final de 120 audios. Únicamente se seleccionaron 10 puntos por sala tanto con la frase musical como con el ruido rosa, dando como resultado final 60 sonidos convolucionados para utilizarse en las pruebas.

Para la elección de los puntos, se tomó en cuenta que la tesis está enfocada en comparar los resultados según el nivel de reverberación de cada sala, el tipo de estímulo y las diferencias de percepción en cada eje por individual. A continuación, se observa gráficamente los puntos seleccionados.

Tabla 7.

Puntos seleccionados para el proceso de encuesta.

Puntos seleccionados
R1
Y2
R3
Y4
R5
B1
G2
B3
G4
B5

Para un mejor entendimiento se muestran capturas de pantalla de los puntos simulados en Sketch Up, con diferentes vistas.

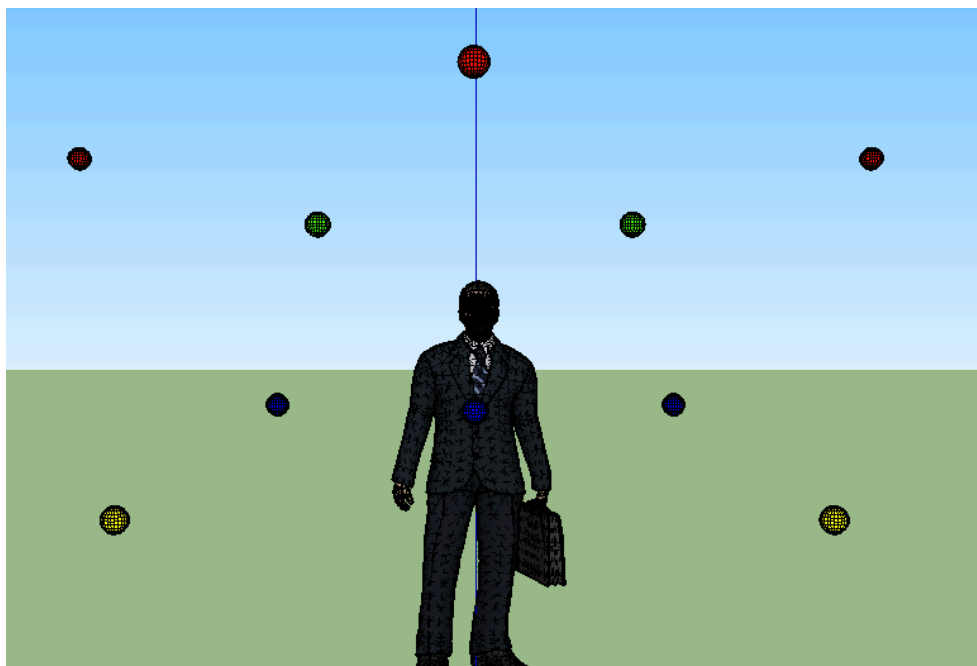


Figura 29. Vista frontal de los puntos seleccionados.

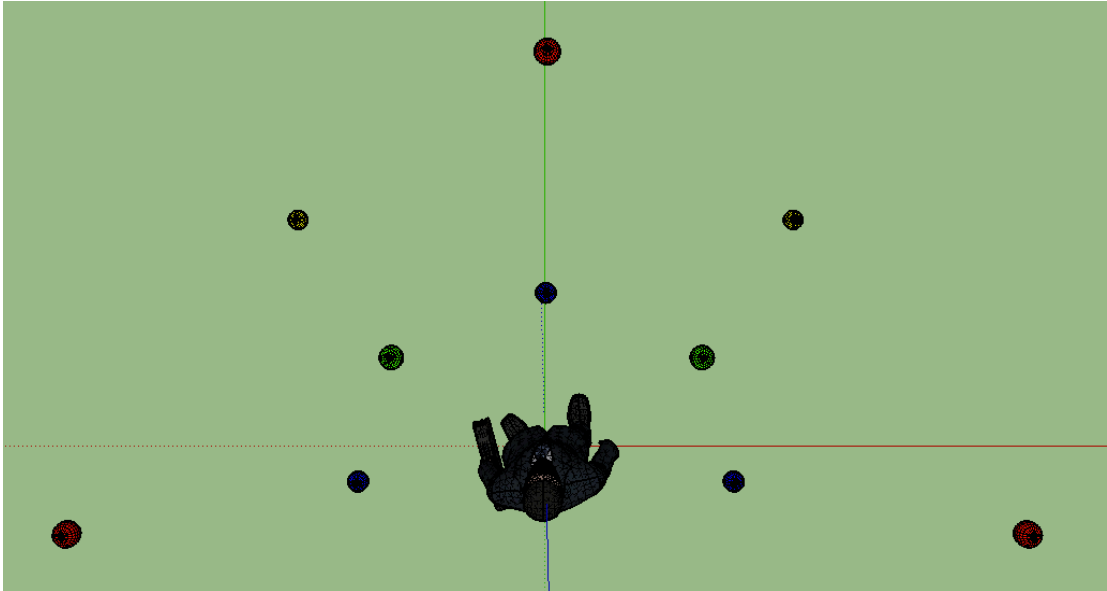


Figura 30. Vista desde arriba de los puntos seleccionados.

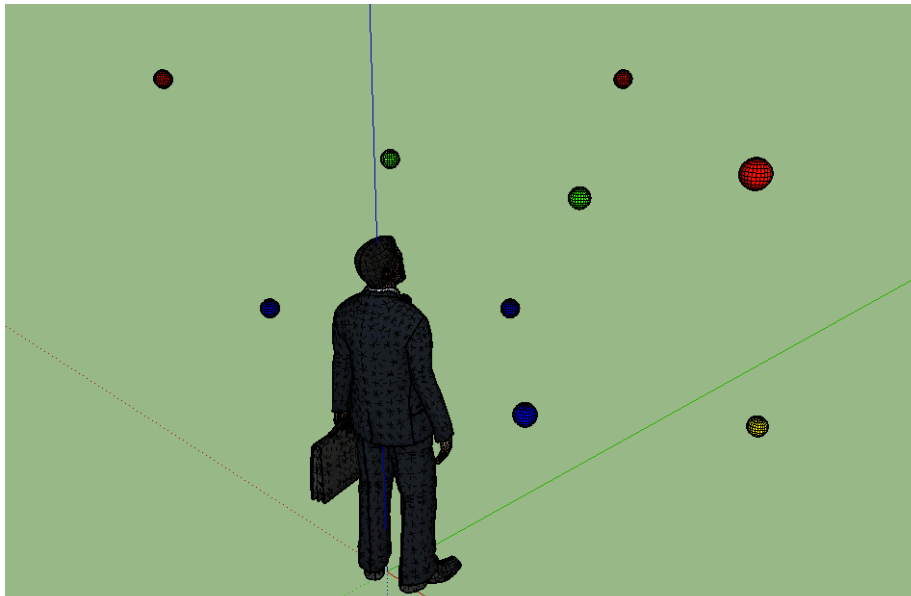


Figura 31. Vista isométrica de los puntos seleccionados.

3.8.2 Instrumentación

- Computadora portátil MacBook Pro
- Sennheiser HD 280 Pro⁷

3.8.3 Cálculo de la población

Para este ensayo se va a considerar como, población total a todos los estudiantes de la carrera de Ingeniería en Sonido y Acústica de la Universidad de las Américas que hayan cursado la materia de entrenamiento auditivo. El cálculo para la muestra poblacional se lo realiza mediante la siguiente ecuación:

$$n = \frac{(Z_{\frac{\alpha}{2}})^2 Npq}{NE_p^2 + (Z_{\frac{\alpha}{2}})^2}$$

Ecuación 1. Fórmula correspondiente al método de estimación de proporción poblacional.

Tomada de Galindo, 2008, p. 387.

Donde:

- N corresponde al total de la población seleccionada.
- P y q son valores establecidos en 0,5 que determinan el margen de la muestra.
- $Z_{\frac{\alpha}{2}}$ es el coeficiente normal estándar que, para este caso, tiene un valor de 2 debido a que se trabaja con una confianza de 95.5%.
- E_p es el error estadístico que se ha establecido en 0,1.

⁷ Las especificaciones técnicas de los audífonos se encuentran en el Anexo 6.

El cálculo se lo realiza en un principio con 183 estudiantes dentro de la carrera, obteniéndose una muestra de 30 personas. Considerando que el número de mujeres no es el suficiente (se necesita que sea el 50% de la muestra total) y que para eliminar la incertidumbre es preferible considerar únicamente a los estudiantes que hayan pasado la materia de entrenamiento auditivo, se reduce el número de población a 53 personas.

$$n = \frac{(1,96)^2 \cdot 53 \cdot 0,5 \cdot 0,5}{53 \cdot (0,1075)^2 + (1,96)^2}$$

$$n = 11,4$$

$$n \approx 12$$

El resultado final de muestra poblacional es de 11,4; es decir que para tener igual número de hombres y mujeres se consideraron 12 encuestas.

3.8.4 Evaluación psicoacústica

El ensayo de la evaluación se realizó en la sala EG1 debido a que la misma simula a una cámara anecoica, es decir, que sus tiempos de reverberación son bajos y no influirán en los sonidos convolucionados.

Como primer paso se realizó una audiometría para evaluar la salud auditiva de cada encuestado. Posteriormente se reprodujo los audios sin convolución, de ruido y música, para establecer un volumen que será el mismo para cada sala y para cada estímulo, puesto que el nivel de presión sonora varía con la distancia y geometría de las salas. Debido a que los encuestados son personas que conocen el tema se les explicó que deben localizar de donde proviene el sonido en diferentes salas que tienen su propia reverberación. Se comenzó a reproducir los audios convolucionados por salas primero un estímulo y luego otro. El orden de reproducción de los audios fue aleatorio para cada prueba. En promedio la prueba tomó un total de tiempo 20 min a 30 min dependiendo de la persona.



Figura 32. Fotografías de encuestados realizando el ejercicio.

4 Resultados y análisis

4.1 Audiometrías

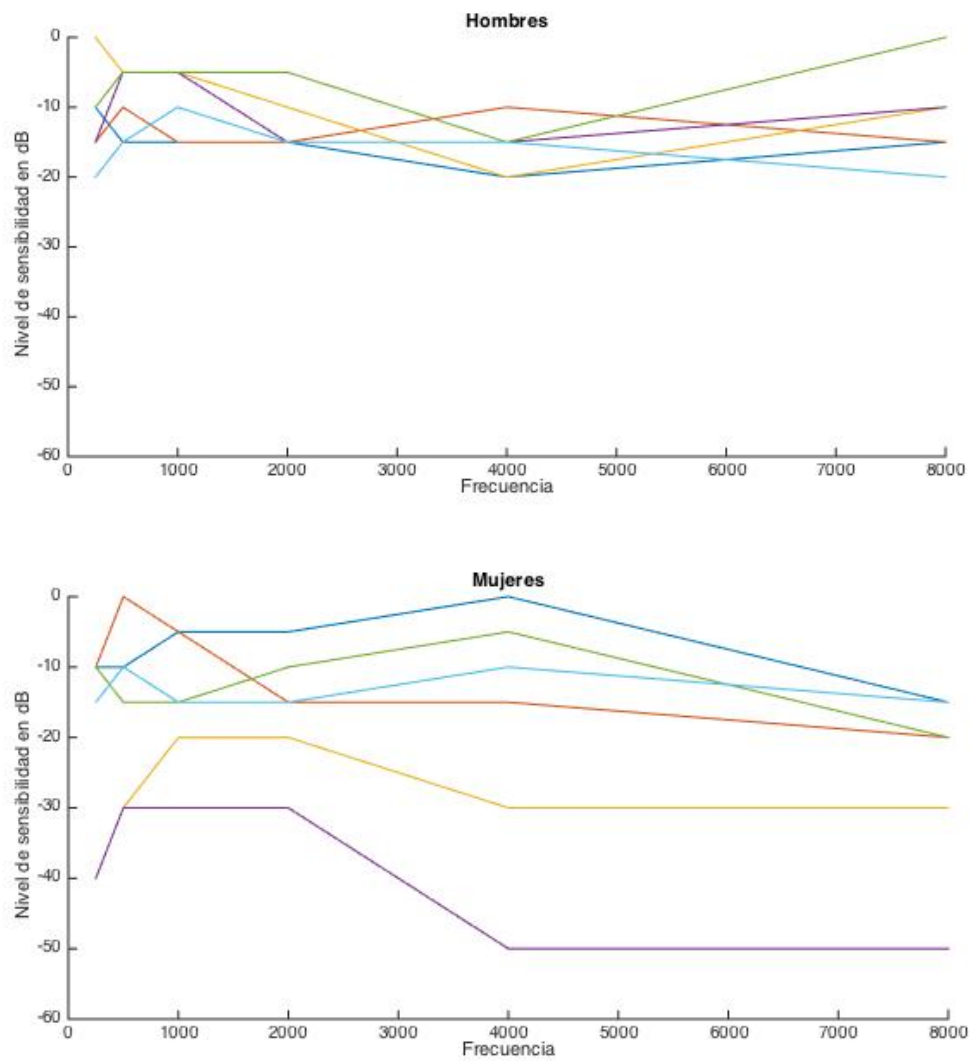


Figura 33. Gráficos correspondientes a las audiometrías realizadas a los participantes de la encuesta.

En la Figura 33 se observa como, en las audiometrías, los hombres tienen un comportamiento similar a diferencia de las mujeres en donde se nota dos casos en donde la audición se encuentra fuera del rango normal.

4.2 Aciertos en los tres términos



Figura 34. Porcentaje de aciertos totales considerando todos los factores de estudio.

En la Figura 34 se consideró ambos sexos, todas las salas, todos los estímulos y los puntos seleccionados. Se observa que los aciertos de la evaluación psicoacústica, en los tres términos, solo representan el 20% del total de la población que es un porcentaje bajo.

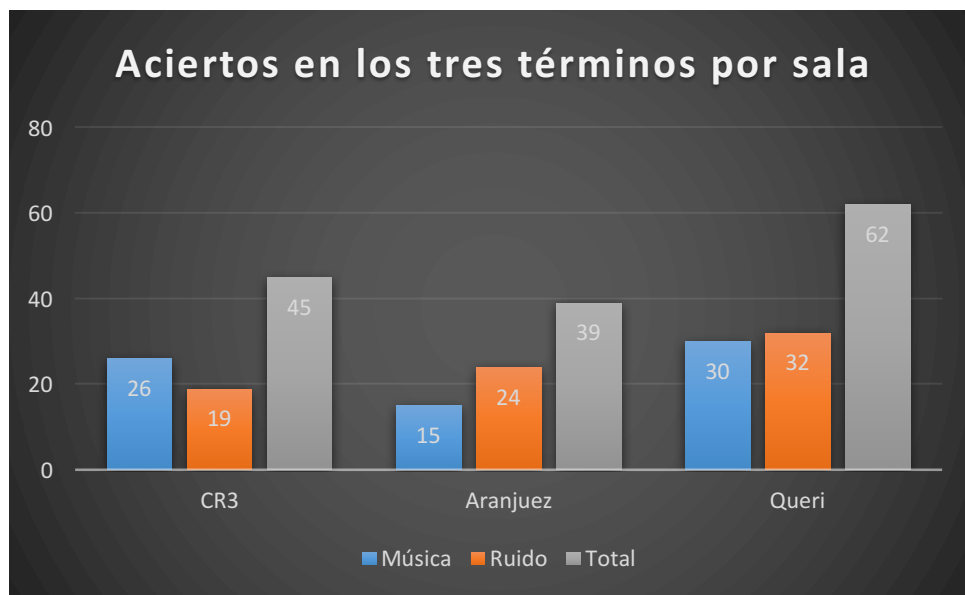


Figura 35. Número de aciertos en los tres términos divididos por salas.

En el caso de los aciertos por sala en los tres términos se tiene un mayor número de aciertos de los dos estímulos en la sala Queri, sala que tiene un tiempo de reverberación medio. En segundo lugar, está la sala CR3 y por último con menor número de aciertos la sala Aranjuez, que tiene mayor tiempo de reverberación.

Se observa que en la sala Queri los aciertos para música y ruido son casi iguales. En el CR3 se tiene mayores aciertos con música y en Aranjuez con ruido.

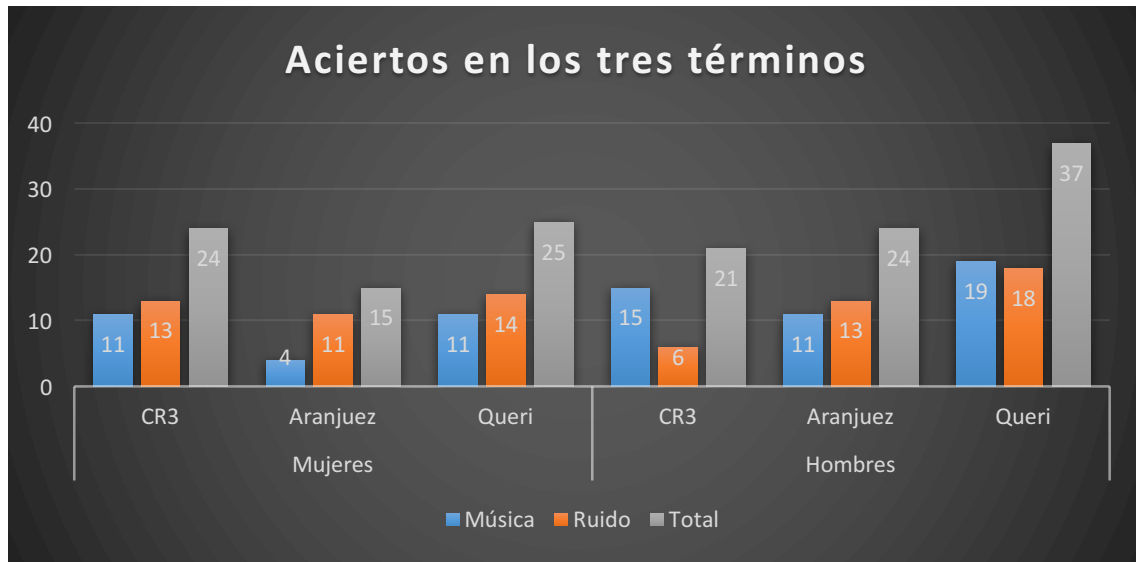


Figura 36. Número de aciertos en los tres términos divididos por sexos.

En la Figura 36 se puede ver una clasificación por sexo, por sala y por estímulos; además la suma de los últimos.

Con respecto al sexo, se observa que predomina el hombre en todos los estímulos y salas puesto que tiene mayor número de aciertos. En la mayoría de los casos predomina el ruido sobre la música, excepto en el caso de los hombres en la sala CR3, donde la mayoría de aciertos son de música. Otro caso a resaltar es el de las mujeres en Aranjuez, donde el número de aciertos en música es mínimo.

Para la población total, la sala con mayores aciertos es Queri y la de menor, en el caso de las mujeres, es Aranjuez y en el de los hombres CR3. Cabe aclarar que CR3 y Aranjuez tienen una diferencia mínima de aciertos en el sexo masculino.

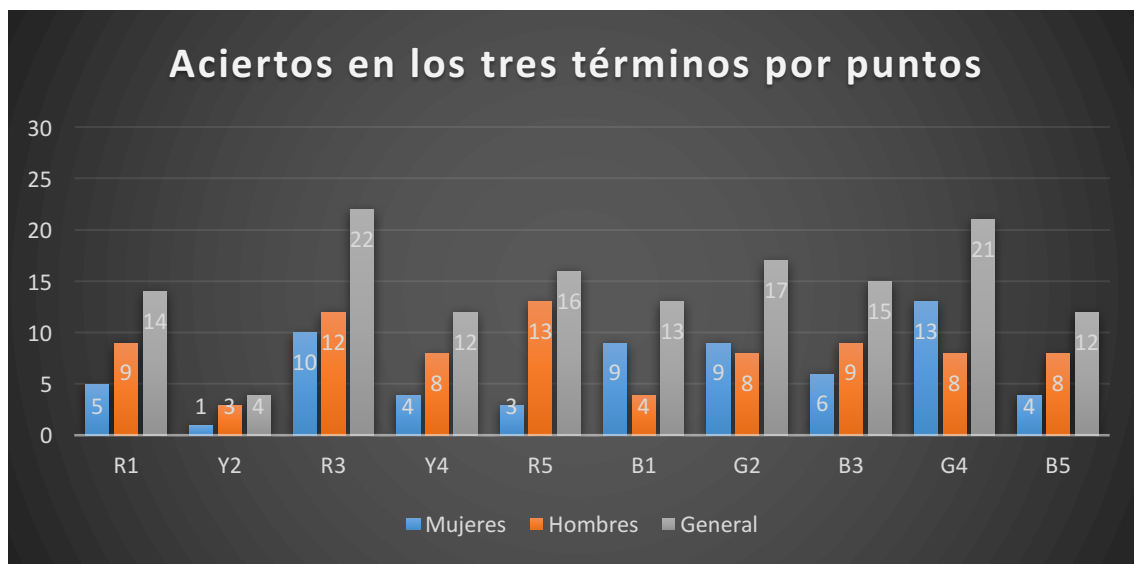


Figura 37. Número de aciertos en los tres términos divididos por puntos específicos.

Los puntos con mejores resultados, sin importar el sexo, son R3 y G4 mientras que punto con menor número de aciertos es Y2. Al clasificar por sexo, las mujeres tienen mejor resultado en el punto G4 y los hombres en el punto R5. Los peores resultados en hombres son en el punto B1 y las mujeres en el R5.

Los resultados de las mujeres en la mayoría de los casos son mucho menores al los resultados de los hombres.

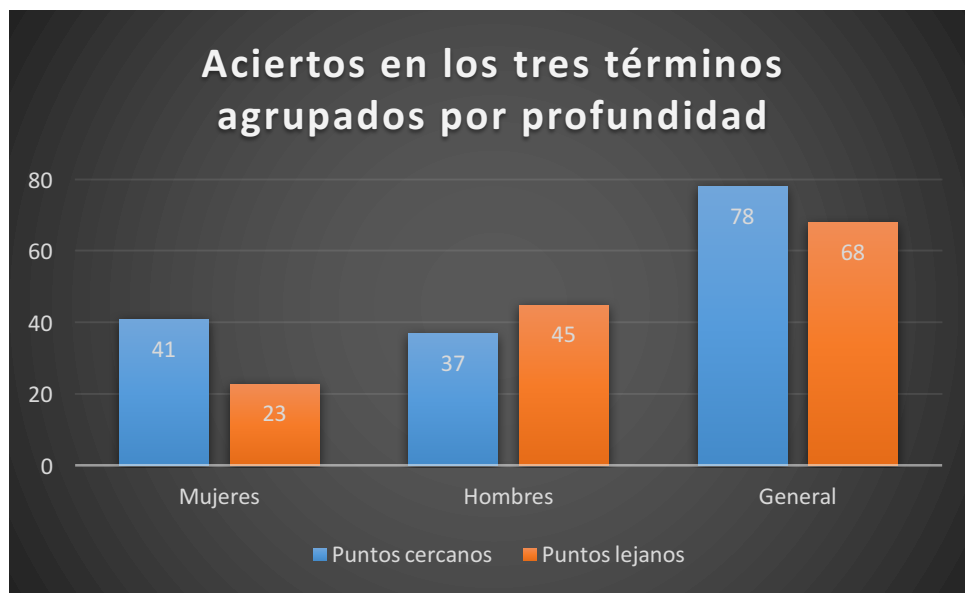


Figura 38. Número de aciertos en los tres términos divididos por puntos lejanos y cercanos.

Al realizar una partición por puntos cercanos y lejanos se observa que no existe una diferencia notable entre los mismos. La diferencia es de 10 aciertos, sin embargo, se puede observar que las mujeres tienden a acertar con los puntos cercanos más que con los lejanos mientras que con los hombres sucede de manera inversa, pero con poca diferencia.

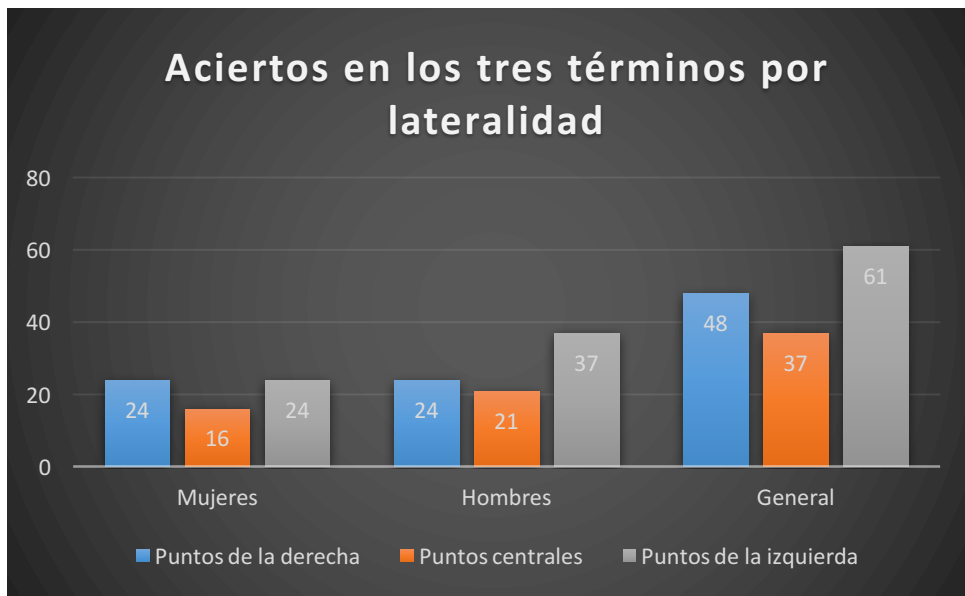


Figura 39. Número de aciertos en los tres términos divididos por puntos en la zona derecha e izquierda.

La Figura 39 divide los puntos según su posición en azimut; se tiene puntos a la izquierda, puntos centrales y puntos a la derecha. La cantidad de puntos centrales es menor a los de la derecha e izquierda.

Los puntos de la izquierda tienen mayores aciertos en comparación con las otras zonas en el caso de los hombres, mientras que para las mujeres se tiene la misma cantidad de aciertos para derecha e izquierda. De manera general hay más aciertos para el lado izquierdo, seguido del lado derecho y por último los puntos centrales.

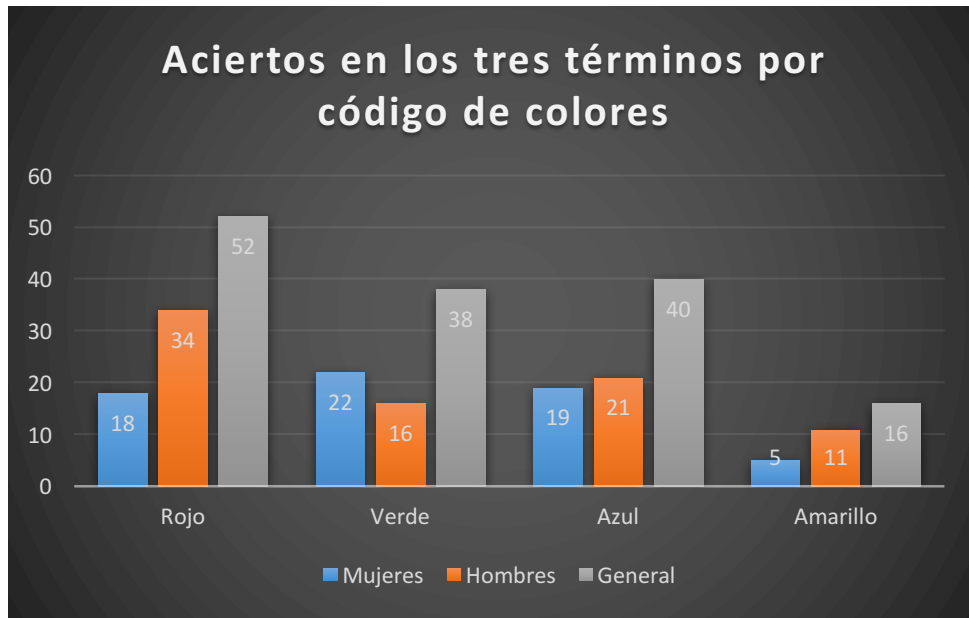


Figura 40. Número de aciertos en los tres términos divididos por la condificación de colores con respecto a los puntos.

La Figura 40 muestra una clasificación por colores de los puntos que se diferencian tanto alturas como distancias. Es importante señalar que tenemos más audios convolucionados en los ejes rojo y azul, a diferencia de los ejes verde y amarillo. En este caso los mayores aciertos se encuentran en los puntos rojos, los cuales tienen mayor altura y también distancia. Si se observa los resultados por sexo, los hombres tienen más aciertos en los rojos y las mujeres en los verdes.

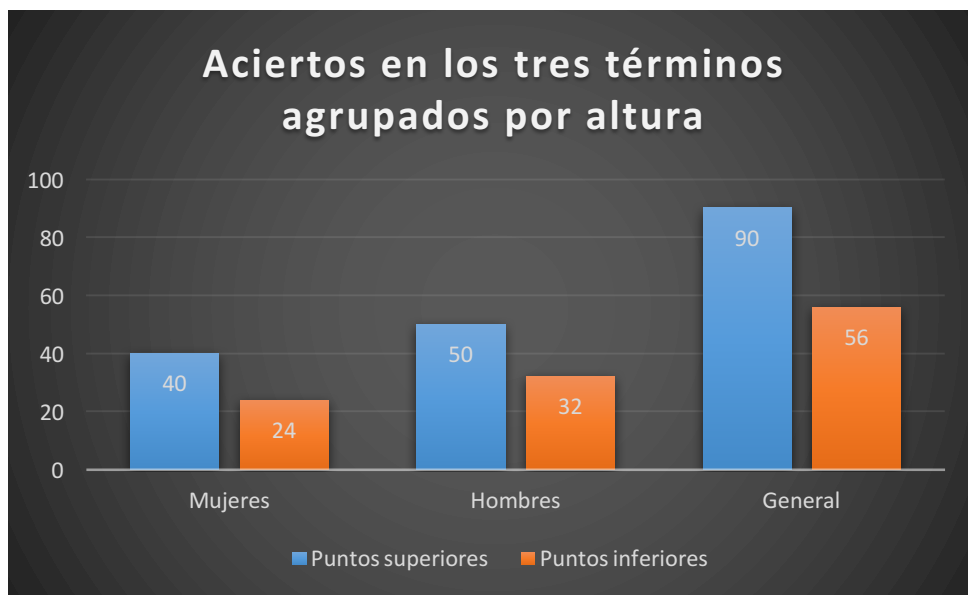


Figura 41. Número de aciertos en los tres términos divididos por sexo.

En la Figura 41 se puede analizar por sexo los aciertos tanto en los puntos inferiores como en los superiores. Se aprecia que los puntos superiores tienen más aciertos que los inferiores, esto tanto para hombres como para mujeres.

4.3 Aciertos en azimut

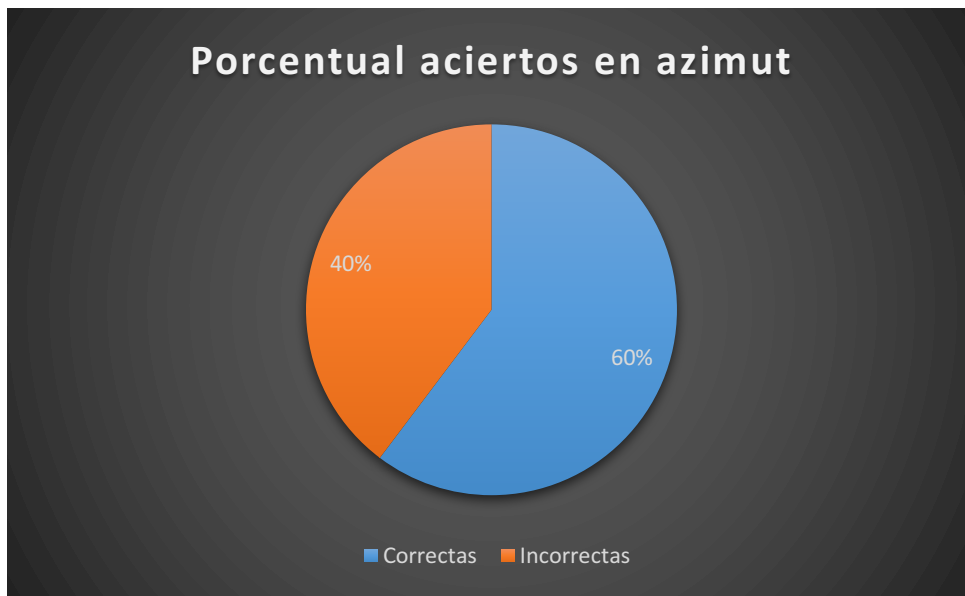


Figura 42. Porcentaje de aciertos globales en el término azimut.

En cuanto al término de azimut se obtuvo un porcentaje de 60% de aciertos, el cual es representativo.

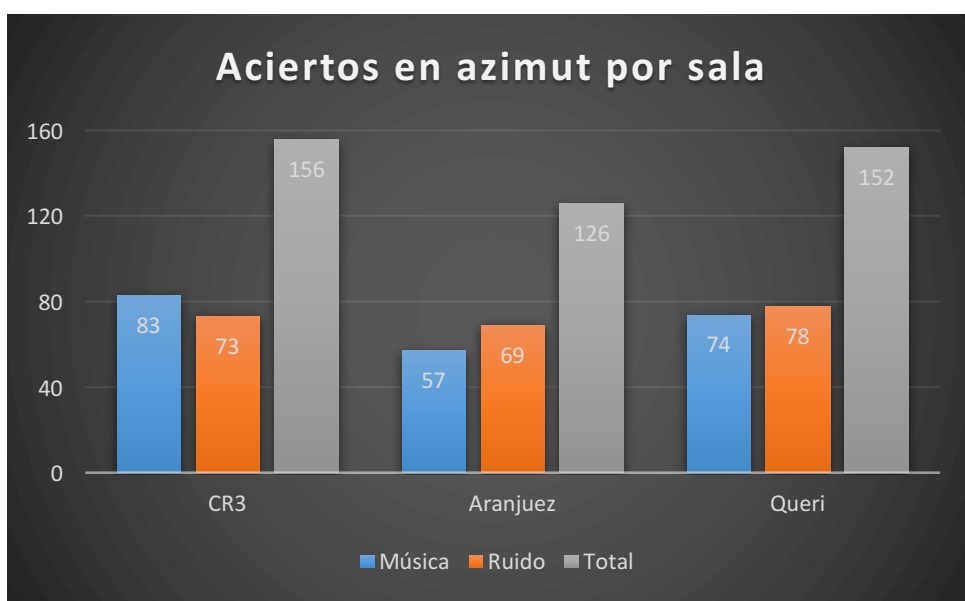


Figura 43. Porcentaje de aciertos en el término azimut divididos por salas.

En estos resultados generales por salas se observa que existe mayor dificultad para reconocer este término en Aranjuez, el cual tiene el tiempo de reverberación más alto. Con respecto a las otras dos salas, existe un reconocimiento similar tanto para el CR3 como para Queri. Se observa que no existe una tendencia sobre cual estímulo se reconoce mejor ya que tienen diferencias mínimas.

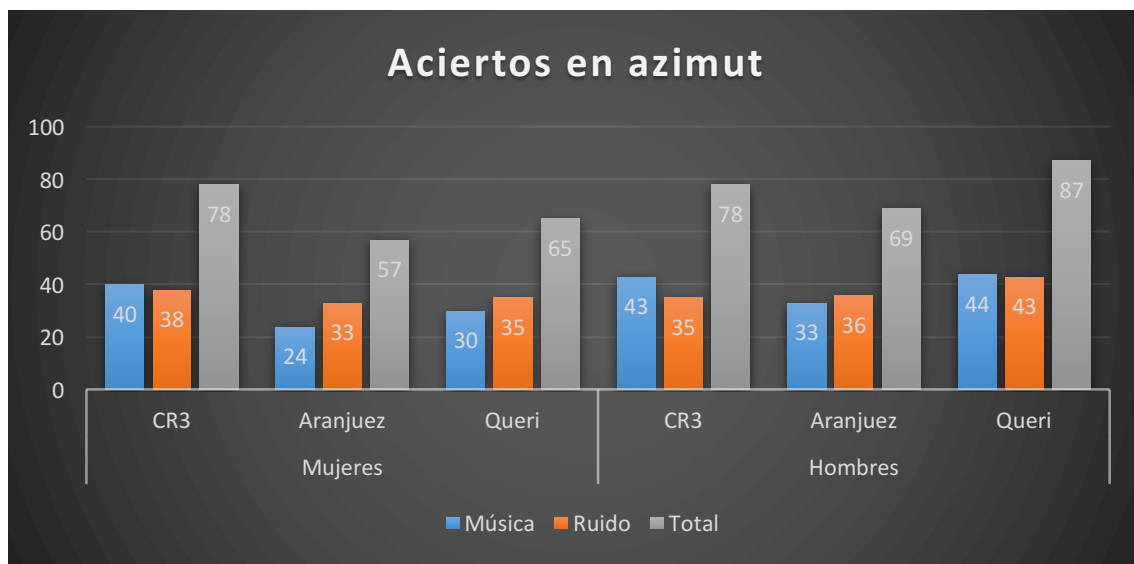


Figura 44. Porcentaje de aciertos en el término azimut divididos por sexo.

En la Figura 44 se observan los resultados por sexo tanto de hombres como de mujeres. La sala que más aciertos tuvo por parte de las mujeres es el CR3 mientras que los hombres poseen mayor nivel de aciertos en Queri. Se observa que la sala Aranjuez tiene menor número de aciertos para ambos sexos. En cuanto a los estímulos, como ya se mencionó, tienen similares aciertos tanto música como ruido.

El recinto CR3 tiene la misma cantidad de aciertos tanto para hombres como para mujeres y el nivel de acierto con música está por sobre el ruido.

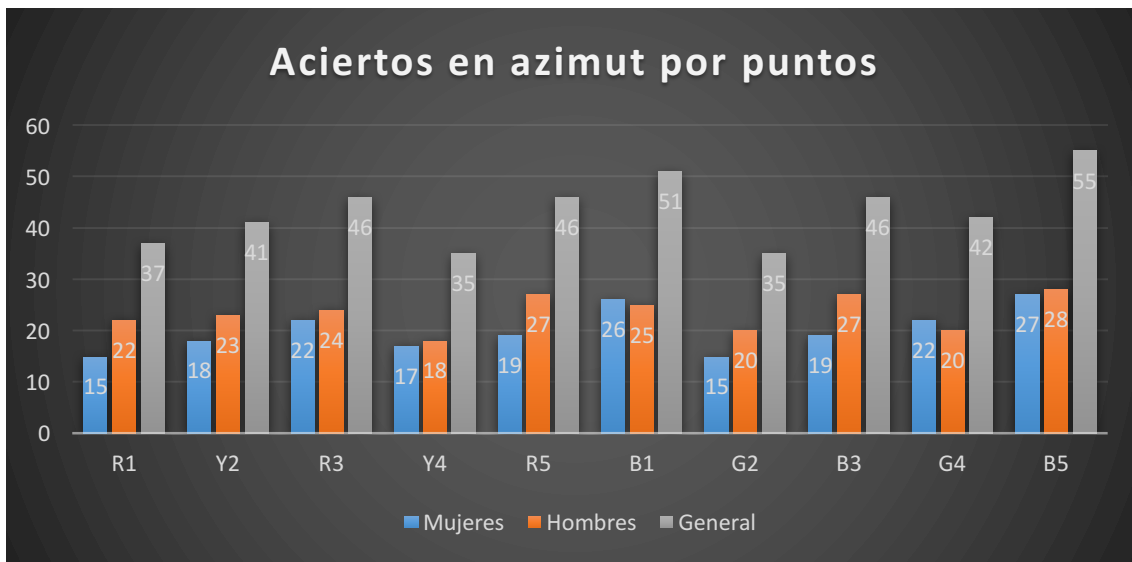


Figura 45. Porcentaje de aciertos en el término azimut divididos por puntos.

Los encuestados acertaron más en los puntos B1 y B5, mientras que acertaron menos a los puntos Y4 y G2. Los resultados de hombres y mujeres se asemejan mucho en todos los puntos.

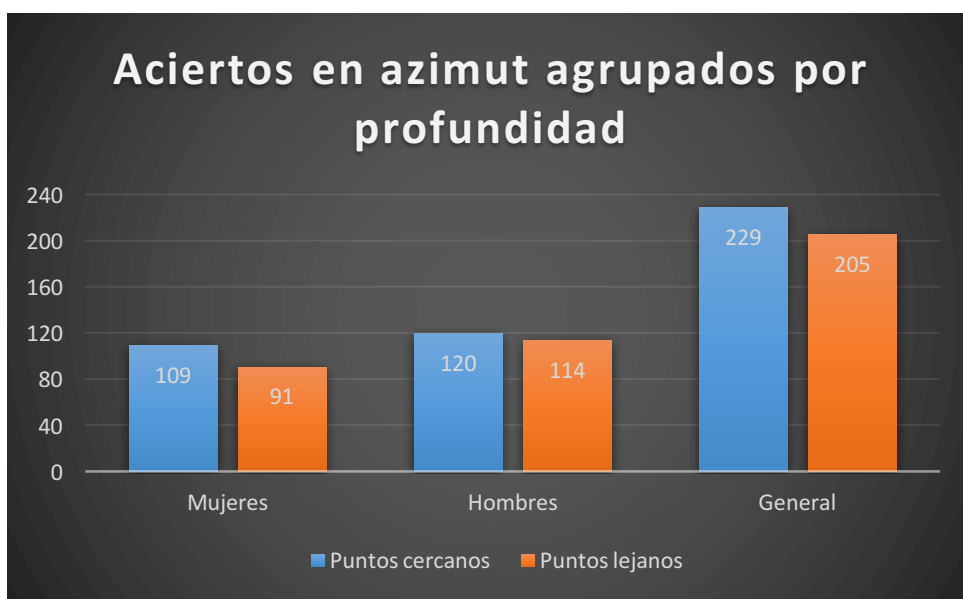


Figura 46. Porcentaje de aciertos en el término azimut divididos por género y distancia de puntos.

En la Figura 46 se observa que los puntos cercanos están sobre los puntos lejanos por 24 aciertos únicamente. Esta tendencia está marcada por las mujeres puesto que en que los puntos cercanos tienen mayor diferencia de aciertos con los lejanos.

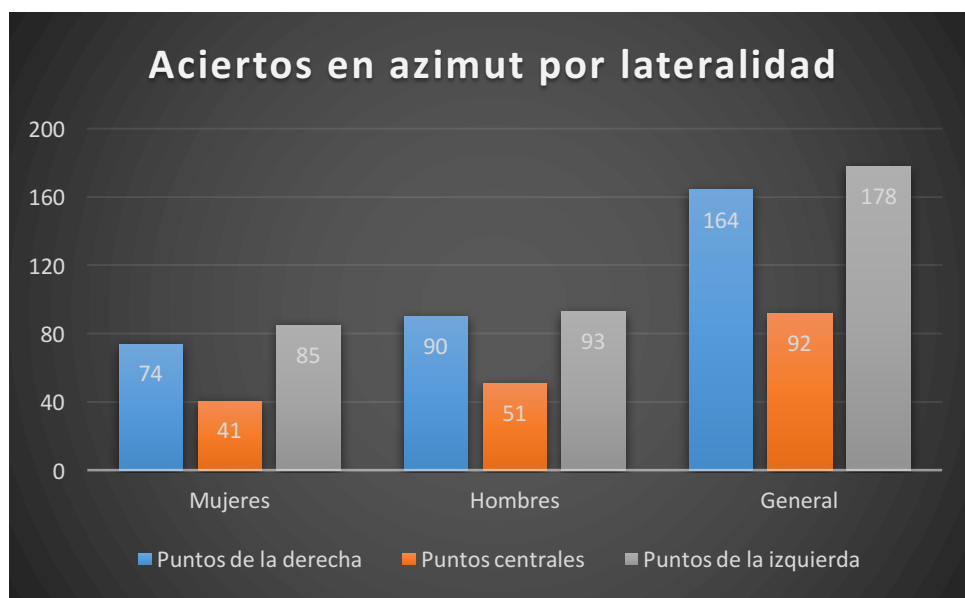


Figura 47. Porcentaje de aciertos en el término azimut divididos por sexo y zonas izquierda, central y derecha.

En el gráfico se observa que, para el término de azimut, son más reconocibles los puntos de la izquierda que los de la derecha independientemente del sexo de los participantes. Para hombres el resultado de aciertos de los puntos de izquierda y derecha son mayores.

4.4 Aciertos en elevación

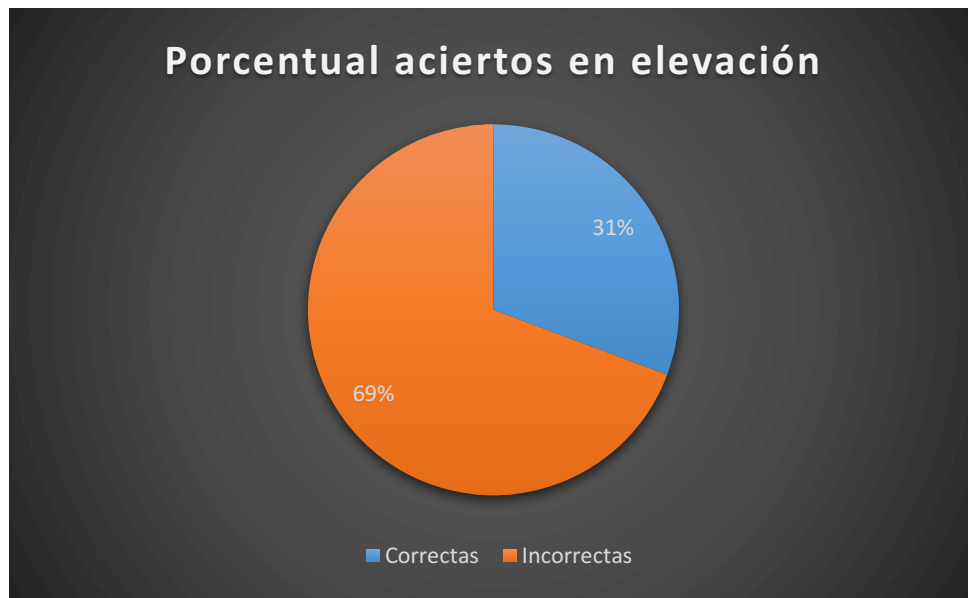


Figura 48. Porcentaje de aciertos globales en el término elevación.

Para la elevación se observa un porcentaje de aciertos del 31% que no representa la mayoría de la población.

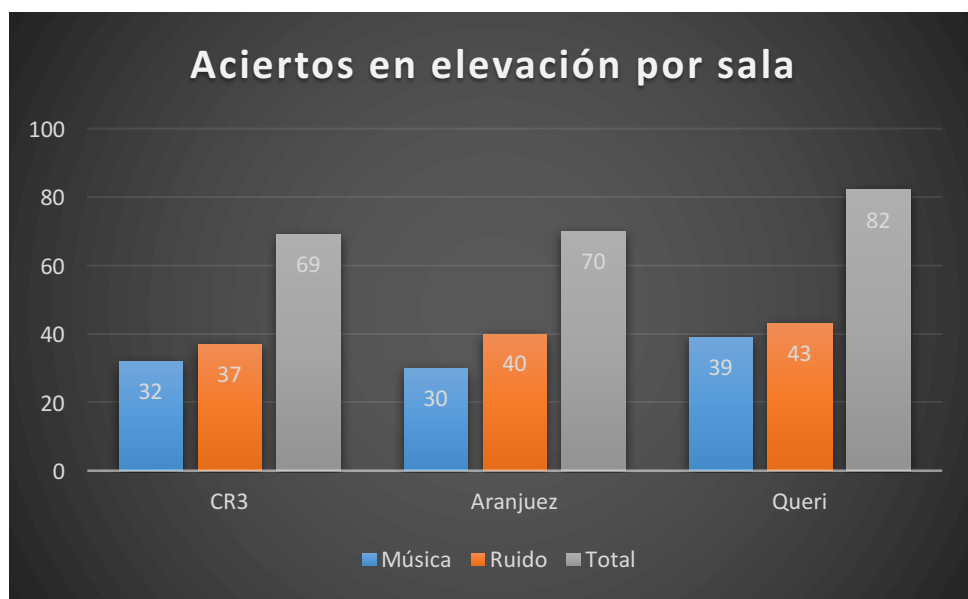


Figura 49. Porcentaje de aciertos en el término elevación divididos por salas.

Al igual que para el término del azimut, en el término de elevación con mayor número de aciertos esta la sala Queri seguida de CR3 y Aranjuez, únicamente se diferencian por 12 aciertos. La mayoría de aciertos están dados por el estímulo de ruido rosa.

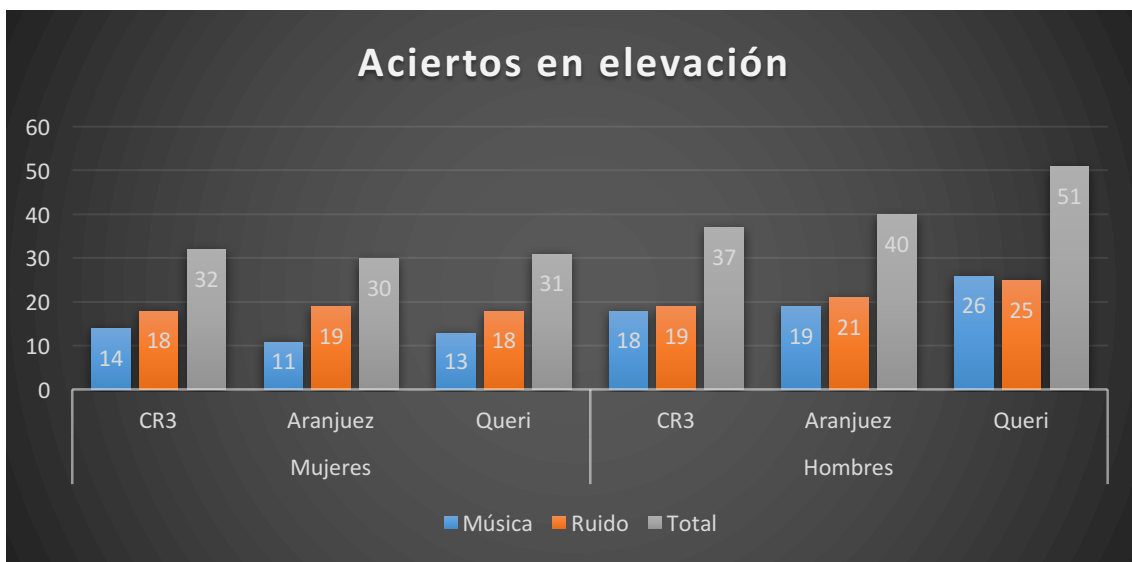


Figura 50. Porcentaje de aciertos en el término de elevación divididos por sexo.

La Figura 50 demuestra que la cantidad de aciertos para las tres salas, por parte del sexo femenino, se diferencia por uno o dos aciertos, y el estímulo sonoro con más aciertos es el ruido.

En el caso de los hombres, la sala Queri predomina en el número de aciertos seguida esta la sala Aranjuez y, por una diferencia no significativa, finalmente el CR3. Los estímulos sonoros son parejos en cuanto a los aciertos, en este sexo.

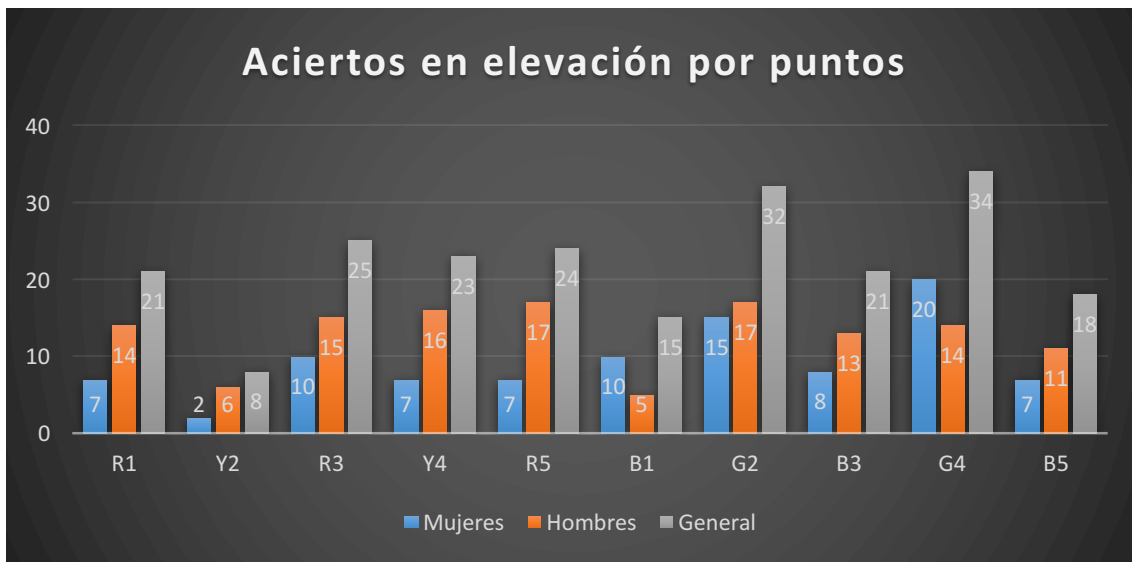


Figura 51. Porcentaje de aciertos en el término altura divididos por puntos.

Los puntos con mayor número de aciertos son G2 y G4 que se encuentran a una altura cercana a la cabeza y en ángulos cerrados. El punto con peores resultados es Y2. Las mujeres tienen mayores aciertos en el punto G4 y el G2, y menor cantidad en Y2. Por otro lado, los resultados de los hombres no tienen mayor diferencia; se destacan B1 y Y2 como los puntos con peores resultados.

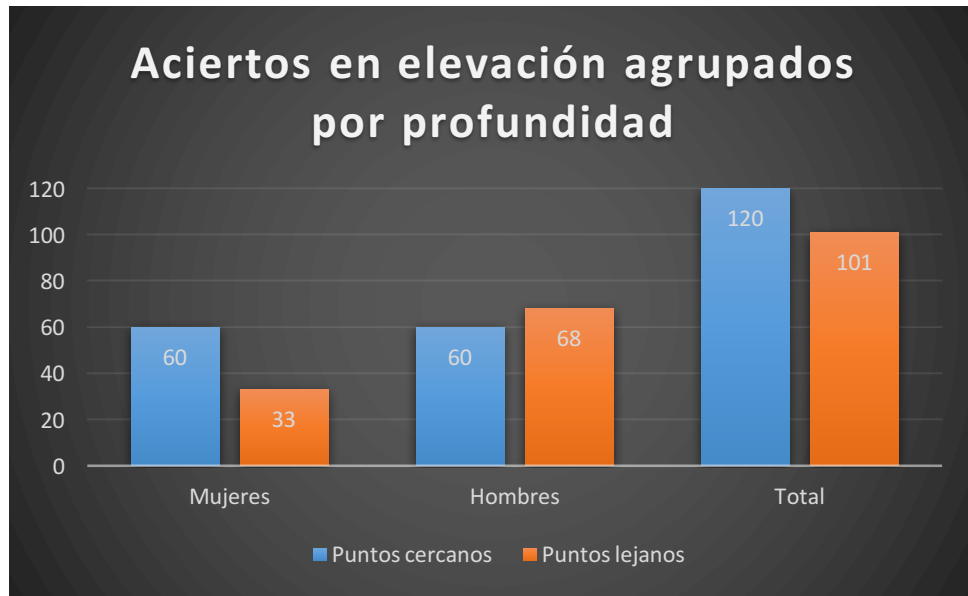


Figura 52. Porcentaje de aciertos en el término altura divididos por puntos cercanos o lejanos.

Para reconocer la elevación de una fuente sonora, la mayoría de la población acertó con puntos más cercanos. Esto se muestra con mayores aciertos en el caso de la mujer mientras que los hombres tienen un comportamiento similar para los puntos cercanos y los puntos lejanos.

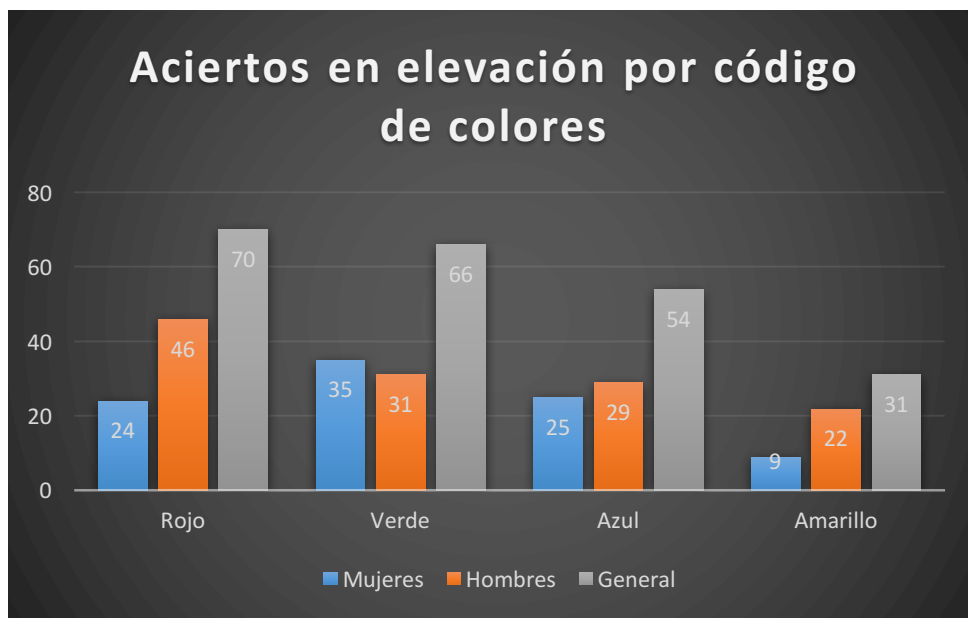


Figura 53. Porcentaje de aciertos en el término altura divididos por eje.

La Figura 53 indica una tendencia; mientras más abajo se encuentre la fuente sonora más difícil es de localizar la misma. Hay una excepción en esta tendencia en el sexo femenino, pues el mayor número de aciertos se presentan en los puntos de color verde, seguidos del rojo, azul y por último el amarillo.

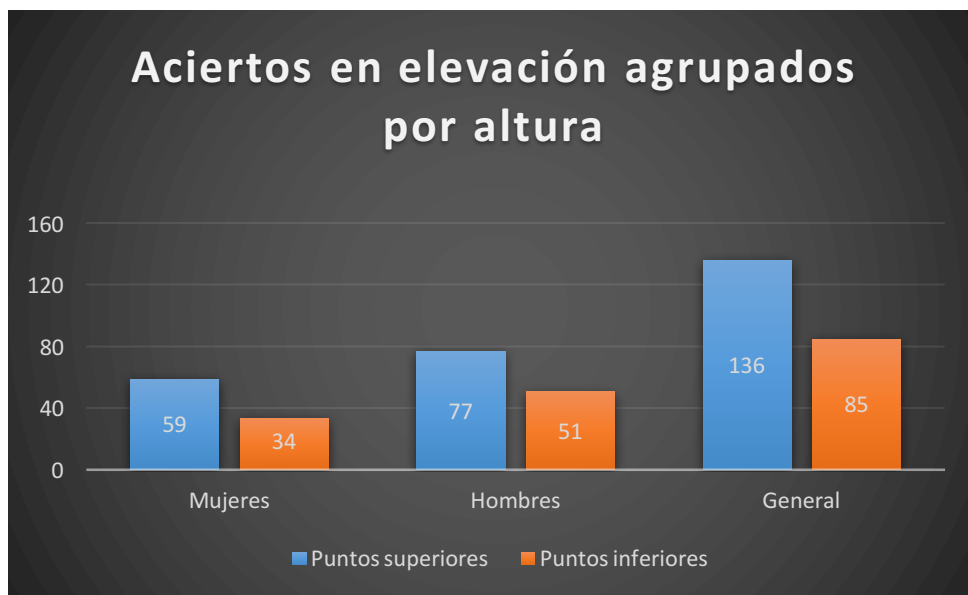


Figura 54. Porcentaje de aciertos en el término altura divididos por puntos superiores e inferiores.

La Figura 54 indica cómo, tanto para hombres como para mujeres, los aciertos son más de la mitad para los puntos superiores. Esta misma tendencia se cumple para hombres como mujeres individualmente.

4.5 Aciertos en distancia

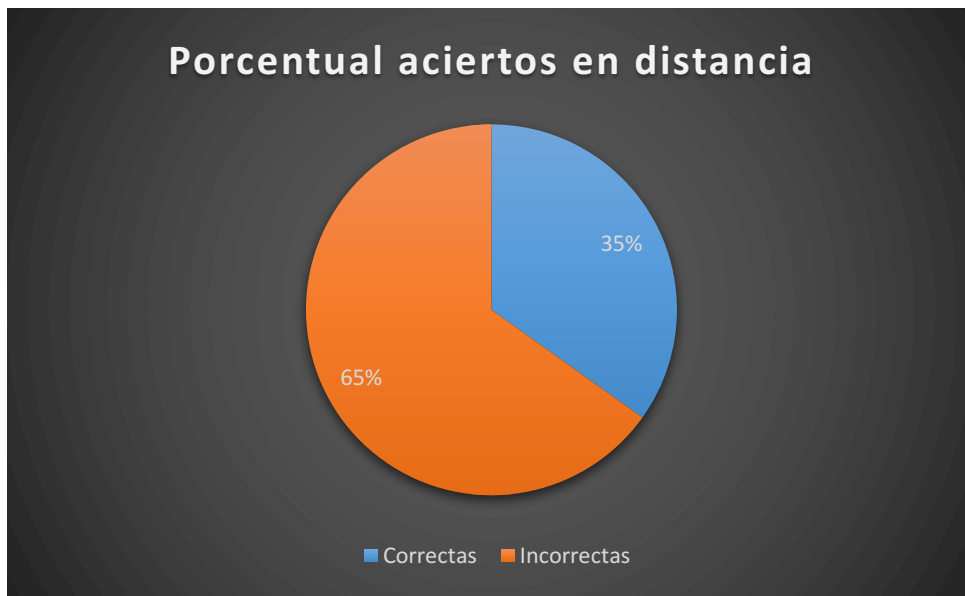


Figura 55. Porcentaje de aciertos globales en el término distancia.

En términos de la distancia se aprecia un porcentaje de 35%, que es un poco mayor al de la elevación.

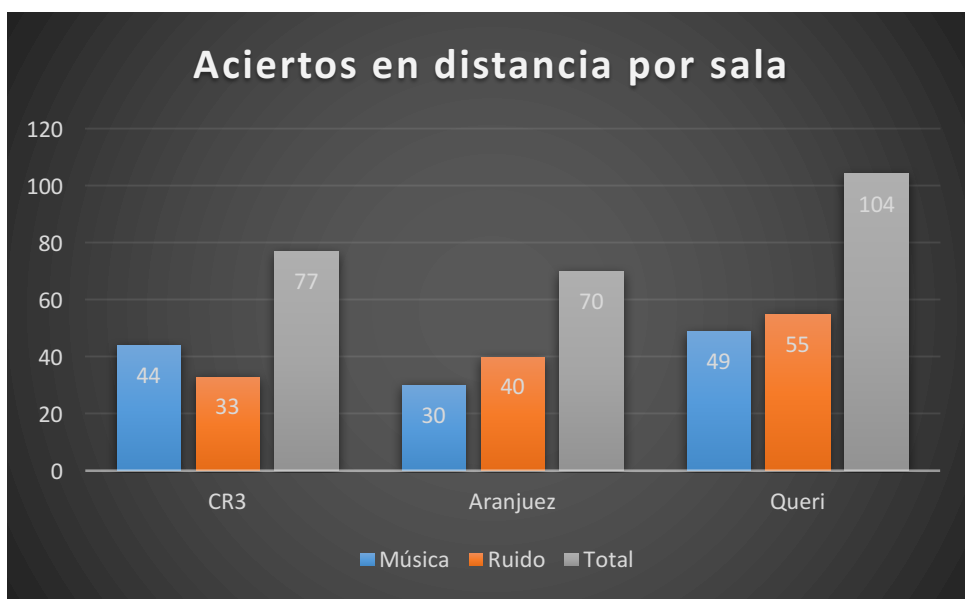


Figura 56. Porcentaje de aciertos en el término distancia divididos por salas.

Al igual que en otros términos, respecto a la distancia, la sala con mayores aciertos es Queri. Se ve de nuevo la influencia de la reverberación para el porcentaje de aciertos.

En cuanto a los estímulos, estos son similares en Aranjuez a diferencia de Queri, en donde predomina el ruido mientras que en el CR3 predomina la música.

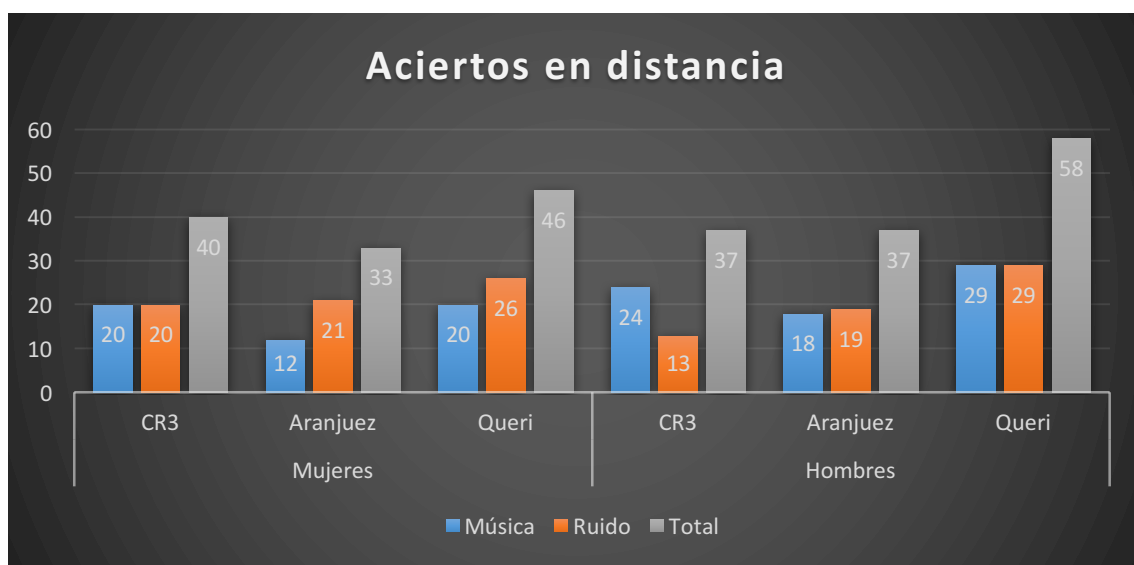


Figura 57. Porcentaje de aciertos en el término distancia divididos por sexo.

La sala que predomina para el sexo masculino es Queri mientras que las salas CR3 y Aranjuez quedan en segundo plano con el mismo número de aciertos. En cuanto a estímulos en la sala CR3 se evidencia más aciertos con música.

Las mujeres tienen mayores aciertos en Queri y luego está CR3, sala en la cual música y ruido tienen el mismo número de aciertos. En las demás salas el ruido tiene más aciertos.

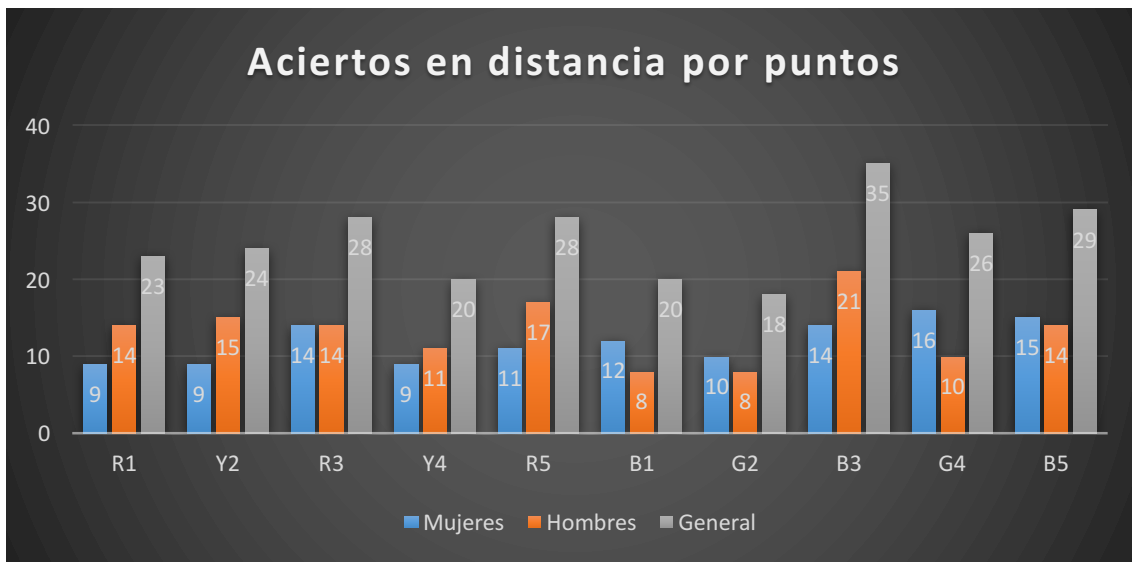


Figura 58. Porcentaje de aciertos en el término distancia divididos por puntos.

El punto que predomina para los hombres es el B3, mientras que el más bajo es el B1 y G2, los cuales se encuentran a menor distancia. Para las mujeres no se observa una gran diferencia entre todos los puntos. En general el punto B3 tiene más aciertos mientras que el G2 es el de menor número de aciertos.

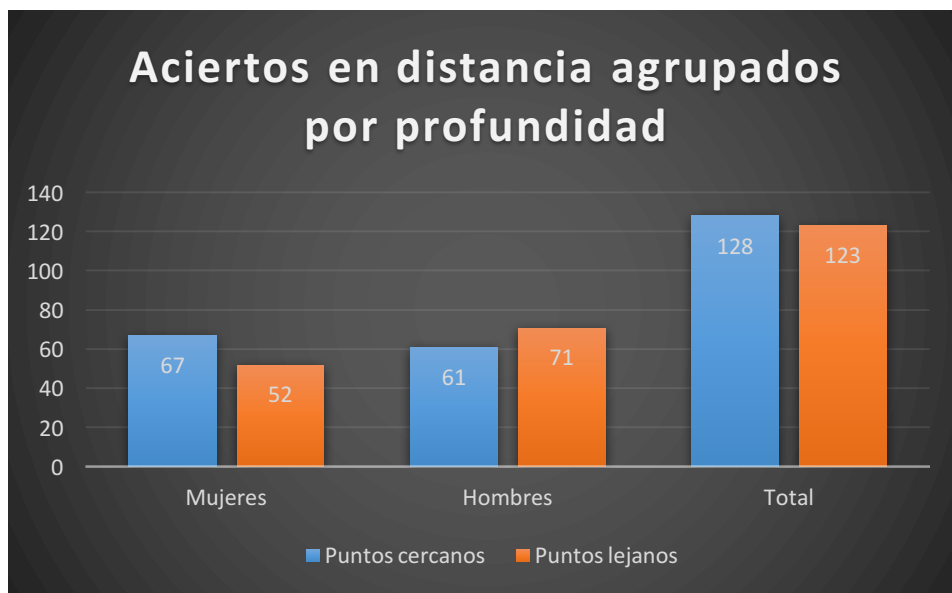


Figura 59. Porcentaje de aciertos en el término distancia divididos por puntos cercanos y puntos lejanos.

En el gráfico se observa que no existe mayor diferencia entre los aciertos de los puntos cercanos y los lejanos. En las mujeres sobresalen, por poco, los puntos cercanos y en los hombres lo contrario, es decir los puntos lejanos.

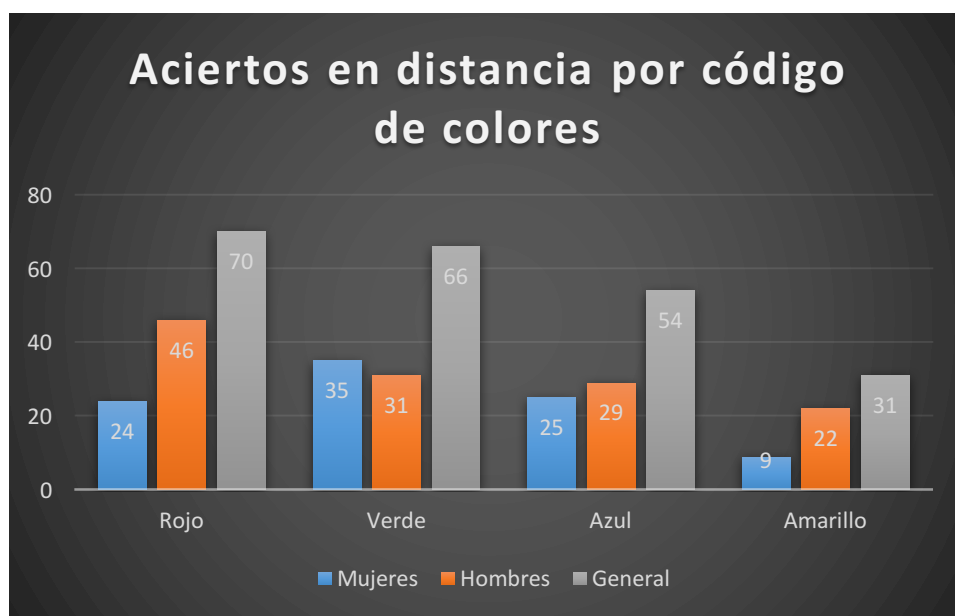


Figura 60. Porcentaje de aciertos en el término distancia divididos por ejes.

Con respecto a distancia en mujeres, tal como sucedió en el término de elevación, se cumple que color con mayor número de aciertos es el verde seguido del azul, luego el rojo y por último el amarillo.



Figura 61. Porcentaje de aciertos en el término distancia divididos por puntos superiores e inferiores.

En cuanto al término de la distancia se puede observar que los encuestados tienen más de la mitad de aciertos en los puntos de la parte superior, esto independientemente del sexo.

4.6 Por elección de puntos

Independientemente de los resultados, es interesante observar cómo cambian las elecciones de ejes y puntos específicos según el ambiente al que son expuestos los encuestados. En esta sección se hace énfasis en esta comparativa.

Se ha tomado el punto más seleccionado por el encuestado, independientemente si es correcto o incorrecto, y se ha marcado una tendencia para el estudio.

4.6.1 Mujeres

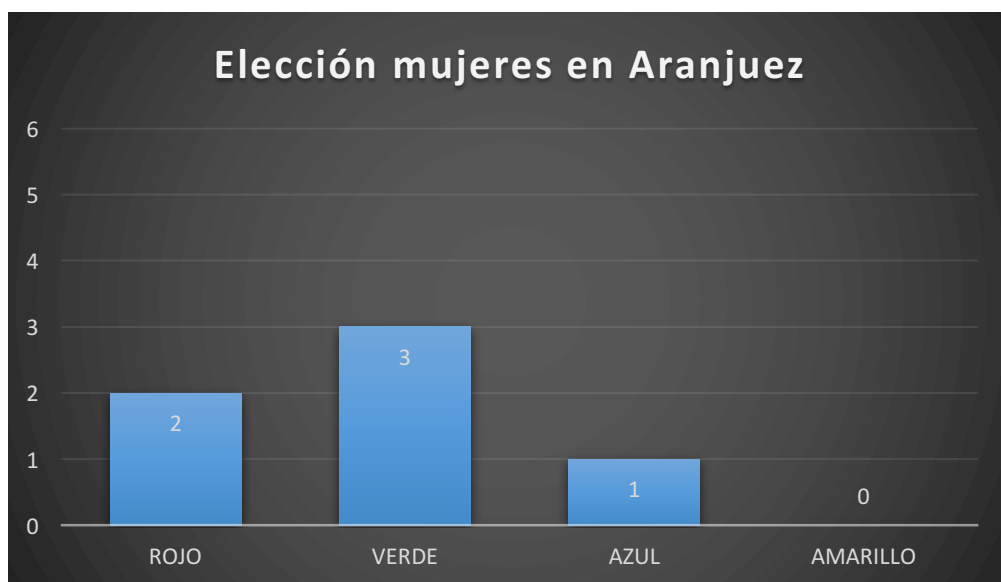


Figura 61. Elección de eje por parte de mujeres en Aranjuez.

Tomando en cuenta las 6 encuestas seleccionadas de mujeres, se observa como en el entorno Aranjuez, las elecciones preferidas son los puntos de color rojo y verde, ubicándose estos puntos en la parte superior.

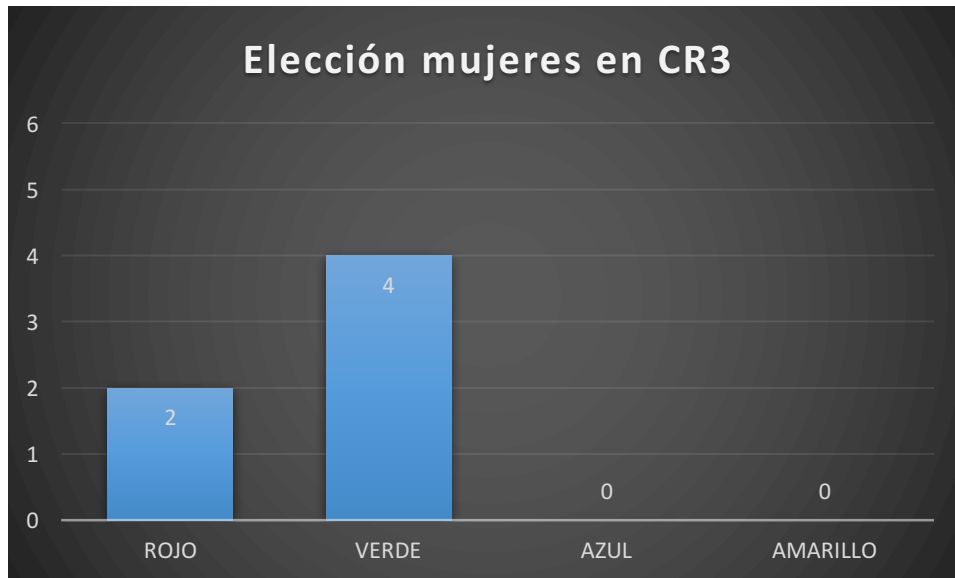


Figura 62. Elección de eje por parte de mujeres en CR3.

En el entorno CR3 se observa como existe una elección única de los puntos que están sobre el oído, a diferencia del anterior entorno en donde existía, por lo menos, una elección de los puntos azules.

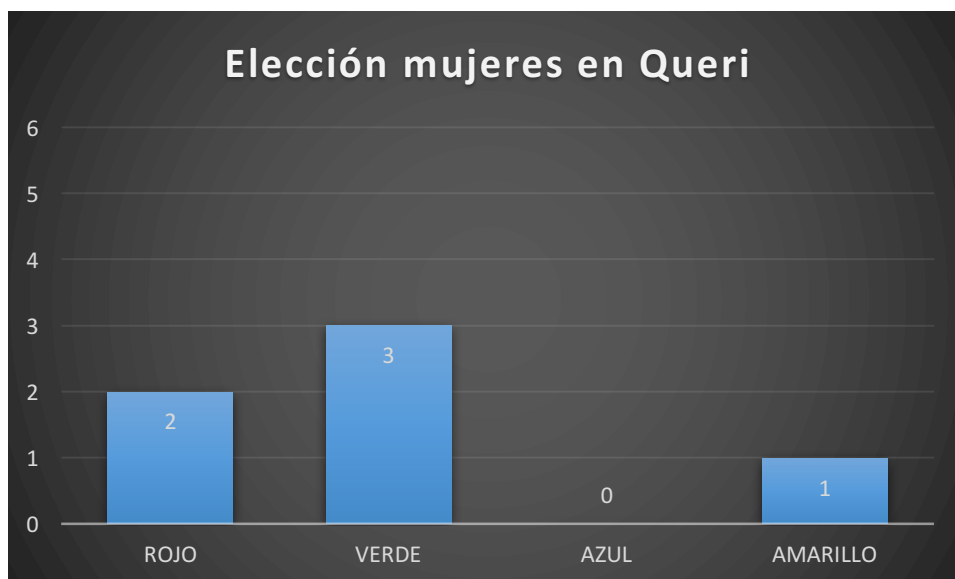


Figura 63. Elección de eje por parte de mujeres en Queri.

Al igual que en los dos entornos anteriores se observa cómo los encuestados seleccionaron mayoritariamente puntos que están sobre el oído.

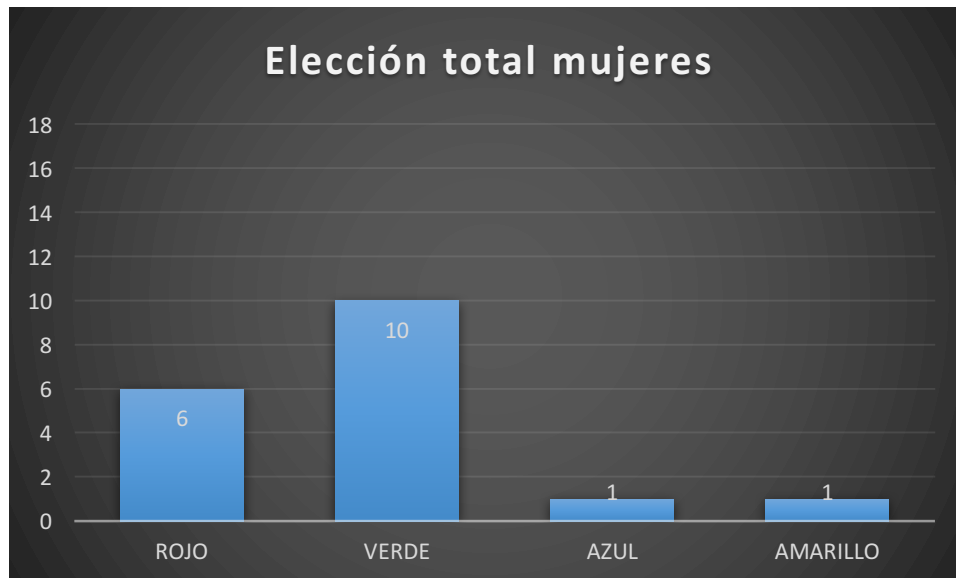


Figura 64. Elección de eje por parte de mujeres en todos los entornos.

Se observa de manera general que las mujeres han seleccionado mayoritariamente puntos, ubicados encima del oído sin importar el entorno o nivel de reverberación.

4.6.2 Hombres

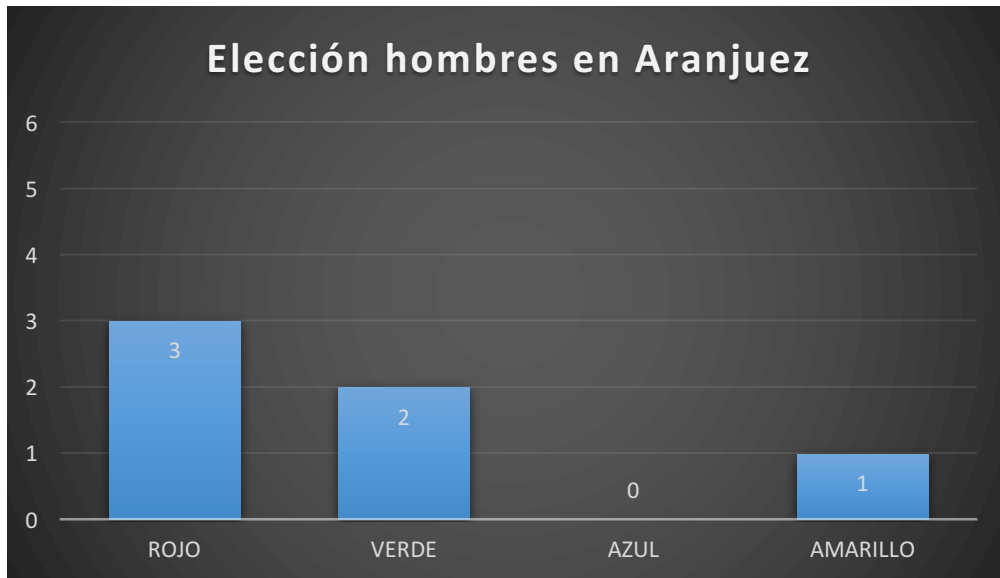


Figura 65. Elección de eje por parte de hombres en Aranjuez.

En el caso de los hombres se mantiene una tendencia similar, pero con mayoría los puntos de color rojo, el cual tiene una distancia y altura más lejana. Se observa que en esta sala hay una mayoría de elección de los puntos lejanos.

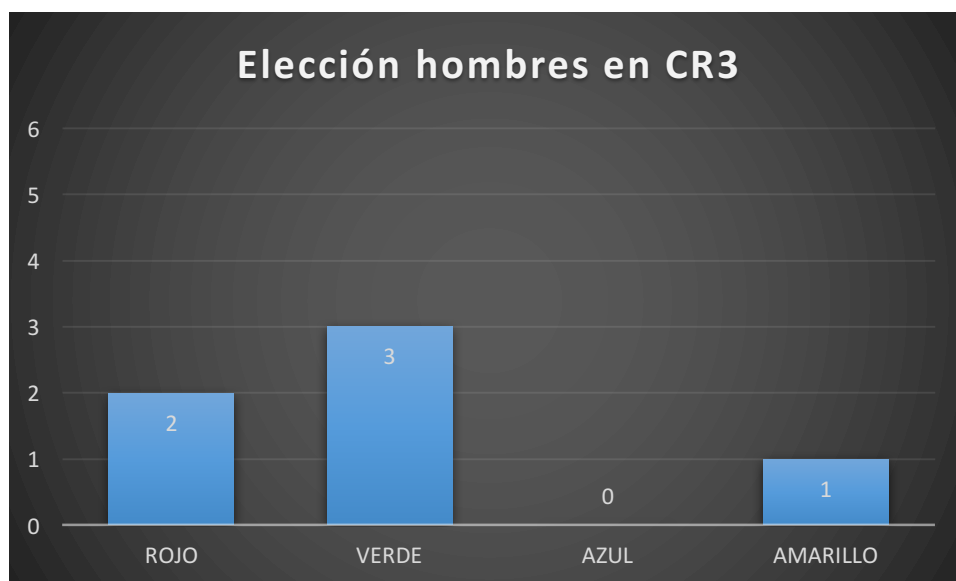


Figura 66. Elección de eje por parte de hombres en CR3.

En este entorno se aprecia una preferencia por el color verde, a diferencia del caso anterior, no hay una mayoría de elección por los puntos lejanos.

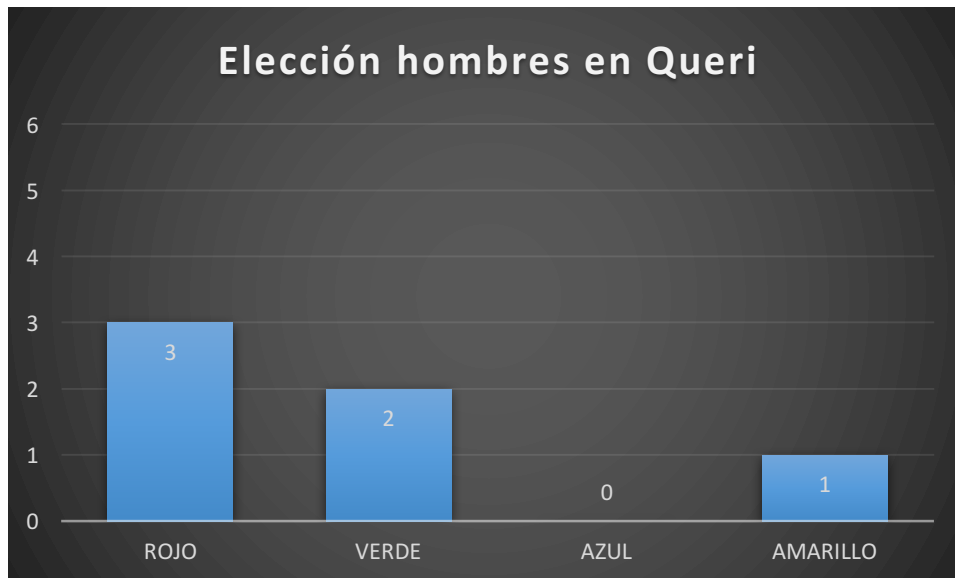


Figura 67. Elección de eje por parte de hombres en Queri.

En este entorno se marca una tendencia completamente igual a lo visto en Aranjuez. Vale recalcar que estos dos entornos tienen un nivel de reverberación más elevado que CR3.



Figura 68. Elección de eje por parte de hombres en todos los entornos.

En el caso general para los hombres se observa la misma tendencia que en las mujeres; estos seleccionaron puntos que están ubicados por encima de su oído. El color rojo es predominante y no existió ninguna elección por el color azul.

4.7 Por punto más seleccionado

4.7.1 Mujeres



Figura 69. Elección de eje por parte de mujeres en todos los entornos.

Tomando en consideración las 6 encuestas de mujeres se tomó los puntos más seleccionados durante el proceso de encuesta. Tal como nos muestra la gráfica, el punto R1 tiene tres elecciones, lo cual representa la mitad, mientras que el punto G2 tiene dos elecciones y el punto G4 solo una. Estos resultados son coherentes con lo expuesto en la elección de los puntos; existe una preferencia por puntos que están encima del nivel del oído humano. Es también destacable la tendencia que se tiene únicamente por estos tres puntos.

4.7.2 Hombres

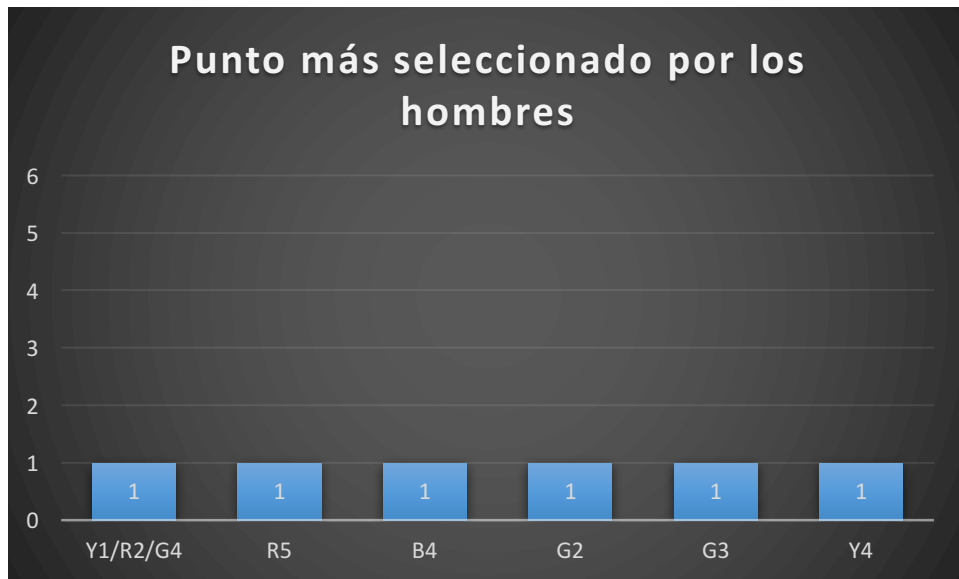


Figura 70. Elección de eje por parte de hombres en todos los entornos.

Tomando en cuenta las 6 encuestas seleccionadas de hombres se observa un comportamiento distinto a lo visto en el caso de las mujeres. No existe una preferencia por algún punto en específico. Esto demuestra que, en el caso de los hombres, no hay preferencias de algún tipo.

5 Conclusiones y recomendaciones

5.1 Conclusiones

Una vez realizado el análisis estadístico con diferentes variables podemos concluir algunos datos que se aborda primero de manera global, es decir a partir de aciertos en los tres términos y luego de manera individual o por cada término; esto dependiendo de las variables seleccionadas. A manera de conclusión global, se puede decir que existe dificultad por parte de los encuestados al identificar de donde proviene un sonido, sin embargo, los resultados de acierto tienen tendencias definidas que a continuación se van a exponer.

Al tomar como variable el tipo de recinto y considerando los tres términos, los mejores resultados corresponden al recinto cuyo tiempo de reverberación es de 1 s (Queri), de manera que dentro de este nivel de reverberación la percepción se ve favorecida, contrario a niveles muy bajos (aproximadamente de 0,4 s) o muy altos (mayores a 2,5 s).

En términos de azimut, se observa que existe un porcentaje similar de aciertos tanto en la sala CR3 como en Queri, por lo que se reafirma la idea expuesta por De la Torre (2016, p. 71) en la cual nos expone que, en azimut, las personas empiezan a experimentar dificultades para la percepción a partir de tiempos de reverberación mayores a 1 s.

Con respecto al término de elevación, existe una mayoría de aciertos en la sala Queri, sin embargo, el margen de aciertos entre recintos se estrecha, de manera que se reafirma la dificultad general que tienen las personas para identificar este término independientemente del nivel de reverberación. Esto puede deberse a la anatomía propia del ser humano, la cual favorece a la percepción en el término

de azimut por la ubicación de los oídos dentro de este eje, contrario al eje en altura.

En términos de distancia vuelve a ser predominante la mayoría del recinto Queri, este se ve muy favorecido por un tiempo de reverberación medio de aproximadamente 1 s. El motivo principal puede estar en las primeras reflexiones que ayudan al cerebro a tener una referencia espacial para la ubicación de fuente en distancia, sin embargo, si el tiempo de reverberación es muy elevado, el número y nivel de reflexiones sonoras tendrá un efecto negativo generando confusión y desubicación.

El segundo criterio es el sexo de los encuestados. Analizando el porcentaje de aciertos en los tres términos, se observa en el caso de las mujeres, un número de aciertos mayor tanto en Queri como en CR3, lo cual presume que dentro de un intervalo de aproximadamente 0,4 a 1 s de tiempo de reverberación la audición femenina se ve beneficiada al localizar una fuente sonora. En el caso de los hombres, el panorama se reduce únicamente a tiempos de reverberación cercanos a 1 s, ya que se observa como en el ambiente CR3 existe un decaimiento bastante notorio del porcentaje de aciertos en comparación con los resultados femeninos, siendo esta la principal diferencia entre ambos sexos.

Analizando el término azimut dentro de este criterio, se observa que las mujeres identifican de mejor manera este término dentro del ambiente CR3, es decir con un ambiente cuyo tiempo de reverberación se asemeja a los 0,4 s. El sexo masculino se mantiene en la tendencia anteriormente descrita; identifica cada término de mejor manera dentro del ambiente cercano al tiempo de reverberación de 1 s.

Se puede concluir que existen diferencias notorias con respecto a la percepción de fuente sonora entre hombres y mujeres. Las mujeres poseen un nivel de precisión menor de manera global, sin embargo, este no se verá drásticamente afectado a diferencia de los hombres, que únicamente manejan un nivel de precisión alto dentro de un ambiente de alrededor de 1 s de tiempo de reverberación.

Tomando como criterio el tipo de estímulo, la población no presenta una preferencia entre el ruido rosa y la música, en los tres términos y de manera individual, es por un pequeño porcentaje que en el término de elevación el ruido rosa tiene una mayoría de aciertos con respecto a la música. Si se observa el número de aciertos por sexo, los hombres en azimut, distancia y en los tres términos aciertan al tener como estímulo la música. Esto sugiere que, al escuchar un tema conocido, cuya sonoridad ya es familiar por el participante de sexo masculino, es más fácil localizar de donde proviene el sonido. En cuanto a las mujeres, estas tienden a acertar con el estímulo del ruido rosa. Esto indica que el sexo femenino no tiene preferencia por un sonido familiar y su audición se enfoca completamente en captar la localización del sonido en base a su rango de frecuencias.

Al realizar el análisis correspondiente al porcentaje de aciertos en los tres términos, clasificando a los puntos por colores y dividiéndolos en ejes superiores e inferiores, se observa un mayor porcentaje de aciertos en los puntos superiores. Esto debido a que en la cotidianidad no se acostumbra a poner equipos de sonido ni fuentes que emitan sonido cerca del piso. Lo mismo ocurre para los términos individuales de distancia y especialmente en elevación que es el eje que estamos analizando, esto se vuelve irrelevante para el término de azimut, en donde el mayor número de aciertos los tiene el color azul, que son puntos cercanos a los oídos pero abajo de los mismos, una hipótesis para esto sería que en este punto existían más puntos de selección por lo cual existía

mayor posibilidad de aciertos; esto se repite con menor porcentaje en los aciertos con los tres términos.

En cuanto al análisis de cercanía de puntos, en tres términos y en los términos individuales, no existe mayor diferencia entre los puntos cercanos con los lejanos; por lo cual podemos decir que los participantes pueden diferenciar puntos cercanos y puntos lejanos de igual manera. Si hacemos un análisis por sexo se puede ver que las mujeres tienden a acertar la ubicación de una fuente cuando esta está a menor distancia, mientras que para los hombres es un poco irrelevante esta variable.

En cuanto al término de azimut la mayoría de población reconoce con exactitud el ángulo independiente de los otros términos. Al igual que el estudio realizado por De la Torre (2016), se obtuvieron buenos resultados para este término puesto que las respuestas son más cerradas y no existía un margen de error grande.

Para la elevación y la distancia los encuestados tienen dificultad de acertar en estos términos, al no conocer la estructura física de la sala, los materiales de los cuales está hecha o una referencia de cómo cambia el volumen o en si el estímulo dependiendo de las variaciones de los términos hacen que las personas no puedan localizar la fuente.

5.2 Recomendaciones

Este estudio plantea una serie de conclusiones generales que pueden ser profundizadas en estudios futuros dentro de varias ramas de sonido y acústica.

Las conclusiones que se plantean con respecto a los términos de distancia y elevación pueden ser esclarecidas a través de estudios de estos términos de forma individual. En el caso de la elevación se puede establecer un rango angular en el que las reflexiones principalmente del techo afecten a la percepción de este término.

En este estudio se utilizó una guía visual para todos los encuestados de manera que se podría simular las salas reales con todos sus materiales y estructuras. Un aspecto interesante para profundizar es la influencia que tiene este ámbito dentro del reconocimiento; podría resultar interesante realizar comparaciones entre personas que tengan guía visual y quiénes no para tener datos más precisos.

Respecto a los puntos que se utilizó para la grabación de respuesta impulsiva y la convolución de los sonidos, se utilizó únicamente puntos frontales. Sería interesante un estudio que incluya puntos posteriores, de manera que se analice la percepción en 360 grados y se puedan realizar comparaciones entre percepción frontal y trasera.

Con respecto a las diferencias obtenidas entre hombres y mujeres, se puede plantear estudios que analicen las diferencias entre el sistema auditivo y la percepción a nivel neuronal. Esto ayudaría a esclarecer el motivo por el que las mujeres tienen menor precisión en general, pero su percepción no varía tan

drásticamente con tiempos de reverberación bajos al contrario de lo que sucede con hombres.

Finalmente, en este estudio se tuvo que realizar una compensación de nivel mediante software externo, por lo que para futuros estudios se deberá tener especial cuidado en la calibración de los micrófonos del sistema binaural y realizar una prueba previa para ver el funcionamiento normal del sistema.

REFERENCIAS

- Arias, C., y Ramos, O. A. (2003). Audición Espacial en Ambientes Reverberantes: Aspectos Teóricos Relevantes. *Revista Interamericana de Psicología*, 37(5010), 373–382.
- Asociación Española de Normalización y Certificación. (2008). *UNE en ISO 3382: Medición de parámetros acústicos en recintos*. Madrid, España: AENOR
- Balan, O., Moldoveanu, F., Morar, A., y Asavei, V. (2014). Experiments on training the sound localization abilities: A systematic review. In *The International Scientific Conference eLearning and Software for Education* (pp. 34–41). Bucharest.
- Bomhardt, R. (2017). *Anthropometric Individualization of Head-Related Transfer Functions Analysis and Modeling: Logos Verlag Berlin*. Recuperado el 5 de octubre de 2017 de <https://books.google.com.ec/books?id=PtszDwAAQBAJ>
- Carrizo, M. (2014). *Fourier y el procesamiento digital de señales*. Recuperado el 23 de diciembre de 2017 de <http://lcr.uns.edu.ar/fvc/NotasDeAplicacion/FVC-Martin%20Carrizo.pdf>
- De la Torre, J. A. (2016). Evaluación psicoacústica del efecto de precedencia de una señal acústica con distintos niveles de reverberación, mediante auralización y estímulos fusionados. *Tesis de la Universidad de Las Américas*.
- Domingo, A. (2014). *Apuntes de Acústica*. Recuperado el 15 de febrero de 2018 de <http://oa.upm.es/23098/1/amd-apuntes-acustica-v2.1.pdf>
- Europa Press. (2011). *¿Por qué 'We are the champions', de Queen, es la canción más pegadiza de la historia?*. Recuperado el 26 de septiembre de 2017 de <http://www.20minutos.es/noticia/1176872/0/cancion/pegadiza/queen/>

- Galindo, E. (2008). *Estadística: Método y Aplicaciones*. Segunda Edición. Quito: Prociencia
- Ghamdan, L., Ismail Shoman, M. A., Elwahab, R. A., y El-Hadid Ghamry, N. A. (2017). Position estimation of binaural sound source in reverberant environments. *Egyptian Informatics Journal*. <https://doi.org/10.1016/j.eij.2016.05.002>
- Goldstein, E. B. (2008). *The Blackwell Handbook of Sensation and Perception*. Wiley. Recuperado el 15 de diciembre de 2017 de <https://books.google.com.ec/books?id=Fs-5McBOqU4C>
- McMullan, R. (2012). *Environmental Science in Building*. Palgrave Macmillan. Recuperado de <https://books.google.com.ec/books?id=A8gcBQAAQBAJ>
- Metro. (s.f.). *1,67 cm es la estatura promedio de los ecuatorianos*. Recuperado el 20 de septiembre de 2017 de <https://www.metroecuador.com.ec/ec/noticias/2016/07/26/167-cm-estatura-promedio-ecuatorianos.html>
- Neumann. (s.f.). *Dummy Head KU 100*. Recuperado el 1 de enero de 2018 de https://www.neumann.com/?lang=en&id=current_microphones&cid=ku100_data
- Nilsson, M. E., y Schenkman, B. N. (2016). Blind people are more sensitive than sighted people to binaural sound-location cues, particularly inter-aural level differences. *Hearing Research*, 332, 223–232. <https://doi.org/10.1016/j.heares.2015.09.012>
- Oberem, J., Masiero, B., y Fels, J. (2016). Experiments on authenticity and plausibility of binaural reproduction via headphones employing different recording methods. *Applied Acoustics*, 114, 71–78. <https://doi.org/10.1016/j.apacoust.2016.07.009>
- Paquier, M., Côté, N., Devillers, F., y Koehl, V. (2016). Interaction between auditory and visual perceptions on distance estimations in a virtual

environment. *Applied Acoustics*, 105, 186–199.
<https://doi.org/10.1016/j.apacoust.2015.12.014>

Pourmohammad, A., y Ahadi, S. (2012). TDE-ILD-HRTF-Based 2D Whole-Plane Sound Source Localization Using Only Two Microphones and Source Counting. *Ijeee.Org*, 2(3), 307–313. Retrieved from <http://www.ijeee.org/papers/104-1127.pdf>

Pulkki, V., y Karjalainen, M. (2014). *Communication acoustics: an introduction to speech, audio and psychoacoustics*. Recuperado el 5 de noviembre de 2017 de <https://ebookcentral-proquest-com.bibliotecavirtual.udla.edu.ec>

Sennheiser. (2016). *HD 280 PRO*. Recuperado el 8 de enero de 2018 de <https://en-us.sennheiser.com/hd-280-pro>

Zhong, X., Sun, L., y Yost, W. (2016). Active binaural localization of multiple sound sources. *Robotics and Autonomous Systems*, 85, 83–92.
<https://doi.org/10.1016/j.robot.2016.07.008>

Zhu, P., Mo, F., Kang, J., y Zhu, G. (2015). Comparisons between simulated and in-situ measured speech intelligibility based on (binaural) room impulse responses. *Applied Acoustics*, 97, 65–77.
<https://doi.org/10.1016/j.apacoust.2015.04.005>

ANEXOS

Anexo 1 – Acrónimos

- DOA: Direction of arrival - *Dirección de llegada.*
- HpTF: Headphone transfer function - *Función de transferencia de audífonos.*
- HRIR: Head Related Impulse Response - *Respuesta de impulso relacionada con la cabeza.*
- HRTF: Head related transfer function - *Función de transferencia relacionada con la cabeza.*
- ILD: Interaural level difference - *Diferencia de nivel interaural.*
- ITD: Interaural time difference - *Diferencia de tiempo interaural.*
- STI: Speech Transmission Index - *Índice de transmisión del habla.*

Anexo 2 – Medición y normativas correspondientes

Para medir el tiempo de reverberación existe la norma UNE EN ISO 3382 (2008) parte 1, orientada a salas de espectáculos, y la parte 2 para recinto ordinarios. Esta norma tiene varias metodologías, pero solamente describimos la respuesta impulsiva a nivel de ingeniería.

Para este tipo de método se requiere tener una fuente omnidireccional que excite todo el recinto y que genere un nivel de presión sonora de 45 dB por encima del ruido de fondo, para poder medir el T_{30} , en caso de ser necesario. Para el ensayo únicamente se requiere de dos personas en la sala.

El nivel de ingeniería determina:

- 6 combinaciones fuente-micrófono
- 2 o más posiciones de fuente.
- 2 o más posiciones de micrófono.
- Un rango de frecuencias de 100 Hz a 5000 Hz.

En cuanto a las posiciones:

- Se requiere tener mínimo 2 m de distancia entre las posiciones de micrófono.
- Se requiere tener mínimo 1 m de distancia entre el micrófono y cualquier superficie reflejante.
- Evitar posiciones simétricas, variar el micrófono en los tres ejes.

Anexo 3 – Tiempos de reverberación

CR3

Datos del recinto	
Recinto:	Estudio de grabación - CR3
Fecha:	8/10/17
Dirección:	Granados y Colimes
Observaciones relevantes:	Se utilizó paneles móviles para cubrir las ventanas del estudio, evitando las reflexiones.

Frecuencia (Hz)	Posición 1			Posición 2			Promedio (s)
	Toma 1	Toma 2	Toma 3	Toma 1	Toma 2	Toma 3	
50	0,32	0,33	0,49	0,53	-	-	0,42
63	-	0,51	0,54	0,37	0,61	0,31	0,47
80	0,38	0,43	0,53	-	0,40	0,40	0,43
100	0,39	0,36	0,42	0,32	0,31	0,33	0,36
125	0,39	0,35	0,36	0,32	0,30	0,36	0,35
160	0,50	0,39	0,38	0,40	0,41	0,39	0,41
200	0,41	0,39	0,40	0,42	0,36	0,40	0,40
250	0,40	0,38	0,37	0,37	0,40	0,32	0,37
315	0,40	0,34	0,36	0,36	0,31	0,38	0,36
400	0,32	0,32	0,30	0,28	0,28	0,27	0,30
500	0,39	0,36	0,51	0,42	0,35	0,44	0,41
630	0,34	0,35	0,33	0,36	0,37	0,33	0,35
800	0,41	0,33	0,30	0,33	0,35	0,24	0,33
1000	0,30	0,33	0,31	0,37	0,35	0,36	0,34
1250	0,38	0,39	0,34	0,38	0,34	0,34	0,36
1600	0,33	0,38	0,31	0,32	0,42	0,25	0,34
2000	0,37	0,39	0,38	0,46	0,32	0,39	0,39
2500	0,36	0,35	0,36	0,36	0,38	0,36	0,36
3150	0,35	0,30	0,32	0,34	0,33	0,31	0,33
4000	0,36	0,35	0,32	0,32	0,34	0,32	0,34
5000	0,33	0,34	0,34	0,36	0,32	0,35	0,34
						TR60 (100 - 5k)	0,36

Magic

Datos del recinto	
Recinto:	Estudio de grabación - Magic Estudio
Fecha:	20/8/17
Dirección:	Ulloa y Rumipamba
Observaciones relevantes:	Se utilizó paneles en la pared lateral izquierda para evitar las reflexiones dadas por la pared de piedra.

Frecuencia (Hz)	Posición 1			Posición 2			Promedio (s)
	Toma 1	Toma 2	Toma 3	Toma 1	Toma 2	Toma 3	
50	-	-	-	-	-	-	0,00
63	0,74	-	-	-	-	-	0,74
80	0,58	-	0,54	0,72	0,62	0,61	0,61
100	-	-	0,88	-	0,69	0,77	0,78
125	0,65	0,70	0,64	0,67	0,72	0,66	0,67
160	-	1,04	0,61	0,80	0,70	0,61	0,75
200	0,57	0,77	0,56	0,77	0,59	0,67	0,66
250	0,52	0,52	0,48	0,46	0,47	0,50	0,49
315	0,50	0,53	0,53	0,48	0,58	0,43	0,51
400	0,48	0,36	0,37	0,48	0,46	0,41	0,43
500	0,77	0,62	0,55	0,70	0,61	0,57	0,64
630	0,80	0,63	0,56	0,68	0,76	0,65	0,68
800	0,78	0,45	0,68	0,68	0,56	0,52	0,61
1000	0,59	0,51	0,65	0,56	0,45	0,53	0,55
1250	0,50	0,43	0,58	0,46	0,47	0,48	0,49
1600	0,45	0,42	0,43	0,38	0,44	0,36	0,41
2000	0,56	0,48	0,43	0,50	0,49	0,40	0,48
2500	0,53	0,49	0,45	0,55	0,51	0,39	0,49
3150	0,52	0,44	0,43	0,41	0,45	0,39	0,44
4000	0,36	0,45	0,38	0,44	0,40	0,39	0,40
5000	0,39	0,39	0,42	0,37	0,34	0,34	0,38
TR60 (100 - 5k)							0,55

Queri

Datos del recinto	
Recinto:	Auditorio Queri - Universidad de las Américas
Fecha:	5/10/17
Dirección:	Queri
Observaciones relevantes:	Se realizó la medición con sillas y equipos en la localidad.

Frecuencia (Hz)	Posición 1			Posición 2			Promedio (s)
	Toma 1	Toma 2	Toma 3	Toma 1	Toma 2	Toma 3	
50	-	-	-	-	-	-	0,00
63	-	-	-	-	-	-	0,00
80	-	-	-	-	-	-	0,00
100	-	-	-	-	-	-	1,81
125	-	-	-	-	-	-	1,81
160	-	-	-	-	1,81	-	1,81
200	1,47	1,67	1,40	1,32	1,77	0,89	1,27
250	1,46	1,47	1,34	1,48	1,40	1,31	1,19
315	1,21	1,25	1,17	1,29	1,29	1,20	1,04
400	1,19	1,16	1,07	1,10	1,19	1,41	0,95
500	1,01	0,93	1,06	0,96	0,96	1,06	0,82
630	1,07	1,19	0,96	1,02	1,05	1,05	0,88
800	1,07	1,09	0,99	0,99	1,12	1,06	0,88
1000	1,13	1,09	1,07	1,09	1,20	1,36	0,93
1250	1,17	1,29	1,39	1,20	1,07	1,33	1,02
1600	1,23	1,36	1,18	1,09	1,14	1,03	1,00
2000	1,29	1,41	1,25	1,27	1,27	1,21	1,08
2500	1,40	1,30	1,51	1,45	1,36	1,08	1,17
3150	1,88	1,54	1,29	0,86	1,38	0,91	1,16
4000	1,29	1,25	1,25	1,39	1,27	1,10	1,08
5000	1,35	1,18	1,16	1,03	1,20	0,85	0,99
						TR60 (100 - 5k)	1,16

Aranjuez

Datos del recinto	
Recinto:	Sala comunal - Edificio Aranjuez
Fecha:	19/8/17
Dirección:	Ignacio Bossano y Av. 6 de diciembre
Observaciones relevantes:	Espacios para yacusi, turco y cocina en la locación.

Frecuencia (Hz)	Posición 1			Posición 2			Promedio (s)
	Toma 1	Toma 2	Toma 3	Toma 1	Toma 2	Toma 3	
50	-	-	-	-	-	-	0,00
63	-	-	-	-	-	-	0,00
80	-	-	-	-	-	-	0,00
100	3,74	4,44	-	-	-	-	4,09
125	3,50	2,89	2,69	2,18	2,60	2,82	2,78
160	1,93	3,28	2,37	2,16	2,57	2,57	2,48
200	3,27	2,77	2,81	2,76	3,00	2,90	2,92
250	3,10	3,02	2,66	3,77	3,07	2,71	3,06
315	2,98	3,21	3,14	3,00	3,40	2,88	3,10
400	3,37	2,89	2,87	3,65	2,98	3,06	3,14
500	2,95	3,00	2,86	3,19	2,66	3,07	2,96
630	3,07	3,40	2,99	3,01	2,99	3,06	3,09
800	3,13	3,06	2,88	3,03	3,19	3,01	3,05
1000	3,07	3,15	3,03	3,03	3,05	2,75	3,01
1250	2,87	2,81	2,81	2,78	2,86	2,92	2,84
1600	2,86	2,71	2,48	1,89	2,79	2,50	2,54
2000	2,71	2,61	2,59	2,58	2,47	2,59	2,59
2500	2,12	2,18	2,40	2,29	2,09	2,39	2,25
3150	1,81	1,86	1,83	2,09	1,92	1,80	1,89
4000	1,93	1,82	1,81	1,97	1,82	1,80	1,86
5000	1,81	1,54	1,75	1,72	1,62	1,62	1,68
						TR60 (100 - 5k)	2,74

Anexo 4 – Especificaciones técnicas del dummy head KU100

Para la grabación binaural utilizamos un dummy head de la marca Neumann modelo KU100.



Figura 71. Dummy head KU100. Tomada de Neumann, s.f.

A continuación, se observará la respuesta en frecuencia de los micrófonos de la misma. Esta respuesta es bastante plana hasta los 5000 Hz y tiene una caída pronunciada en 8500 Hz aproximadamente.

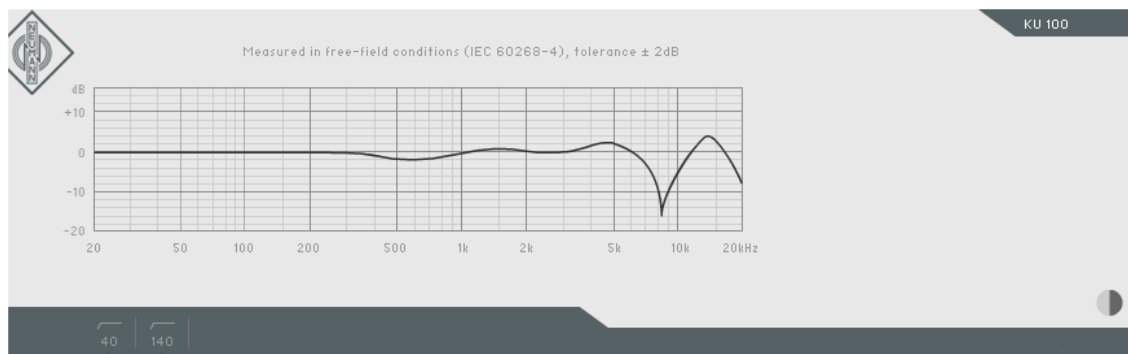


Figura 72. Respuesta en frecuencia correspondiente a dummy head KU100 . Tomada de Neumann, s.f.

Anexo 5 – Máximos absolutos

Con música

Puntos	Con música								
	Aranjuez			CR3			Queri		
	L	L (+ 5,1 dB)	R	L	L (+ 5,1 dB)	R	L	L (+ 5,1 dB)	R
R1B	10,14	18,28	17,65	3,70	6,65	12,03	3,99	7,17	9,66
R2B	9,50	17,15	16,00	2,58	4,66	6,32	2,51	4,55	7,87
R3B	8,36	15,08	18,91	5,25	9,51	9,96	2,74	4,96	5,16
R4B	8,94	16,07	19,90	3,77	6,78	5,75	2,98	5,36	5,68
R5B	10,20	18,46	17,52	5,31	9,55	8,06	2,33	4,17	3,86
G1B	6,41	11,54	13,19	4,80	8,65	10,86	3,66	6,58	11,05
G2B	9,42	16,94	18,35	7,18	12,93	14,36	3,16	5,67	8,79
G3B	9,40	16,96	17,83	4,63	8,33	7,64	3,16	5,69	6,48
G4B	7,73	13,91	15,32	4,76	8,52	9,16	3,67	6,63	5,92
G5B	11,77	21,17	23,71	7,00	12,63	9,85	6,16	11,05	7,71
B1B	8,18	14,73	16,16	4,70	8,47	15,30	4,26	7,68	13,17
B2B	8,47	15,25	18,38	6,15	11,09	14,81	3,39	6,06	10,17
B3B	10,38	18,73	21,26	5,79	10,42	12,37	4,90	8,77	10,48
B4B	7,76	13,96	17,15	9,90	17,74	18,42	6,38	11,43	8,97
B5B	11,94	21,47	22,04	6,83	12,36	9,57	9,84	17,70	14,29
Y1B	7,35	13,22	15,68	4,50	8,23	9,40	3,51	6,34	9,54
Y2B	8,30	14,88	14,57	4,42	7,95	11,92	2,56	4,62	6,98
Y3B	6,82	12,29	13,75	3,20	5,87	7,20	4,58	8,22	9,23
Y4B	10,74	19,34	20,69	3,87	6,96	6,30	5,89	10,56	5,73
Y5B	9,15	16,47	17,47	3,29	5,92	6,08	3,82	6,89	6,06

Con ruido rosa

Puntos	Con ruido rosa					
	Aranjuez		CR3		Queri	
	<i>L (+ 5,1 dB)</i>	<i>R</i>	<i>L (+ 5,1 dB)</i>	<i>R</i>	<i>L (+ 5,1 dB)</i>	<i>R</i>
R1B	14,44	15,63	5,95	8,35	4,45	6,59
R2B	11,74	14,10	4,16	4,80	4,07	6,99
R3B	12,68	14,90	7,27	7,36	3,71	4,47
R4B	14,23	15,21	4,88	5,27	5,59	5,11
R5B	12,76	12,77	7,36	7,15	4,01	3,08
G1B	11,42	13,86	6,71	8,41	6,13	10,32
G2B	13,92	19,52	7,81	10,79	4,39	8,07
G3B	12,85	17,89	5,46	6,99	4,32	5,30
G4B	11,99	13,47	5,81	6,18	6,31	4,25
G5B	19,72	19,71	9,80	7,72	10,59	6,28
B1B	11,89	13,92	6,54	9,91	4,84	8,91
B2B	12,60	18,23	8,18	12,31	3,47	6,98
B3B	15,59	16,96	8,31	9,53	6,93	7,97
B4B	14,49	15,24	13,54	12,03	10,84	6,83
B5B	16,02	15,19	10,26	7,31	11,38	8,41
Y1B	12,69	15,43	7,97	9,57	4,69	8,50
Y2B	11,85	13,21	6,31	9,70	3,90	7,02
Y3B	11,30	12,59	5,62	6,43	5,74	6,60
Y4B	15,16	17,67	6,57	5,08	6,97	5,21
Y5B	16,27	16,60	5,69	4,76	5,44	4,49

Nota: En esta tabla únicamente se presentan los valores con la compensación hecha

Anexo 6 – Especificaciones técnicas de los audífonos

Los audífonos Sennheiser HD 280 Pro son un modelo semi-cerrado con respuesta plana para monitoreo profesional.



Figura 73. Audífonos Sennheiser HD 280 pro. Tomado de la página oficial de Sennheiser.

Su respuesta en frecuencia tiene una tendencia plana, de manera que favorece a que en el proceso de encuesta no se presenten coloraciones que pudieran influir en los resultados.

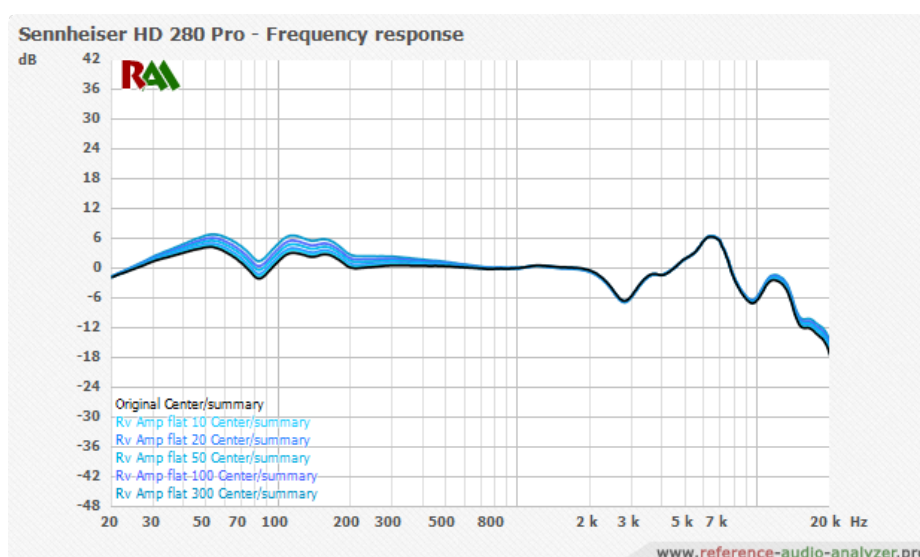


Figura 74. Respuesta en frecuencia audífonos Sennheiser HD 280 pro. Tomado de la página oficial de Sennheiser.

Headphone HD 280 PRO

Acoustic principle	dynamic, closed
Impedance	64 Ω
Ear coupling	circumaural
Frequency response	8 - 25,000 Hz
Ambient noise attenuation	≤ 32 dB
THD	<0.1% (@ 1 kHz, 100 dB SPL)
Weight (w/o cable)	approx. 285 g
Contact pressure	approx. 6 N
Length of coiled cable	≥ 1.3 m / ≤ 3 m
Operating temperature	-10 °C to +55 °C
Storage temperature	-20 °C to +70 °C
Relative humidity	$\leq 90\%$ @ 40 °C

Figura 75. Especificaciones técnicas de audífonos Sennheiser HD 280 pro. Tomado de la página oficial de Sennheiser.

