



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA Y SISTEMAS DE TELECOMUNICACIÓN

PROYECTO FIN DE GRADO

TÍTULO: Desarrollo de un Protocolo para la Realización de Experimentos de Percepción Subjetiva

AUTOR: Víctor de Tejada Molera

TITULACIÓN: Ingeniería de Sonido e Imagen

TUTOR: Rubén Fraile Muñoz

DEPARTAMENTO: Ingeniería Telemática y Electrónica

VºBº

Miembros del Tribunal Calificador:

PRESIDENTE: Gerardo Pérez Villalon

TUTOR: Rubén Fraile Muñoz

SECRETARIO: Antonio Pedrero González

Fecha de lectura: 28 de octubre, 2020

Calificación:

El Secretario,

Resumen

La psicoacústica es el campo que estudia cómo percibimos los seres humanos el sonido. Esto se realiza mediante la creación de experimentos subjetivos de audio donde un determinado número de personas realizan, en situaciones controladas, una serie de pruebas con las que se extraen unas ciertas conclusiones.

Desafortunadamente, no existe un método homogéneo para realizar los experimentos. Existen recomendaciones provenientes de organismos internacionales como la EBU (*European Broadcasting Union*) o la UIT (*Unión Internacional de Telecomunicaciones*), pero sólo aportan información para casos muy específicos y técnicos.

Por otro lado, si se analiza una muestra de diferentes estudios, se aprecia una gran variedad en los procedimientos seguidos, el número de personas utilizadas o el análisis de los datos (en algunos casos ni siquiera se mencionan qué consideraciones se han tomado). Esto no hace más que poner de manifiesto la necesidad de elaborar un protocolo para la creación de experimentos subjetivos de audio con los que establecer una serie de recomendaciones mínimas para los tipos de test más comunes, que siga, dentro de lo posible, las recomendaciones de los organismos internacionales y que sea fácilmente interpretable y ampliable.

Para lograr este objetivo, se realiza un primer lugar un trabajo para definir los elementos que aparecen en los experimentos psicoacústicos más comunes y se analiza una amplia selección de bibliografía relacionada con este campo. Con este análisis se extraen los elementos más comunes y se establece una división de los estudios subjetivos de audio en cinco categorías: Inteligibilidad, sensibilidad, reverberación, localización y molestia. Dentro de cada una de estas categorías se establecen unas recomendaciones generales, comunes a todos los participantes, y otras específicas atendiendo al número de participantes necesarios y su experiencia, la duración del experimento, el tipo de señal a utilizar, el tipo de test o el tipo de análisis estadístico, entre otros.

Después, se realiza un ejemplo de aplicación de protocolo con un experimento que tiene como objetivo determinar si la percepción del número de cantantes en una pieza depende más de las variaciones en frecuencia o de las variaciones en amplitud. Ésto se realiza haciendo comparaciones de audios naturales con audios artificiales que simulan dichas variaciones. Los resultados se comparan con los obtenidos en un estudio similar realizado anteriormente y se observa que se obtienen resultados similares, pero el estudio realizado con el protocolo profundiza más en el análisis estadístico de los datos y se obtienen mayores conclusiones.

Con esto, queda comprobada la utilidad del protocolo y se dejan abiertas nuevas líneas de investigación como la creación de subcategorías dentro de las ya creadas y profundizar más dentro de los diferentes tipos de test estadístico que pueden utilizarse para el análisis de los datos.

Abstract

Psychoacoustics is the field that studies how humans perceive sound. This is done through the creation of subjective audio experiments where a determined number of people perform, under controlled situations, a series of tests where certain conclusions are obtained.

Unfortunately, there is no a common method to perform these experiments. Recommendations exist from international organizations such as the EBU (*European Broadcasting Union*) or the ITU (International Telecommunication Union), but they only provide information for very specific and technical scenarios.

On the other hand, if a sample of different studies is analyzed, lots of differences can be detected in the procedures followed, the number of people who participate or the data analysis (in some cases it is not even mentioned what considerations have been made). This only highlights the necessity to develop a protocol for the creation of subjective audio experiments. This protocol establishes a series of minimum recommendations for the most common types of tests, which follow, as much as possible, the recommendations from international organizations that also are easily interpretable and expandable.

In order to achieve this objective, the first task is to define the elements that appear in the most common psychoacoustic experiments and a wide selection of literature related to this field is analyzed. Then the most common elements are extracted and a division for the subjective audio studies is established: intelligibility, sensitivity, reverberation, localization and annoyance. Within each of these categories, some general recommendations and other specific ones are established according to the number of participants required and their experience, the length of the experiment, the type of signal used, the type of test or the type of statistical analysis, among others.

Then, an example of protocol application is made with an experiment that tries to determine if the perception of the number of singers in a piece depends more with frequency variations or with amplitude variations. This is done by making comparisons of natural audios with artificial audios that simulate such variations. The results are compared with those obtained in a similar study carried out previously and it is observed that similar results are obtained, but the study done with the protocol has a deeper statistical analysis of the data and greater conclusions are obtained.

This proves the utility of the protocol and open future lines of research such as the creation of subcategories in those already made and to go deeper into the different types of statistical tests that can be used for data analysis.

Índice general

1. Introducción	9
2. Antecedentes	11
3. Especificaciones y restricciones de diseño	13
4. Definiciones	15
4.1. Definiciones de elementos psicoacústicos	15
4.2. Definiciones de categorías	17
4.3. Definiciones de tipos de tipos de test	18
4.4. Definiciones de tipos de análisis de datos	19
5. Búsqueda de información	21
5.1. Fuentes de información	21
5.2. Criterio y selección de estudios	22
6. Resumen de los estudios seleccionados	25
6.1. Recomendaciones y normativas	25
6.2. Artículos de revistas científicas y conferencias de congresos	26
6.3. Libros de referencia	29
7. Clasificación de los estudios	31
7.1. Inteligibilidad	31
7.2. Sensibilidad	31
7.3. Reverberación	31
7.4. Localización	31
7.5. Molestia	34
7.6. Recomendaciones y normativas	34
8. Diseño del protocolo	37
8.1. Recomendaciones generales	37
8.2. Identificación de los rasgos del ensayo	38
8.3. Recomendaciones específicas	39
8.4. Criterio para la toma de decisiones	42
9. Justificación de las recomendaciones	43

10. Ejemplo de aplicación del protocolo	47
10.1. Objetivo del estudio	47
10.2. Características del estudio según el protocolo	47
10.3. Procedimiento del estudio	49
10.4. Material del estudio	49
10.5. Resultados del estudio	50
10.5.1. Resultados de la parte 1	50
10.5.2. Resultados de la parte 2	50
10.6. Análisis de los resultados	52
10.6.1. Parte 1	52
10.6.2. Parte 2	54
10.7. Conclusiones del estudio	55
11. Comparación con el estudio original	57
11.1. Características del estudio original	57
11.2. Comparación de resultados	58
12. Conclusiones	59
A. Resultados completos de la parte 1 del experimento	61
B. Resultados completos de la parte 2 del experimento	71
C. Análisis de costes	91

Índice de figuras

4.1. Relación de tipos de categorías con elementos psicoacústicos	18
10.1. Número de oyentes percibidos en función del número de oyentes teóricos para audios modificados en amplitud.	50
10.2. Número de oyentes percibidos en función del número de oyentes teóricos para audios modificados tonalmente.	51
10.3. Número de oyentes percibidos en función del número de oyentes teóricos para audios naturales.	51
10.4. Respuestas acumuladas para la percepción de qué tipo de audio dispone de más cantantes al compararse dos tipos de audio.	52

Índice de cuadros

7.1. Características de los estudios de inteligibilidad acústica.	32
7.2. Características de los estudios de sensibilidad acústica.	33
7.3. Características de los estudios de reverberación acústica.	34
7.4. Características de los estudios de localización acústica.	35
7.5. Características de los estudios de molestia auditiva.	35
7.6. Características de las normativas y recomendaciones.	36
8.1. Características de los estudios de sensibilidad acústica.	42
10.1. Valores de <i>Alpha-Risk</i> en función del número de personas y el número de respuestas positivas.	49
10.2. Número de respuestas seleccionadas para cada tipo de audio cuando es comparada con otro de los tipos.	52
10.3. Resultado del análisis de resultados de la parte 1 por el método Wilcoxon.	53
10.4. Número de respuestas para cada tipo de audio cuando es comparado con otro.	54
10.5. Numero de veces que debe ser escogida una opción necesarias para obtener un determi- nado <i>Alpha-Risk</i> teniendo un total de 115 comparaciones.	55
10.6. Número de veces que una opción es escogida en una comparación suprimiendo las res- puestas dudosas y cálculo de los valores necesarios para distintos valores de <i>Alpha-Risk</i>	55
11.1. Comparación entre estudio original y siguiendo protocolo	57
C.1. Costes asociados a las recursos utilizados en el proyecto	92

Lista de acrónimos y símbolos

A continuación, se presenta una lista de los acrónimos y símbolos utilizados en orden alfabético:

- 2-AFC: *Two-Alternative Forced Choice*, traducido como sistema de elección forzada entre dos opciones.
- Alcons%: *Articulation Loss of Consonants*, traducido al castellano como Índice de Pérdida de Articulación de Consonantes.
- ANOVA: *Analysis of Variance*, sistema de Análisis de la varianza en castellano.
- α : También llamado “*alpha-risk*”. Es la probabilidad de concluir que una diferencia perceptible existe cuando en realidad no es así.
- β : Se le llama también “*beta-risk*”. Es la probabilidad de concluir que una diferencia no perceptible existe cuando en realidad no es así.
- C_{80} : Índice de Claridad sonora a 80 milisegundos.
- D_{50} : Índice de Definición sonora a 50 milisegundos.
- EBU: *European Broadcasting Union*. En castellano se traduce como Unión Europea de Radiodifusión.
- ICA: *International Congress of Acoustics*.
- ISO: *International Organization for Standardization*. Organismo internacional de estandarización en español.
- ITU: *International Telecommunication Union*. En castellano se traduce como Unión Internacional de Telecomunicaciones (UIT).
- log: Logaritmo decimal (en base 10).
- L_p : Nivel de presión sonora.
- MUSHRA: *MULTi Stimulus text with Hidden Reference*, traducido al castellano como test multi-estímulo con referencia oculta.
- P_1 : Presión medida en un punto.
- P_0 : Presión de referencia.
- SAQI: Spatial Audio Quality Inventory.
- STI: *Speech Transmission Index*, traducido generalmente como índice de Transmisión de Discurso.
- UIT-R BS: Recomendación de la Unión Internacional de Telecomunicaciones para el ámbito de sistemas de radiodifusión. Viene acompañado por un número que identifica a la recomendación concreta.
- UNE-EN: Una Norma Española - Asociación Española de Normalización. Al igual que las UIT-R, vienen acompañadas por un número que identifica la norma concreta.
- UPM: Universidad Politécnica de Madrid.

Capítulo 1

Introducción

La psicoacústica es una rama de la acústica que pretende analizar cómo percibimos los seres humanos el sonido. La forma habitual para medirlo consiste en la realización de test auditivos.

A la hora de realizar dichos test hay que realizar un muestreo de la población con el que se desarrolla un estudio estadístico. Esto lleva a que una persona que pretenda realizar un experimento en este campo deba resolver previamente una serie de cuestiones como: ¿cuántas personas debo escoger?, ¿de qué edad?, ¿deben estar entrenadas en este tipo de test?, ¿durante cuánto tiempo pueden estar realizando las escuchas antes de presentar fatiga?

Para resolver este tipo de cuestiones, se pretende realizar un protocolo para las mediciones subjetivas de audio que sirva de guía para futuros estudios. En él se dan respuesta, entre otras, a las preguntas ya mencionadas para los tipos de experimentos subjetivos más comunes que puedan darse.

Este objetivo se consigue, en primer lugar, realizando un estudio del “estado del arte”. Se revisan distintos tipos de documentos como libros, normativas, recomendaciones, artículos de revistas científicas o conferencias, etc. De esta forma se consiguen extraer los elementos clave que conforman los estudios subjetivos de audio.

A continuación, tras establecer las definiciones mínimas necesarias para fijar las bases del protocolo, se establece una división en categorías de los diferentes tipos de test subjetivos de audio. Dentro de cada categoría, se establecen una serie de recomendaciones, en primer lugar generales y después específicas, para los elementos más importantes para la realización de los experimentos. Algunos ejemplos son la selección de los participantes, los periodos de excitación, la duración del experimento, el tipo de test a realizar, el análisis de los datos, etc.

Todas las recomendaciones siguen, en la medida de lo posible, las recomendaciones de los organismos internacionales y están en concordancia con los procedimientos seguidos por los estudios analizados. Para ello, se emplea un capítulo íntegramente en justificar las decisiones que se han tomado en la redacción del protocolo.

Una vez realizado el protocolo, se comprueba la eficacia del mismo realizando un test subjetivo de audio sobre una de las categorías definidas. El experimento sigue las recomendaciones establecidas anteriormente y después se comprueban los resultados obtenidos con otro test con un objetivo similar. De esta forma, quedan reflejadas las fortalezas del protocolo y se comentan posibles líneas de actuación para mejorarlo en un futuro.

Capítulo 2

Antecedentes

A la hora de analizar el estado del arte, se ha observado una ausencia clara de estandarización dentro de los procedimientos necesarios para realizar estudios subjetivos de audio.

En primer lugar, se han observado algunas recomendaciones provenientes de algunos organismos internacionales como la ITU o la EBU ([1], [2] y [3]). Por lo general, aportan información sobre aspectos o áreas muy concretas de los estudios psicoacústicos y no entran a analizar las diferentes realidades que pueden encontrarse en su realización. Esto impide que sean fáciles de aplicar para la mayoría de las áreas de estudio que pueden aplicarse.

Por otro lado, en los libros de referencia consultados, la información que se recoge sí que suele disponer de información más general sobre la realización de estudios psicoacústicos, pero suelen estar centrados en aspectos específicos de la escucha humana como la fatiga auditiva [4] y [5]. En otros casos, la información se centra más en definir los elementos psicoacústicos más que en dar recomendaciones sobre cómo medirlos o aplicarlos [6] y [7].

Por último, al analizar diferentes estudios procedentes de revistas científicas o congresos internacionales, se observa que los procedimientos seguidos por las personas que lo realizan son muy dispares entre sí. En algunos casos no se encuentra información sobre las características de los participantes (la edad, su experiencia, etc.). En otros muchos casos, no se especifica el tipo de análisis de datos que se ha realizado, la duración de los experimentos o cómo ha sido la interacción de los oyentes con el propio experimento. Algunas de esta gran variedad de casos pueden observarse en ([8], [9] y [10], entre otros).

Estas inconsistencias, y la ausencia de una fuente única a la que acudir a la hora de preparar un ensayo psicoacústico, hacen ver la relevancia de crear un documento que aglutine las características más recomendables para la realización de este tipo de experimentos, de forma que incluya las características más generales que deben reflejarse, pero profundice también en aspectos más específicos para los tipos de test más comunes.

Capítulo 3

Especificaciones y restricciones de diseño

Como el ámbito de los estudios subjetivos de audio es tan amplio, conviene determinar una serie de especificaciones y límites dentro del diseño del protocolo. En este caso, se ha decidido establecer las siguientes consideraciones:

En primer lugar, se decide que el protocolo debe englobar los tipos de medidas subjetivas de audio más comunes basados en test estadísticos. De esta forma, se consigue que nuestro estudio ofrezca soluciones para las áreas donde se están realizando actualmente la mayoría de los estudios psicoacústicos.

Por otro lado, el protocolo debe seguir, dentro de lo posible, las recomendaciones que los organismos internacionales como la ITU (*International Telecommunication Union*) o la EBU (*European Broadcasting Union*) tienen publicadas al respecto. Así se consigue que el procedimiento seguido por un futuro estudio siga un procedimiento normalizado acorde con estas organizaciones.

También se considera que este protocolo debe ser fácilmente interpretable en futuros estudios. La información se presenta de forma sencilla según cada tipo de experimento y además permite que se pueda seguir desarrollando los campos acción del protocolo en futuras ampliaciones.

Por último, se especifica que este protocolo debe contener toda la información necesaria para que, en sí mismo, permita diseñar los experimentos y evaluar sus resultados sin que sea necesario acudir a otras fuentes.

Capítulo 4

Definiciones

A la hora de desarrollar el protocolo, es necesario establecer, en primer lugar, una serie de definiciones con las que especificar cada uno de los elementos que se evalúan en él. De esta forma, se establecen cuatro grupos de definiciones: definiciones de elementos psicoacústicos, definiciones de tipos de test, definiciones de categorías de estudios psicoacústicos y definiciones de los diferentes tipos de análisis de datos que pueden utilizarse.

4.1. Definiciones de elementos psicoacústicos

En este capítulo se definen aquellos elementos que pueden ser sujetos de medición dentro de un estudio psicoacústico.

Nivel sonoro

También referido como nivel de presión sonora. Se define bajo la ecuación (4.1):

$$L_p = 20 * \log\left(\frac{P_1}{P_0}\right) \quad (4.1)$$

Siendo P_1 la presión medida en el punto de estudio y P_0 la presión de referencia (típicamente $20[\mu Pa]$).

Sonoridad

La sonoridad es la percepción subjetiva sobre si un determinado sonido suena “más alto o más bajo”. A esta percepción se la conoce coloquialmente como ”volumen”.

Inteligibilidad

La inteligibilidad es la capacidad para comprender las palabras que se están transmitiendo. Se suele medir cuantitativamente para salas mediante los índices *STI* y *Alcons %*.

Por un lado, *STI* se define como el índice de transmisión de discursos, o *Speech Transmission Index* en inglés. Este índice divide la inteligibilidad de una sala en varias categorías en función de si es más o menos óptimo para la transmisión dentro de una escala de 0 a 1. Siendo 0 una muy mala inteligibilidad y 1 una inteligibilidad excelente.

Por otro lado, *Alcons %* es el índice de pérdida de articulación de consonantes, *Articulation Loss of Consonants* en inglés. En este índice la inteligibilidad se expresa en forma de porcentaje. Siendo 0 % la máxima inteligibilidad y 100 % la mínima. Al igual que con el índice *STI*, se divide el rango en zonas cualitativas de inteligibilidad.

Claridad

La claridad se define como la capacidad de distinguir sucesos sonoros muy próximos entre sí. Existen varios índices para tratar de dar un valor cuantitativo a este elemento. Uno de los más habituales es el C_{80} que mide la diferencia de energía sonora entre los primeros 80ms y la energía del sonido reverberante a partir de dicho instante de tiempo. Este criterio se define para elementos distintos a la palabra.

Definición

La definición es la capacidad de distinguir cada uno de los elementos de un discurso. Al igual que con la claridad, se utilizan muchos índices para medirlo. Uno de los más habituales es el llamado D_{50} que mide la diferencia de energía sonora de los primeros 50ms frente al resto de la energía reverberante. Se usa principalmente para la palabra.

Por expresarlo de otra forma, la Definición evaluaría situaciones en las que se realizan conferencias, ponencias y presentaciones; mientras que la Claridad evaluaría situaciones de conciertos, películas, etc., donde la prioridad no es tanto si podemos distinguir cada una de las palabras, sino si podemos distinguir los instrumentos y/o efectos que componen el evento sonoro.

Similitud

Percepción subjetiva en la cual dos señales sonoras se consideran más o menos iguales entre sí.

Degradación sonora

Es la pérdida de calidad auditiva que sufre una señal. Esto puede deberse a la imperfección del sistema de transmisión, a las compresiones del sistema de grabación, la distancia entre el emisor y el receptor, etc.

Un ejemplo de Degradación sonora sería la producida al grabar un instrumento musical mediante un micrófono con una respuesta que no es plana en todo el espectro sonoro, ser codificada por un sistema *lossy* (con pérdidas) y ser reproducido de nuevo con un altavoz con una respuesta en frecuencia no plana.

Tiempo de reverberación

El tiempo de reverberación se define como la cantidad de tiempo que transcurre entre el instante en el que una fuente sonora deja de emitir hasta que el nivel de presión sonora decae 60[dB] respecto del que había con la fuente en funcionamiento.

Reverberación

La reverberación es la sensación de persistencia de un sonido en un determinado ambiente o sala. No debe confundirse con el tiempo de reverberación definido anteriormente.

Sonido anecoico

Un sonido anecoico es aquella señal sonora que no contiene información del entorno en el que se encuentra.

Sonido reverberante

Un sonido reverberante es aquella señal sonora que contiene la información de la señal directa y del entorno en el que se encuentra.

Fatiga auditiva

Sensación que se produce al exponerse a uno o varios sucesos sonoros durante un periodo prolongado de tiempo. Esta sensación puede causar que el oyente deje de percibir los sucesos sonoros de forma correcta produciendo resultados no válidos. Es, por tanto, una situación a evitar y habrá que establecer medidas para evitarlo como pausas, limitar tiempos de exposición, etc.

Molestia auditiva

Aversión subjetiva que produce un suceso auditivo al ser captada por un oyente. Al contrario que la fatiga auditiva, no es causada necesariamente por el tiempo de exposición, pero un evento “no molesto” puede acabar convirtiéndose en uno debido a ello.

4.2. Definiciones de categorías

Los diferentes elementos psicoacústicos definidos anteriormente quedan agrupados dentro de diferentes categorías. En este apartado se definen las diferentes categorías que se han formado tras el estudio del estado del arte.

Inteligibilidad

Se define la categoría de Inteligibilidad como aquella que engloba todos los estudios psicoacústicos que tienen como objetivo el estudio de la comprensión de la información sonora; ya sea para la palabra, la música u alguna otra forma. Dentro de esta categoría quedan incluidos todos aquellos experimentos que pretendan estudiar comportamientos relacionados con la claridad, la definición o la propia inteligibilidad en las posibles situaciones que puedan darse.

Sensibilidad

Esta categoría se define como aquella que engloba los estudios que pretenden estudiar las diferencias cuantitativas entre dos o más sucesos sonoros, a priori, similares. De esta forma, quedan incluidos los proyectos cuyo objeto de estudio sean las diferencias de nivel sonoro, sonoridad, similitud, degradaciones, etc.

Reverberación

Esta categoría es la que engloba aquellos experimentos que pretenden estudiar aspectos relacionados con la sensación de persistencia de un sonido hasta que desaparece. Quedan incluidos los estudios de percepción de duración de una señal, tiempos de reverberación y diferencias de tiempo entre sonidos anecoicos y reverberantes, entre otros.

Localización

Es la categoría que aúna los diferentes proyectos que tienen como foco de estudio la procedencia y posición de determinados estímulos sonoros.

Molestia

Esta categoría incluye a los estudios psicoacústicos que pretenden evaluar la molestia auditiva que se producen por ciertos sonidos o situaciones sonoras.

Como puede deducirse, ambos tipos de definiciones están relacionados entre sí. Estas relaciones pueden apreciarse de forma más sencilla en la figura 4.1.

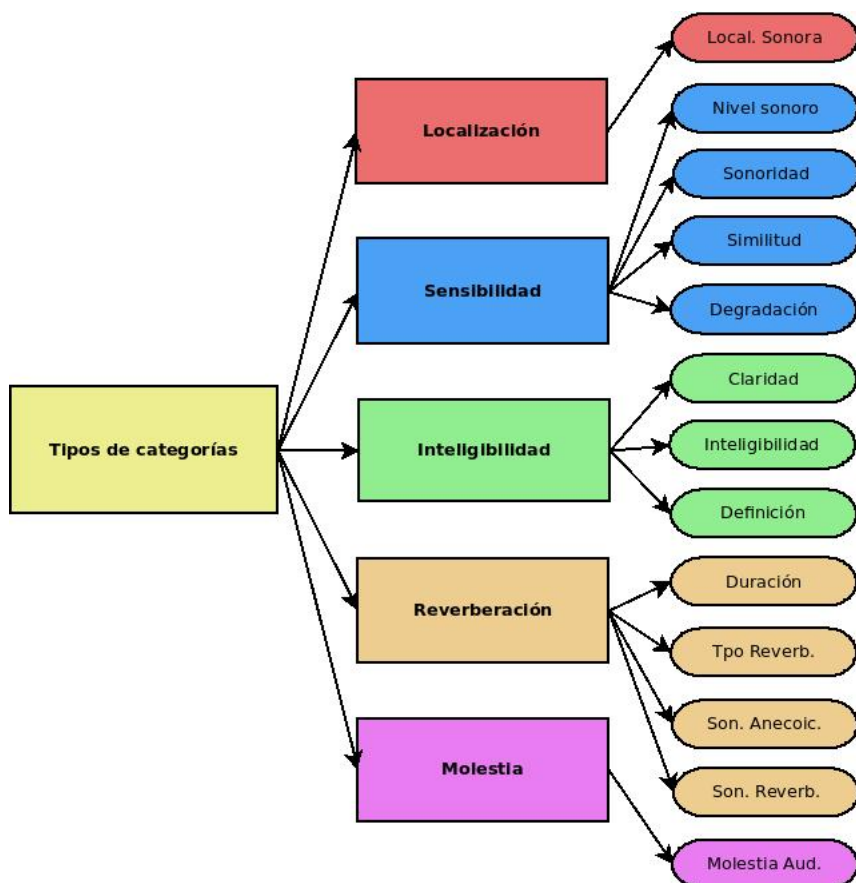


Figura 4.1: Relación de tipos de categorías con elementos psicoacústicos

4.3. Definiciones de tipos de tipos de test

Los tipos de test son los diferentes procedimientos que pueden utilizarse a la hora de recoger la información en un determinado estudio. Cada tipo de test tiene sus fortalezas y debilidades, y de ellas dependerá que sean más o menos apropiadas para un experimento u otro.

En este apartado se describen los más habituales.

Test binomial

Engloba todos aquellos tipos de test en los que la respuesta está entre dos opciones: Sí/No, A/B, Bien/Mal, etc. Este tipo de test son útiles cuando se quiere saber si se da o no una determinada condición, pero no interesa recoger aspectos intermedios. Por el contrario, no son apropiados cuando

se quiere tener un gran abanico de opciones. Un tipo de test de esta clase muy habitual en estudios psicoacústicos es el denominado 2-AFC.

- 2-AFC: Son las siglas del test *Two-Alternative Forced Choice*. Consiste en dar al usuario dos estímulos y “forzarle” a escoger entre una de las dos opciones según lo que se pretenda estudiar. Las condiciones de orden, repetición o comentarios post-respuesta dependen del experimento en sí.

Test de escalas numéricas

Los test de escalas numéricas son aquellos en los que el participante del experimento asigna un valor comprendido dentro de una escala a una o varias características psicoacústicas.

Las escalas pueden ser continuas o discretas; enteramente positivas o centradas en cero con valores positivos y negativos. Este último tipo es especialmente útil cuando se quiere hacer una comparación entre dos elementos complementarios.

Por otro lado, los tests pueden hacerse de forma secuencial; evaluando diferentes aspectos uno detrás de otro, o bien se pueden evaluar en paralelo. Esta forma suele denominarse como la técnica *MUSHRA*. Dicha técnica es de las más utilizadas para ciertos estudios subjetivos de audio tal y como recomiendan las normas ISO.

- MUSHRA: Son las siglas de *MUlti Stimulus test with Hidden Reference and Anchor*. Según [2] en este método, el sujeto del estudio puede pasar de una señal de referencia a cualquiera de los estímulos del test. Después evalúa dicho estímulo dentro de una escala de 0 a 100. Dentro de las señales de prueba se encuentra de nuevo la señal de referencia (referencia oculta) y dos señales de patrón (patrón medio y bajo). Ambas son una versión filtrada en paso bajo de la señal original con una frecuencia de corte de 7 y 3.5kHz respectivamente.

Test de posiciones

Este tipo de test está especialmente diseñado para cuando se realizan estudios de localización sonora. A los sujetos se les facilita un sistema en el que pueden escoger una determinada posición. De esta forma, se puede obtener la desviación en grados respecto al centro de referencia o el altavoz del que se percibe que proviene una señal, por poner algunos ejemplos.

4.4. Definiciones de tipos de análisis de datos

Una vez que se ha realizado un tipo de test, hay que realizar un análisis de los datos obtenidos con el fin de extraer las conclusiones pertinentes del estudio realizado. En este apartado se resumen los sistemas más habituales para llevarlo a cabo.

Estadística descriptiva

La estadística descriptiva es aquella que se utiliza para describir o resumir información característica de una determinada muestra de datos. Para ello, se vale de herramientas tales como la media aritmética, media geométrica, la varianza, la desviación estándar y de gráficas como los histogramas.

Estadística paramétrica

La estadística paramétrica es aquella que se utiliza cuando se presupone que una determinada muestra de datos sigue una determinada distribución de probabilidad determinados por un número finito de parámetros. Dentro de ella, uno de los sistemas más habituales el Sistema ANOVA.

- Sistemas ANOVA (*Analysis of Variance*): Son un conjunto de modelos estadísticos que permiten estudiar las diferencias entre las medias de diferentes grupos de muestras. Generalmente se utilizan cuando las muestras siguen una distribución normal.
 - T-test: Es un tipo de test estadístico que se utiliza para comparaciones de pares de experimentos cuando ambos tienen la misma, aunque desconocida, varianza.
 - F-test (ANOVA): Es un tipo de test estadístico que compara los diferentes factores que producen las desviaciones en diferentes experimentos.

Estadística no paramétrica

La estadística no paramétrica, al contrario que la paramétrica, no asume que una muestra de datos siga alguna distribución de probabilidad, o en caso de seguir alguna, no determina qué parámetros la determinan. Dentro de esta categoría, existen numeros métodos. Algunos de los más utilizados son:

- Wilcoxon signed-ranked test: Es un tipo de test estadístico que se utiliza para comparaciones de pares de experimentos basándose en categorías (rankings); dicha comparación se realiza se calculando la media de ambos experimentos y se compara la desviación de uno con la de otro. Después, se toman diferentes combinaciones de los valores individuales para cada experimento y se vuelven a comparar las medias. De esta forma se determina si esas variaciones son significativas o no.
- Estadística esférica: Es un tipo de estadística que se centra en el estudio de direcciones dentro de distribuciones esféricas.

Capítulo 5

Búsqueda de información

Para la realización del protocolo es necesario realizar una búsqueda considerable de información. Para ello, conviene explicar las fuentes de información que se han usado, el criterio que se ha seguido para seleccionar dichas fuentes y cómo se han seleccionado. En este capítulo se analizan todas estas cuestiones.

5.1. Fuentes de información

Para el desarrollo del protocolo se han utilizado, principalmente, dos sistemas de búsqueda de información:

Por un lado, se ha utilizado el servicio *Google Scholar*. Se trata de un buscador de documentos técnicos y científicos. Este sistema resulta muy útil ya que se pueden realizar búsquedas por términos, tipos de documentos, etc. Lo que permite un primer filtro de información de forma sencilla.

Por otro lado, se ha utilizado el sistema *Ingenio* de la Universidad Politécnica de Madrid (UPM). Este sistema también permite filtrar resultados por tipo de documentos, fecha, términos, campos de estudio, etc. Además cuenta con la ventaja de que indica si es posible consultar físicamente dichos documentos y, al ser un servicio interno de la UPM, se tiene la seguridad de que se va a poder acceder al documento en cuestión.

Utilizando ambos sistemas, se han consultado cuatro tipos de documentos:

- Normas y recomendaciones: están escritas por organizaciones internacionales. Determinan criterios para realizar de forma correcta los experimentos psicoacústicos y, por tanto, se tratan de documentos muy necesarios para el protocolo. En nuestro caso de estudio, se han analizado las recomendaciones de la *International Telecommunication Union* (ITU) y la *European Broadcasting Union* (EBU).
- Libros técnicos: contienen información muy contrastada y tienen una gran fiabilidad. Por el contrario, son documentos que suelen carecer de concreción y son menos recientes que otros documentos.
- Artículos de revistas indexadas: son el siguiente eslabón en cuanto a fiabilidad y suelen abordar temas más concretos y recientes que los libros técnicos.
- Conferencias de congresos: contienen la información más concreta y reciente. Permiten analizar mejor cómo se están realizando los estudios en la actualidad. Por contra, son los documentos que menos fiabilidad tienen y hay que analizarlos más detenidamente.

5.2. Criterio y selección de estudios

A la hora de realizar una selección de las fuentes consultadas, se han seguido criterios distintos en función del tipo de documento puesto que la información que presentan es distinta en cada caso.

Libros

El criterio para la selección de libros viene determinado por sus definiciones de conceptos acústicos, los tipos de ensayos que proponen y los límites que sugieren para evitar la fatiga auditiva. Por otro lado, también se han tenido en cuenta aquellas fuentes que proporcionan información sobre los distintos tipos de modelos estadísticos.

Para su búsqueda, se han introducido en *Google Scholar* y el sistema *Ingenio* términos como “Auditory Perception Methods”, “Auditory Fatigue”, “Psychoacoustics”, etc.

En este tipo de documentos, los ejemplos de estudios concretos no son muy habituales por lo que no se ha considerado una prioridad a la hora de seleccionarlos.

De esta forma se han conseguido encontrar 4 libros: [4], [6], [7] y [11].

Artículos de revistas y conferencias de congresos

Con estos dos tipos de documentos se ha seguido el mismo criterio. Para obtener estos documentos se ha procedido a realizar búsquedas en Google Scholar de documentos con términos como “Auditory Perception”, “Auditory Fatigue”, “Subjective test” o “Psychoacoustics”. De esta forma, se han obtenido más de 50 resultados.

Después, se han leído los resúmenes de los estudios para confirmar que aplicaban a los campos deseados, ya sea directamente desde Google Scholar o desde la plataforma Ingenio UPM si no era posible. También se revisaba las referencias de dichos estudios para ver si hacían mención a otros que pudieran ser de interés. Por poner un ejemplo, varios artículos han sido extraídos de [12]. Al finalizar este proceso, se ha conseguido reducir el número a un total de 27 casos.

Debido a la enorme variedad de estudios subjetivos de audio, se ha optado por analizar su contenido en varios grupos. Estos grupos aportan información específica sobre los aspectos que se han creído más relevantes para la realización del protocolo. De este modo, se consigue filtrar más rápidamente los elementos válidos de los no válidos. Los grupos que se han considerado son los siguientes:

- Medida: determina que aspecto psicoacústico se ha pretendido medir en el estudio.
- Tipo test: determina el tipo de test subjetivo que se ha utilizado durante el estudio.
- N° de participantes: determina el número de personas que han participado en el test. Si el estudio está dividido en varias partes, se especifica el número de personas para cada una de ellas.
- Tipo de oyente: determina si los oyentes que participan en el estudio son expertos o no expertos en la realización de test subjetivos de audio.
- Duración de la señal: determina cuánto dura la señal que se ha utilizado durante el test.
- Duración de la sesión: determina cuánto tiempo ha durado una sesión del estudio para un determinado participante. También se incluye la duración de los descansos, si los hubiera.
- Entrenamiento previo: determina si los participantes del test han realizado algún entrenamiento previo a la realización del test.

- Análisis de datos: determina el tipo de análisis estadístico que se ha utilizado con los datos obtenidos del test.
- Tipo de estímulo: determina qué señal o tipo de señal se ha utilizado en la realización del test.
- Se muestra una sola vez, lo modifica el usuario, etc: da información sobre el rol que han tenido los participantes durante el test. Algunos ejemplos pueden ser: la posibilidad de repetir las señales por parte del participante, el número de veces que puede hacerlo, si puede escoger el orden de reproducción de las mismas, etc.

De todos los estudios que han pasado estos filtros, algunos no han podido utilizarse dentro del protocolo ya que, pese a tratar de los temas adecuados, no incluían la información suficiente de los campos indicados. Es el caso de [13], [14] y [15]. De esta forma, el número final de artículos de revistas y conferencias de congresos quedó reducido a 24 casos válidos.

Normativas y recomendaciones

Para este último tipo de documentos, se ha seguido el mismo criterio de búsqueda que para los artículos de revistas científicas y conferencias de congresos añadiendo los campos de “Normativa” y “Recomendación”. A continuación se leía el resumen de las mismas para seleccionar los que correspondían a los temas implicados en el protocolo obteniéndose de esta forma 7 documentos; siendo uno de ellos el listado de la UIT de normativas para mediciones subjetivas de audio [16].

Después, se ha dividido su información en los mismos campos que en los artículos y conferencias. Esto tiene un doble propósito: homogeneizar la información por un lado y simplificar el contenido de la misma para poder extraer datos más fácilmente para la redacción del protocolo. Durante este procedimiento, se descartaron la recomendaciones [3], [17] y [18] porque toda la información de ellas hacía referencia a otro de los documentos y no habría aportado información nueva, o bien ya habían sido suprimidas por el organismo. De esta forma, se reduce la lista definitiva a 3 documentos: [1], [2] y [19].

Durante partes más avanzadas del proyecto, se encontró una normativa relacionada con el análisis de datos en los ensayos que analizan diferencias y similitudes entre dos o tres muestras [20]. Esto sería de utilidad en la parte de aplicación del protocolo y también fue incluido.

Capítulo 6

Resumen de los estudios seleccionados

En este capítulo se resumirán brevemente cada uno de los estudios seleccionados. Dichos resúmenes se harán atendiendo a la clasificación de los documentos expuesta en el capítulo anterior.

6.1. Recomendaciones y normativas

- *Recomendación UIT-R BS.1116-3: Métodos para la Evaluación Subjetiva de Pequeñas Degradaciones en los Sistemas de Audio* [1]: en este documento se hacen recomendaciones sobre cómo proceder a la hora de realizar estudios subjetivos sobre pequeñas degradaciones. En él se sugiere utilizar escalas numéricas; un número de participantes de mínimo 20 personas, expertas y con entrenamiento previo. Los participantes deben poder conmutar los diferentes estudios de forma individual. La duración de la señal debe estar entre 15 y 25 segundos de duración. La duración de la sesión no debe sobrepasar los 30 minutos y no realizarse más de 15 experimentos por sesión. El análisis de los datos se recomienda hacer mediante sistemas ANOVA.
- *Recomendación UIT-R BS.1534-3: Método para la Evaluación Subjetiva del Nivel de Calidad Intermedia de los Sistemas de Audio* [2]: en este documento se hacen recomendaciones generales para el estudio subjetivo de sistemas de calidad intermedia. Se sugiere utilizar el sistema *MUSH-RA*; un número de participantes de mínimo 20 personas, expertas y con entrenamiento previo. Los oyentes deben poder pasar de la señal de referencia a cualquier otra señal. La duración de la señal debe estar comprendida entre los 10 y los 12 segundos. La duración de la sesión sigue las mismas reglas que [1]. Lo mismo ocurre con el análisis de los datos.
- *Recomendación UIT-R BS.1284-2: Métodos Generales para la Evaluación Subjetiva de la Calidad de Sonido* [19]: este documento da unas pautas generales para la evaluación subjetiva de la calidad del sonido. Se sugiere la utilización de escalas numéricas; un número de participantes de mínimo 20 personas si no son expertos y de 10 si lo son. En ambos casos se recomienda el uso de un entrenamiento previo. Los participantes deben poder conmutar entre las señales. La duración de la señal debe estar comprendida entre los 15 y los 20 segundos, mientras que la duración de la sesión entre los 15 y 20 minutos. En caso de necesitar que sean más largas, deben hacerse pausas de la misma duración. El análisis de datos debe hacerse mediante medias e intervalos de confianza.
- *Norma UNE-EN ISO 10399:2018: Análisis sensorial. Metodología. Ensayo dúo-trío* [20]: en esta normativa se presentan el número de personas necesario, respecto del total, para que se considere que existe una diferencia o similitud perceptible entre dos o tres muestras. La probabilidad de

que esa diferencia o similitud exista se mide con las variables α y β respectivamente, y su valor depende tanto del número de personas participantes del ensayo, como del número de respuestas positivas.

6.2. Artículos de revistas científicas y conferencias de congresos

- *Subjective Evaluation of New Room Acoustic Measures* [8]: este estudio realiza experimentos sobre sonoridad, claridad, reverberación, etc. Utiliza un sistema 2-AFC. Participaron 10 personas expertas con entrenamiento previo. La duración de las señales era de 15 segundos. Los usuarios podían manejar las señales y reproducirlas a voluntad. El análisis de los datos se hizo utilizando medias y regresiones convencionales.
- *Simulated Transfer Path Accuracy vs. Sound Perception* [9]: este estudio analiza la similaridad entre señales acústicas. Utiliza un sistema de escala del 1 a 10. Participan 14 personas. No hay información sobre si son expertas ni si ha habido entrenamiento previo. El oyente sólo puede controlar si decide repetir la escucha una vez por experimento. Las señales tienen una duración de 2 segundos. No hay información sobre la duración de la sesión. Para el análisis de datos se ha utilizado el sistema ANOVA.
- *Just Noticeable Differences of Spatial Cues in Echoic and Anechoic Acoustical Environments* [10]: este estudio analiza diferencias entre el nivel y duración de la señal entre sonidos anecoicos y reverberantes. Utiliza un sistema de test 2-AFC. Participan 8 personas expertas y no expertas sin información de entrenamiento previo. Los oyentes no pueden controlar la reproducción de los estímulos, pero pueden decidir si repetirlos. Las señales tienen una duración de 2 segundos. No hay información sobre la duración de la sesión. El análisis de datos se realiza mediante medias y regresiones convencionales.
- *The Sensitivity of Listeners to Early Sound Field Changes in Auditoria* [21]: este estudio realiza experimentos sobre la sensibilidad en campo cercano. Utiliza el sistema en que se anota el valor mínimo en el que detecta diferencias con respecto a la señal original. Participaron 10 personas (de las que se usan 7 al final) todas ellas expertas. No hay información de entrenamiento previo. Las señales duraban entre 5 y 8 segundos. Los sujetos no tenían control sobre la reproducción de las señales. El análisis de datos se hizo mediante F-Test.
- *Some Factors Affecting Multi-Channel Listening* [22]: este estudio analiza el tema de la inteligibilidad acústica. Utiliza un sistema de test Sí/No. Participan un total de 96 personas (48 por parte del experimento). No hay información sobre su experiencia ni de entrenamiento previo. Los oyentes no tienen control sobre los audios y no hay repeticiones. Las señales tienen una duración de 15 segundos. La sesión dura aproximadamente 5 minutos. El análisis de los datos se realiza mediante medias y regresiones lineales.
- *The Nature and Distribution of Errors in Sound Localization by Human Listeners* [23]: este estudio analiza experimentos de localización acústica. El test se realiza marcando la posición referenciada desde el altavoz del que se cree que procede el sonido. Participan 19 personas sin información sobre si son expertas o no. Sí reciben entrenamiento previo. La señal se repite de forma indefinida hasta que el oyente da una respuesta, pero no tiene control sobre el orden. Las señales tienen una duración de 150ms. No hay información sobre la duración de la sesión. Para el análisis de datos se utilizan métodos estadísticos esféricos.

- *Subjective and Objective Assessments of Noise Barriers in Terms of the Loudness Level* [24]: este estudio analiza la sonoridad de señales acústicas. Se utiliza un test de escala numérica positiva. Participan 10 personas no expertas. No hay información sobre entrenamiento previo. El usuario puede escoger repetir la señal hasta 10 veces. Tiene una duración de 3 segundos. La sesión tiene una duración de unos 15 minutos (14 tests). El análisis de los datos se realiza mediante medias convencionales.
- *Subjective and Objective Quality Assessment of Audio Source Separation* [25]: este estudio realiza varios experimentos acústicos relacionados con la sensibilidad. Utiliza un sistema de test *MUSHRA*. Participan 23 personas expertas con entrenamiento previo. Los oyentes pueden seleccionar los audios las veces que quieran, tal y cómo se recomienda para el sistema *MUSHRA*. La duración de las señales es de 5 segundos. No hay información sobre la duración de la sesión. El análisis de los datos se realiza mediante el sistema ANOVA.
- *Perceptive and Objective Evaluation of Calibrated Room Acoustic Simulation Auralizations* [26]: este estudio analiza situaciones de tiempo de reverberación, claridad, tonalidad, etc. Utiliza un sistema de escala continua comparativo entre dos señales A y B. Participan 26 personas expertas con entrenamiento previo. Las señales se repiten automáticamente, pero los oyentes pueden escoger cuando empezar uno de los test. La duración de las señales es de 13 segundos. La duración de la sesión es de 10 minutos aproximadamente habiendo pasado por 20 situaciones distintas. El análisis de los datos se realiza mediante el sistema ANOVA.
- *A Spatial Audio Quality Inventory (SAQI)* [27]: este estudio es el único que no se centra en un caso concreto, sino que su objetivo es definir los distintos aspectos psicoacústicos que pueden estudiarse, de los más generales a los más concretos, y establecer una serie de términos opuestos sobre los que pueden evaluarse. Sobre ellos se obtienen varias definiciones sobre los elementos psicoacústicos más representativos.
- *Audibility of Spectral Dips and Peaks in Broadband Noise* [28]: este estudio analiza diferencias de niveles entre señales. Utiliza un sistema de test 2-AFC. Participan 12 personas no expertas sin entrenamiento previo. Los oyentes no tienen control sobre las señales. Dichas señales tienen una duración de 1 segundo. La sesión tiene una duración de 85 minutos con descansos entre cada test. El análisis de datos se realiza mediante el sistema ANOVA.
- *A Just Noticeable Difference in C50 for Speech* [29]: este estudio realiza experimentos sobre la inteligibilidad y la claridad. El test se realiza mediante un sistema de igual/diferente. Participan 10 personas no expertas con entrenamiento previo. El usuario no tiene control sobre la reproducción de las señales. Las señales tienen una duración indefinida; se encuentran en bucle hasta que el usuario da su respuesta. La sesión duraba 21 experimentos distintos; no hay información sobre la duración media de los mismos. El análisis de los datos se hizo mediante medias y regresiones convencionales.
- *Direct-to-Reverberant Energy Ratio Sensitivity* [30]: este estudio analiza las variaciones de la reverberación. El test se realiza mediante un sistema 2-AFC. Participan 6 personas expertas. No hay información sobre entrenamiento previo. Las señales duran 300 milisegundos. El oyente no tiene control sobre los audios. La sesión dura 2 horas, sin información sobre descansos. El análisis de los datos se hace mediante estadística paramétrica siguiendo un modelo Chi cuadrado.
- *About Just Noticeable Differences for Aspects of Spatial Impressions in Concert Halls* [31]: El estudio se centra en el análisis de diferencias de niveles sonoros. El tipo de test utilizado es

un igual/diferente. Participan 12 personas (de las cuales se analizan los de 7) no expertas con entrenamiento previo. Los oyentes no tienen control sobre los estímulos que tienen una duración de 15 segundos. Las sesiones duran 30 minutos más el tiempo de descanso no especificado. El análisis de los datos se realiza mediante estadística paramétrica bajo el modelo de Bernuilly.

- *The Just Noticeable Difference of Center Time and Clarity Index in Large Reverberant Spaces* [32]: Este estudio analiza la inteligibilidad en espacios reverberantes. Utiliza el sistema en que se anota el valor mínimo en el que detecta diferencias con respecto a la señal original. Participan inicialmente 40 personas de las que se usan los datos de 13. Son expertos y no expertos con entrenamiento previo. Los participantes controlan la reproducción de las señales y pueden repetirse cuantas veces se quiera. Las señales tienen una duración de entre 5 y 9 segundos. No hay información sobre la duración de la sesión. El tipo de análisis de datos es un T-Test.
- *Effect of Experimental Design on the Results of Clarity-Index Just Noticeable-Difference Listening Tests* [33]: en este estudio se realizan experimentos sobre la iteligibilidad. Utiliza un sistema de test de igual/diferente. Participan entre 11 y 51 personas en función de la parte del experimento. Todos ellos son expertos y realizan un entrenamiento previo. Los oyentes no controlan la reproducción de las señales, pero puede escoger si repetir las señales que deben escucharse completamente. La duración de las señales oscila entre 10 y 11 segundos. No hay información de la duración de la sesión. El análisis de los datos se realiza mediante medias y regresiones convencionales aplicando estadística paramétrica.
- *Research on the Bandwidth Effect on Binaural Loudness Summation on the Cam Scale* [34]: este estudio realiza un experimento sobre la sonoridad. Aplica un sistema de escala numérica marcando la ganancia aplicada. Participan 5 personas no expertas con entrenamiento previo. Los oyentes no controlan los estímulos, pero se repiten de manera indefinida hasta alcanzar tres veces el mismo nivel de sonoridad. Las señales duran 5 segundos. No hay información de la duración de la sesión, pero se realizan 6 experimentos. El análisis de datos se realiza con medias y regresiones convencionales.
- *Annoyance Penalty of Amplitude-Modulated Sound* [35]: este estudio se centra en el análisis de la molestia auditiva. Utiliza un sistema de escala numérica del 1 al 10. Participan 32 personas no expertas, pero con entrenamiento previo. El oyente no tiene control sobre los audios. La duración de la señal es continua, pero debe escucharse como mínimo 18.5 segundos. La duración de la sesión es de aproximadamente 65 minutos repartidos en 74 casos con descanso. El análisis de datos se realiza mediante medias y regresiones convencionales.
- *Acuity of Spatial Stream Segregation Along the Horizontal Azimuth with Non-Individualized Head-Related Transfer Functions* [36]: este estudio se centra en el análisis de localización de señales acústicas. Se realiza mediante un sistema de acertar/fallar. Participan 18 personas no expertas, pero con entrenamiento previo. Los oyentes no tienen control sobre las señales más allá de las repeticiones que se produzcan por el acierto o el fallo. Las señales duran 20 milisegundos que se repiten hasta un máximo de 6.4 segundos. La sesión dura entre 60 y 90 minutos con descansos entre cada test. Para el análisis de datos se utiliza el sistema Wilcoxon signed-rank.
- *Improved Binaural Speech Inteligibility by Adding Reverberation to the Target Speaker* [37]: este estudio se centra en el análisis de la inteligibilidad acústica. Se realiza lo que se conoce como un *Speech Recognition Test* al 50 % o SRT50 registrando los valores de ganancia necesarios para llegar a dicho punto. Participaron 8 personas no expertas, pero con entrenamiento previo. Los oyentes no tienen control sobre los estímulos. Las señales duran 3.5 segundos repartidos entre

la señal de interferencia y 2.5 segundos de la señal a reconocer. La sesión dura unos 90 minutos con descansos. El análisis de los datos se realiza mediante estadística paramétrica.

- *Feature Analysis of Sound Direction Perception Using Frequency Band-Limited Stimuly: Extension of a Directional Band Model* [38]: este estudio se centra en el análisis de la localización sonora. El test se realiza mediante selección del altavoz de procedencia del sonido. Participan 5 personas. No hay información sobre si son expertas o de si han realizado entrenamiento previo. El usuario no tienen control sobre los estímulos. Las señales tienen una duración de 1 segundo. La sesión dura aproximadamente 3 minutos. Para el análisis de datos se presupone la utilización de medias convencionales.
- *Annoyance Modeling of Construction Noise Using Acoustical Features, Noise Sensitivity and Health Condition* [39]: este estudio se centra en el análisis de la molestia auditiva. Se utiliza un sistema de escala numérica de 11 valores (de 0 a 10). Participan 48 y 60 personas en función de la parte del estudio. No son expertas y no hay información sobre entrenamiento previo. No existe control de los estímulos por parte del oyente. Las señales tienen una duración de 10 segundos. La duración de la señal es de 7 minutos para la primera parte y 2 minutos para la segunda.
- *Spatial Sound Segregation in the Monaural Listening Condition* [40]: este estudio centra su análisis en la localización acústica. Utiliza un sistema de dibujo de la trayectoria del sonido respecto del oyente. Participan 13 personas expertas sin entrenamiento previo. El oyente tiene control del número de repeticiones, pero no del inicio de la reproducción. Las señales tienen una duración de 5 segundos. No hay información de la duración de la sesión. El análisis de datos se realiza mediante estadística paramétrica utilizando medias convencionales y percentiles.
- *Loudness of Ramped and Damped Sounds that are Temporally Shifted Across Ears* [41]: este estudio se centra en el análisis de la sonoridad. Se utiliza un sistema Más alto/Más bajo que la señal de referencia. Participan 16 personas sin entrenamiento previo. No hay información sobre si son expertas. Los oyentes no tienen control sobre la reproducción de las señales. Las señales tienen una duración de aproximadamente 1.5 segundos. No hay información sobre la duración de la sesión, pero se sabe que son 8 tests por persona. El análisis de datos se realiza mediante estadística paramétrica utilizando medias convencionales y percentiles.
- *Annoyance of Impulsive Sounds - A Psychoacoustic Experiment Involving Synthetic Sounds* [42]: este estudio se centra en el análisis de la molestia auditiva. Se utiliza una escala numérica de 11 puntos (de 0 a 10). Participan 40 personas no expertas, pero con entrenamiento previo. No existe control por parte del usuario sobre los estímulos. Las señales tienen una duración de 20 segundos. La sesión tiene una duración de unos 30 minutos. El análisis de los datos se realiza mediante el sistema ANOVA.

6.3. Libros de referencia

- *Hearing: An Introduction to Psychological and Psysiological Acoustics* [4]: este libro se utiliza como base a la hora de definir diferentes elementos psicoacústicos como la fatiga auditiva, la sensibilidad, inteligibilidad, etc.
- *Perceptual Audio Evaluation-Theory, Method and Application* [5]: este libro se utiliza como referencia para varios elementos a considerar en el protocolo. Dentro de él se encuentran recomendaciones sobre la limitación de la duración de las sesiones, que según el mismo deben estar

comprendidas entre los 30 y 40 minutos para evitar la fatiga auditiva de los participantes. También se recomienda el uso de los sistemas ANOVA para el análisis de los datos obtenidos. Por otro lado, da información sobre los diferentes tipos de test que pueden realizarse y que son una base dentro del capítulo de “Definiciones” como son las test binomiales y las escalas numéricas. También hace referencia en varias ocasiones a las normativas y referencias que se aplican a nuestro proyecto.

- *Introduction to Modern Nonparametric Statistics* [6]: este libro se utiliza como referencia para las definiciones de varios de los diferentes tipos de análisis estadísticos que pueden aplicarse para el estudio de los datos obtenidos en los experimentos psicoacústicos, especialmente en el caso de la estadística no paramétrica. Debido a su carácter introductorio, sus contenidos quedan reflejados dentro del capítulo de “Definiciones”.
- *Psychoacoustics: Facts and Models* [7]: este documento aporta información sobre definiciones de diferentes aspectos psicoacústicos como la sonoridad y sobre los diferentes métodos para test psicoacústicos como los test Sí/NO, los métodos 2AFC, entre otros.
- *Spatial Hearing: The Psychophysics of Human Sound Localization* [11]: este libro se utiliza como referencia para la determinación de los límites de tiempo para las señales psicoacústicas y evitar que el oyente se acostumbre a ellas. Según el mismo, no debe superar los 3 minutos y requiere de una pausa de 1-2 minutos para volver a obtenerse resultado adecuados.

Capítulo 7

Clasificación de los estudios

El siguiente paso es clasificar los estudios en diferentes grupos. En primer lugar, se dividen dentro de las categorías ya explicadas en el capítulo de “Definiciones”. En esta clasificación no se han incluidos los libros debido a que su área de estudio es muy general y puede ser aplicado en todos los casos; además que al aportar cada uno elementos sobre aspectos muy específicos, su información no quedaría bien reflejada en las tablas.

7.1. Inteligibilidad

En este grupo se encuentran los estudios relacionados con la inteligibilidad y el entendimiento del habla. Dichos estudios y sus características son los que se presentan en la tabla 7.1.

Como se puede ver, los estudios se corresponden con [22], [29], [32] y [33]. En este caso, no se cree conveniente aplicar ninguna subdivisión adicional, puesto que todas siguen características muy similares entre sí.

7.2. Sensibilidad

Esta categoría engloba a los estudios relacionados con el estudio de la sensibilidad sonora y la sonoridad. Los estudios están presentados en la tabla 7.2.

Los experimentos se corresponde con [9], [10], [21], [24], [25], [28], [31] y [41]. Dentro de esta categoría pueden diferenciarse dos subgrupos: uno para las diferencias de nivel, y otro para el estudio de la sonoridad.

7.3. Reverberación

En esta categoría se engloban los estudios relacionados del estudio de la sensación de reverberación. Dichos estudios están representado en la tabla 7.3.

Como se puede apreciar, los experimentos se corresponden con [8], [10], [26] y [30]. En esta categoría no es posible aplicar ninguna otra subdivisión.

7.4. Localización

En esta categoría se encuentran los estudios relacionados con la localización acústica. Dichos estudios están presentados en la tabla 7.4.

Referencia	1999, JS. Bradley, R. Reich and SG. Norcross	2010, F. Martellotta	2010, MC. Vi-geant and RD. Celmer	1954, J. Egan, E. Carterette and E. Thwing (a)	1955, J. Egan, E. Carterette and E. Thwing (b)	1956, J. Egan, E. Carterette and E. Thwing (c)	1957, J. Egan, E. Carterette and E. Thwing (d)
Medida	Inteligibilidad c50	Inteligibilidad en espacios reverberantes	Inteligibilidad	Inteligibilidad	Inteligibilidad	Inteligibilidad	Inteligibilidad
Tipo test	Same-different	Minimal Changes (Es diferente y cuanto)	Same-different	Ganancia aplicada	Recorte de frecuencia	Ganancia aplicada	Identificación de mensaje (Sí/No)
Nº Participantes	10	40 (13 pasada la prueba)	11-51	18	8	16	96 (48 por parte)
Tipo de oyente	No	Expertos y no expertos	Expertos	nsnc	nsnc	nsnc	nsnc
Duración señal	loop	5s y 9s	10.9s 10.3s	15s (5 aviso y 10 de señal)	15s (5 aviso y 10 de señal)	15s (5 aviso y 10 de señal)	15s (5 aviso y 10 de señal)
Duración sesión	21 casos con descansos entre ellos	nsnc	nsnc	20*15=300s (5min)	20*15=300s (5min)	20*15=300s (5min)	20*15=300s (5min)
Entrenamiento previo	Sí	Sí	Sí	ns/nc	Si	ns/nc	ns/nc
Análisis de datos	Diferencia y regresión	Función psicométrica	Diff Limen 50 % and regression	Medias convencionales y regresiones no lineales	Medias convencionales y regresiones no lineales	Medias convencionales y regresiones no lineales	Medias convencionales y regresiones no lineales
Tipo de estímulo	Male rhyme	Pange Lingua y Theme for Cello	Diferentes grabaciones anecoicas	Dos frases dichas por hombres a la vez	Dos frases dichas por hombres a la vez	Dos frases dichas por hombres a la vez	Dos frases dichas por hombres a la vez
Tipo de experimento (si se muestra una sola vez, lo modifica el usuario, etc.)	No hay repeticiones. Deciden y se pasa al siguiente	Los oyentes controlan el software y pueden repetir cuanto quieran.	En algunos puntos el oyente puede cambiar las señales cuando quieran. En otros, tienen que escuchar las señales enteras. Luego pueden repetir, pero siempre siguiendo el método	El oyente no puede conmutar. Los audios se presentan una vez. Sesiones conjuntas.	El oyente no puede conmutar. Los audios se presentan una vez. Sesiones conjuntas.	El oyente no puede conmutar. Los audios se presentan una vez. Sesiones individuales.	El oyente no puede conmutar. Los audios se presentan una vez. Sesiones conjuntas.

Tabla 7.1: Características de los estudios de inteligibilidad acústica.

Referencia	1993, T. Cox, W. Davies and Y. Lam	2005, IB. Witew, GK. Behler and M. Vorländer	2016, S. Klockgether and S. van de Par	2019, L. Kritly, V. Chmelik et al.	2019, G. Pulvirenti, N. Totaro and E. Parizet	2019, M. Nowak and P. Kokowski	2019, J. Schlittenlacher, R. Zhao and B. Moore	2011, V. E. Emiya, Vincent et al.
Medida	Sensibilidad campo cercano	Diferencias de nivel	Diferencia nivel y tiempo entre anecoico y reverb	Detección de diferencias de nivel	Similaridad	Loudness	Loudness	Varios
Tipo test	Minimal Changes	Same-different	Adaptive 31/2-AFC	AFC (Señal A y B distintas y X igual a alguna de las 2)	Escala 1-10	Absolute Method of Estimation (escala positiva)	Más alto/más bajo que la referencia	MUSHRA (Escala continuas en paralelo)
Nº Participantes	7-10	7-12	8	12	14	10	16	23
Tipo de oyente	Expertos	No	Expertos y no expertos	No	nsnc	No	nsnc	Expertos
Duración señal	8 y 5 s	15s	2s, 2s y 2s	1s	2s	3s	1540ms	5s
Duración sesión	nsnc	30 mins (+ descansos)	nsnc	1h 25min con descansos entre cada test	nsnc	15 min	8 pruebas	nsnc
Entrenamiento previo	ns/nc	Sí	ns/nc	No	nsnc	ns/nc	No	Sí
Análisis de datos	F-test analisis	Función psicométrica y Monte Carlo fit.	Medias y cálculo de errores estándar	ANOVA	ANOVA	Medias geométricas	Medias y percentiles aritméticos	Intervalos de confianza, Regresiones lineales, ANOVA
Tipo de estímulo	Sonido de agua y The Scottish	Trumpet concerto Comet	Guitar, violin y snare drum	Ruido rosa con pico o valle a una determinada frecuencia	Señal simulada y señal grabada	Grabaciones de ruido en laboratorio anecoico con y sin pantalla acústica	Tren de pulsos de tonos de 1khz de diferente longitud	Voces, notas de instrumentos. 80 sonidos diferentes
Tipo de experimento (si se muestra una sola vez, lo modifica el usuario, etc.)	El usuario es pasivo, no interviene en la reproducción de los audios.	Los oyentes no pueden repetir los estímulos.	Los estímulos no son controlados por los oyentes. Se repiten los sucesos varias veces ajustando los valores.	Para cada prueba se muestran 3 señales. Una A y B distintas entre sí, y luego una X que es igual de forma aleatoria a una de las otras 2. Usuario dice a cuál es igual X	Se ponen las dos señales y se hace una única repetición.	Cada señal se repite 10 veces (las 14 señales).	Se le pone una señal de prueba y el oyente juzga si es más alto o más bajo. Según eso se añade o se resta 5dB hasta que se produce la inversión	Los oyentes pueden seleccionar los audios cuanto deseen tal y cómo especifica MUSHRA

Tabla 7.2: Características de los estudios de sensibilidad acústica.

Referencia	1995, GA. Soulodre and JS. Bradley	2002, P. Zahorik	2016, S. Klockgether and S. van de Par	2016, B. Postma and B. Katz
Medida	loudness, clarity, reverberance, bass, treble, envelopment, apparent source width, and overall preference	Variación directo/reverberante para distancia	Diferencia nivel y tiempo entre anecoico y reverb	Tiempo de reverb, coloración, tonalidad, claridad, etc.
Tipo test	Attribute 2-AFC	Adaptive 2-AFC	Adaptive 31/2-AFC	Escala continua A es más o B es más
Nº Participantes	10	6	8	26
Tipo de oyente	Expertos	Expertos	Expertos y no expertos	Expertos
Duración señal	15s	50 ms 300ms 300ms impulse	2s, 2s y 2s	13s
Duración sesión	nsnc	2h	nsnc	10 mins aprox. (20 sesiones*23s + repeticiones)
Entrenamiento previo	Sí	ns/nc	ns/nc	Sí
Análisis de datos	Regresión	Función psychometrica y Chi-square fit test	Medias y cálculo de errores estándar	Estadística paramétrica
Tipo de estímulo	Nozze Figaro	Burts and impulse	Guitar, violin y snare drum	Señal de voz auralizada y señal grabada
Tipo de experimento (si se muestra una sola vez, lo modifca el usuario, etc.)	El usuario puede escoger qué pone dentro de las 10 señales y repetirlas cuanto quiera.	Se hacen preguntas y responde. El oyente no tiene control sobre los audios.	Los estímulos no son controlados por los oyentes. Se repiten los sucesos varias veces ajustando los valores.	Repeticiones indefinidas y selección de inicio.

Tabla 7.3: Características de los estudios de reverberación acústica.

Los experimentos se corresponden con [23], [36], [38] y [40]. En esta categoría no es necesario aplicar más subdivisiones.

7.5. Molestia

Esta categoría engloba los estudios relacionados con los experimentos de molestia auditiva. Estos estudios se presentan en la tabla 7.5.

Dichos estudios se corresponden con [35], [39] y [42]. Al igual que con los casos anteriores no se considera necesario aplicar más subdivisiones.

7.6. Recomendaciones y normativas

Esta categoría engloba las recomendaciones y normativas sobre medidas subjetivas de audio. Están presentadas en la tabla 7.6.

Se ha considerado esta categoría debido al carácter generalista de las normas y que, al final, consisten en la base del protocolo para todas las categorías.

Referencia	1997, S. Carlile, P. Leong and S. Hyams	2019, M. Shiell and E. Formisano	2019, M. Yamada, F. Saze et al	2019, D. Morikawa, D. Kojima and T.Hirahara
Medida	Localización	Localización	Localización	Localización
Tipo test	Precision con sonido parado	Acertar/Fallar	Escoger el altavoz por el que ha salido la señal. (Numerados del 1 a 8). Si no pueden localizarlo se da un 0.	Dibujo de la trayectoria sobre tablet
Nº Participantes	19	18	5	13
Tipo de oyente	nsnc	No	nsnc	Sí
Duración señal	150ms	20ms	1s	5s
Duración sesión	nsnc	60-90 min. Con descansos entre cada test.	20 sesiones (cada sesión son 1s de señal y 2 de espera repetidos 3 veces)	nsnc
Entrenamiento previo	Sí	Si	ns/nc	No
Análisis de datos	Spherical statistics methods	Wilcoxon signed-rank y medias y medianas aritméticas. Correlación Pearson.	No info (medias aritméticas supuestas)	Medias y percentiles aritméticos
Tipo de estímulo	Ruido blanco generado por software	Señales de ruido creadas con Matlab	Ruidode 1/3 de octava	Ruido gaussiano
Tipo de experimento (si se muestra una sola vez, lo modifica el usuario, etc.)	El sonido se repite de forma indefinida hasta que el oyente ha fijado la localización de la que cree que proviene el sonido	La señal sigue un patrón, a la vez que el enmascarante se repite. La duración máxima es de 6.4segs. Si se acierta, se reduce la separación en 1 paso, si se falla, se aumenta la separación en 3 pasos. Se hacen 15 barridos por experimento.	Los estímulos se repiten 3 veces de forma aleatoria en uno de los altavoces	La señal se repite el número de veces que el oyente quiera. No lo activa él/ella. Primero lo hace con una oreja tapada y luego la otra.

Tabla 7.4: Características de los estudios de localización acústica.

Referencia	2019, V. Rajala and V. Hongisto	2019, J. Lee, S. Chang et al.	2019, V. Hongisto and P. Virjonen
Medida	Molestia	Molestia	Molestia
Tipo test	Escala de 0 a 10	Escala de 11 puntos	0 a 10
Nº Participantes	32	48 y 60	40
Tipo de oyente	No	No	No
Duración señal	18.5s	10s	20s
Duración sesión	65 min (74 casos)	40 y 12 experimentos. 400s y 120s aprox.	92 sonidos
Entrenamiento previo	Sí	ns/nc	Sí
Análisis de datos	Medias aritméticas, regresiones lineales, y conversión a escala A	Medias, percentiles y regresiones lineales.	ANOVA
Tipo de estímulo	Señal simulada a través de Matlab	Grabaciones de ruido de construcción	Señales generadas con pseudo ruido a través de MATLAB
Tipo de experimento (si se muestra una sola vez, lo modifica el usuario, etc.)	Se pone la señal de forma indefinida hasta que el usuario marca la casilla (tiene que esperar los 18.5s)	Los estímulos sólo se reproducen una vez. No hay control por parte del usuario	Los estímulos se presentan una vez. No hay control por parte del usuario.

Tabla 7.5: Características de los estudios de molestia auditiva.

Referencia	2015, UIT-R BS.1116-3	2019, UIT-R BS.1284-2	2015, UIT-R BS.1534-3
Medida	Pequeñas Degradaciones	Generales	Ev calidad interm
Tipo test	Escala 1-5	Escala 1-5 Escala 1-7	MUSHRA (Escala
Nº Participantes	20	Escala continua 0-100 10 expertos o 20 no expertos	continuas en paralelo) 20
Tipo de oyente	Expertos	Preferiblemente expertos	Expertos
Duración señal	10-25s	15-20s	10-12s
Duración sesión	20-30 min no más de 10-15 experimentos	15-20 min (si hay que hacer más, poner pausas de la misma duración)	ref a UIT 1116
Entrenamiento previo	Sí	Sí	Sí
Análisis de datos	ANOVA	Medias e intervalos de confianza	Análisis de varianza (ANOVA), medias, etc.
Tipo de estímulo	nsnc	nsnc	nsnc
Tipo de experimento (si se muestra una sola vez, lo modifica el usuario, etc.)	Estudios individuales, permitir conmutar de forma individual	El oyente debe conmutar los estímulos	El oyente puede pasar de la señal de referencia a cualquier otra.

Tabla 7.6: Características de las normativas y recomendaciones.

Capítulo 8

Diseño del protocolo

Una vez establecidas las definiciones mínimas y realizado el análisis de los documentos enfocados en nuestro estudio, puede comenzarse el diseño del protocolo. Este diseño está dividido en tres partes:

- En primer lugar, se establecen unas recomendaciones generales que son independientes del aspecto subjetivo de audio que se quiera realizar.
- Después, se identifican los rasgos y elementos del ensayo sobre los que hay que tomar una serie de decisiones.
- A continuación se presentan una serie de recomendaciones generales en función de la categoría del ensayo.
- Por último, se establece el criterio que se recomienda seguir para la toma de dichas decisiones.

8.1. Recomendaciones generales

Las recomendaciones que se presentan en este apartado se aplican a cualquier tipo de test subjetivo de audio. Para ello, se utilizan las consideraciones generales de las normativas y los límites que fijan los libros de referencia consultados. Del mismo modo, conforman parte de este apartado aquellos elementos comunes analizados en los diferentes artículos y conferencias y que son similares para las diferentes categorías.

- Es recomendable que los estudios se hagan de forma individual, en vez de en grupo.
- También se recomienda que los participantes de los estudios sean expertos o con experiencia en la realización de estudios subjetivos de audio. En caso de no disponer de personas expertas, pueden utilizarse personas sin experiencia.
- Independientemente del grado de experiencia de los participantes, se recomienda que se realice un entrenamiento previo para facilitar el manejo de la interfaz del experimento y saber cómo dar las respuestas adecuadamente.
- En caso de que durante el entrenamiento una persona no aporte un nivel de audición acorde con el objeto del experimento, debe ser sustituido por otro oyente.
- El mínimo de personas que deben realizar el experimento debe rondar la decena. En caso de utilizar personal no experto este número debería aumentarse.

- La duración de las señales, aún dependiendo del tipo de experimento, no debe exceder de 2 minutos. Esto es independiente de que sea una señal única o en bucle. En general, con unos pocos segundos puede ser suficiente. De esta forma se evita que el oyente se acostumbre a la señal y se obtengan resultados no válidos.
- Por lo general, es recomendable que el oyente tenga la capacidad de conmutar entre diferentes estímulos y las repeticiones, salvo que el tipo de estudio lo impida.
- Si se va a repetir la reproducción de una misma señal de forma consecutiva, es recomendable que pasen unos 5-10 segundos para “desacostumbrar” al oyente al sonido.
- La duración de la sesión no debe exceder los 30 minutos para evitar los efectos de la fatiga auditiva. En caso de ser necesario alargar la duración, es recomendable establecer descansos de unos 15-20 minutos.

Estas recomendaciones determinan los mínimos para la mayoría de test psicoacústicos; no obstante, en algunos casos se requiere utilizar otras recomendaciones más específicas o alguna modificación de las anteriores. Para ello, conviene determinar una identificación de los rasgos propios del ensayo que se va a realizar.

8.2. Identificación de los rasgos del ensayo

A la hora de preparar un estudio subjetivo de audio, conviene invertir un tiempo en responder una serie de cuestiones previas. Estas preguntas facilitarán establecer las particularidades del ensayo y poder establecer así unas características que deben seguirse para obtener los mejores resultados posibles.

¿En qué categoría se engloba mi estudio?

Para responder a esta pregunta hay que tener determinado qué elemento o elementos psicoacústicos se quieren analizar en el estudio. Esta pregunta puede parecer sencilla, pero de ello depende que se apliquen unas recomendaciones u otras. En la figura 4.1 ya se mostró un diagrama con la relación entre los elementos psicoacústicos y las distintas categorías que se habían definido para este tipo de ensayos.

Es posible que en un mismo estudio se quieran analizar diferentes aspectos de la psicoacústica. Puede ocurrir que coincidan dentro de la misma categoría. Si esto es así, puede aplicarse el mismo procedimiento para todos ellos, e incluso realizarse de forma paralela. Si se diera la situación de que los elementos forman parte de categorías distintas, se recomienda que se realicen por separado.

¿Qué tipo de test es el más adecuado?

Según el aspecto psicoacústico que se pretende estudiar, puede ser recomendable utilizar un tipo de test u otro. Por lo general, las recomendaciones de la UIT [1], [2] y [19] señalan las escalas numéricas como las más apropiadas. Se pueden utilizar escalas positivas del 0 al 10 si se quiere evaluar cuánto de intenso es un determinado aspecto. También pueden utilizarse otras escalas enteras para que el oyente compare entre dos elementos. Si se van a estudiar varios elementos, la recomendación es la utilización de test *MUSHRA*.

Por otro lado, pueden existir ciertos estudios que, por su estructura o su objetivo, pueden requerir otros tipos de test como señalizaciones de procedencia, selección de señales (2-AFC), entre otros. Es labor de las personas que preparan el ensayo decidir cuál de todas las opciones que se han presentado es la más adecuada.

¿Participantes expertos o inexpertos?, ¿cuántos necesito?

Como ya se comentó en el apartado de “Recomendaciones generales”, a priori es recomendable contar con participantes expertos en la realización de test subjetivos de audio. No obstante, es posible que en algunas situaciones quieran utilizarse muestras que incluyan a personas que no son expertas en este tipo de procedimientos. Independientemente del tipo de personas que se utilicen, es recomendable que realicen un entrenamiento previo. De esta forma se convierte a la gente sin experiencia en personas con experiencia y es más probable que los resultados que se obtengan sean relevantes.

En cuanto al número de personas, las recomendaciones de la UIT [1], [2] y [19] establecen unos mínimos de 10 personas con experiencia o 20 en el caso de que no la tengan. No obstante, para algunas categorías puede ser interesante ampliar este número para obtener mejores resultados.

¿Qué tipo de señal utilizo?

El tipo de señal que se vaya a utilizar es algo más relevante de lo que puede parecer en un principio. A no ser que el objeto del estudio esté relacionado con un tipo de sonidos concretos, algunas señales son más útiles que otras según el tipo de ensayo. Por ejemplo: a la hora de reconocer la localización de un sonido, una señal impulsiva puede ser más difícil de reconocer que una señal más constante en el tiempo como un ruido blanco. Por otro lado, si se quiere estudiar la degradación de calidad subjetiva que sufre una señal entre dos formatos de archivos de audio, un ruido blanco puede no ser la elección más apropiada.

¿Cuál es el tipo de análisis de datos?

El tipo de análisis de datos varía según lo que se pretenda estudiar. Utilizar un sistema u otro conlleva una serie de presunciones que pueden tener efectos en los resultados obtenidos. Las recomendaciones de la UIT [1], [2] y [19] establecen los sistemas de análisis de la varianza (ANOVA) como el sistema por defecto. Este sistema presupone que los datos siguen una distribución normal. En algunos casos, puede presuponerse que los datos no siguen ninguna distribución en concreto y habría que utilizar en su lugar sistemas de estadística no paramétrica. Por otro lado, pueden darse estudios en los que la distribución sea sencilla y utilizar los sistemas de medias, percentiles e intervalos de confianza habituales son suficientes. En caso de haber realizado un test binomial, la recomendación sería aplicar las tablas de la norma ISO10399 [20].

Es por esto por lo que es imprescindible identificar las características de nuestros datos para poder analizarse de forma correcta.

8.3. Recomendaciones específicas

Para ayudar a resolver las preguntas anteriores, se presentan una serie de recomendaciones más específicas según el tipo de categoría en el que se engloba el estudio. Al igual que con las recomendaciones generales, pueden ser modificadas si el experimento así lo requiere.

Inteligibilidad

En esta categoría se recomienda la utilización de tests de escalas numéricas si se quieren realizar estudios cuantitativos. En caso de que quieran estudiarse varios elementos a la vez, se recomienda el uso del test *MUSHRA*. Si se quiere estudiar entre la selección entre señales, la recomendación es la de utilizar test 2-AFC.

El número de participantes mínimo que se recomienda es de 20 personas, preferiblemente expertos, y que realicen un entrenamiento previo. También se recomienda que los oyentes puedan decidir la repetición de las señales o conmutar entre las mismas.

La duración de la señal se recomienda que esté en torno a los 10-20 segundos. La duración de la sesión en su totalidad no debería exceder los 30 minutos. Pequeños descansos entre análisis de señales también se consideran necesarios.

Se recomienda un análisis de datos usando los sistemas ANOVA; no obstante, si no se puede suponer una distribución normal de los datos, por ejemplo si se tienen pocos, puede recurrirse a sistemas de estadística no paramétrica o de comparación de medias. En cualquier caso, es conveniente aplicar el estudio de los intervalos de confianza.

Sensibilidad

En esta categoría, al igual que en la anterior, se recomienda la utilización de escalas numéricas para estudios cuantitativos; la utilización de *MUSHRA* para varios elementos o características en paralelo y, si se quiere estudiar selección entre dos señales, el sistema 2-AFC.

Siguiendo las recomendaciones de la UIT, se determinan unos mínimos entre 10 y 20 personas según si son o no expertas con entrenamiento previo. En caso de tratarse de sensibilidad muy exigente, la recomendación es utilizar un mínimo de 20 personas expertas. El entrenamiento previo sigue siendo recomendable en cualquier caso. Se recomienda también que los oyentes puedan decidir si repetir la reproducción de las señales y la posible conmutación entre las mismas.

La duración de las señales se recomienda que tenga una duración entre los 10 y 15 segundos. La sesión no debería superar los 20-30 minutos. También se recomienda que un oyente no realice más de 15 experimentos por sesión. Los descansos también son recomendables.

Se recomienda seguir un análisis de datos utilizando los sistemas ANOVA para esta categoría, pero, al igual que en el caso anterior, si no se puede suponer una distribución normal de los datos, se puede recurrir a medias convencionales o sistemas de estadística no paramétrica junto al uso de intervalos de confianza.

Reverberación

En la categoría de reverberación acústica se continúa con la recomendación de escalas numéricas, test *MUSHRA* o el sistema 2-AFC según la tipología del estudio.

Se recomienda un número mínimo de entre 10 y 20 personas, preferiblemente expertas, realizando un entrenamiento previo. Se recomienda que los participantes puedan conmutar entre posibles señales y decidir si quieren repetir la reproducción de las mismas.

La duración de las señales no debería exceder los 15 segundos. Al igual que otros estudios, la sesión no debería durar más de 20-30 minutos, y la realización de descansos es también recomendable.

En esta categoría se recomienda prestar especial atención al tipo de distribución que siguen los datos. Si se supone una distribución normal, la recomendación es utilizar los sistemas ANOVA. Si se considera que no siguen ninguna distribución, se recomienda utilizar sistemas de estadística no paramétrica utilizando estudios comparativos de medias como los sistemas de tipo *Wilcoxon*, entre otros. Por último, también se pueden utilizar los estudios estadísticos convencionales de cálculo de medias y regresiones. En cualquier caso, es recomendable el cálculo de los intervalos de confianza.

Localización

Para esta categoría se recomienda un tipo de test numérico donde se asigna un valor numérico en función de la procedencia del estímulo según la posición del oyente. En caso de no poder determinar la procedencia se asigna un valor distinto a los demás (por ejemplo 0).

Se recomienda un mínimo de 20 personas expertas o no expertas para la realización del ensayo. La realización de un entrenamiento previo se considera recomendable. Al igual que en otras categorías, se recomienda que el oyente pueda decidir si repetir la reproducción de la señal o la conmutación entre estímulos.

Para estos estudios, se recomienda la utilización de señales de ruido aleatorio de duración no mayor a 5 segundos. La duración de la sesión no debe exceder los 30 minutos y la realización de descansos se considera recomendable.

En estos ensayos, el análisis de datos recomendable depende de la distribución que se presuponga a los mismos. Suelen ser habituales los métodos de estadística esférica o el sistema *Wilcoxon signed-rank* para estadística no paramétrica para análisis de medias. No obstante, tampoco se descarta la utilización de sistemas de estadística paramétrica. Al igual que en los casos anteriores, se recomienda el cálculo de los intervalos de confianza de los valores medios obtenidos.

Molestia

Para los estudios de molestia acústica, se recomiendan test de escalas numéricas convencionales para estudios individuales o el sistema *MUSHRA* para grupos de estímulos.

Se recomienda grupos más numerosos para este tipo de estudios de unas 30-40 personas. La experiencia no se considera tan crítica en esta categoría, pero sí recomendable. El uso de entrenamiento previo sigue siendo recomendado igualmente. La repetición de las señales tampoco es tan relevante en este caso, pero puede ser recomendable en ciertos estudios. La conmutación entre las mismas se sigue recomendado para experimentos comparativos.

La duración de las señales se recomienda que esté comprendida entre 10 y 20 segundos. La duración de la sesión no debería ser superior a los 20 minutos y el uso de descansos también se cree adecuado.

El análisis de datos recomendado para esta categoría depende también de la distribución que se suponga de los datos. Si se supone una distribución normal, entonces la recomendación es utilizar sistemas ANOVA. También pueden utilizarse sistemas F-Test o T-Test. Por otro lado, pueden utilizarse sistemas estadísticos no paramétricos si no se puede suponer ninguna distribución conocida. Como en el resto de los casos, se recomienda el cálculo de los intervalos de confianza.

Como resumen de las situaciones presentadas se presenta la tabla 8.1.

Categoría	Inteligibilidad	Sensibilidad	Reverberación	Localización	Molestia
Tipo test	Escalas numéricas, MUSHRA o 2-AFC	Escalas numéricas, MUSHRA o 2-AFC	Escalas numéricas, MUSHRA o 2-AFC	Sistema numérico de procedencia. Valor único para imposibilidad de localización	Escala numérica o MUSHRA
Nº Participantes	20	20	10-20	20	30-40
Tipo de oyente	Expertos	Expertos	Expertos	Cualquiera	Cualquiera
Duración señal	10-15s	10-15s	15s max	5s max	10-20s
Duración sesión	30min max	20-30min (no más de 15 experimentos)	20-30min	30min max	20min max
Entrenamiento previo	Sí	Sí	Sí	Sí	Sí
Análisis de datos	ANOVA u otros según distribución de datos	ANOVA u otros según distribución de datos	ANOVA u otros según distribución de datos.	ANOVA u otros según distribución de datos	ANOVA u otros según distribución de datos
Tipo de estímulo	-	-	-	Ruido aleatorio	-
Tipo de experimento (si se muestra una sola vez, lo modifica el usuario, etc.)	Los oyentes pueden conmutar entre señales y repetirlas.	Los oyentes pueden conmutar entre señales y repetirlas.	Los oyentes pueden conmutar entre señales y repetirlas.	Los oyentes pueden conmutar entre señales y repetirlas.	Repetición y conmutación según objeto del estudio.

Tabla 8.1: Características de los estudios de sensibilidad acústica.

8.4. Criterio para la toma de decisiones

A la hora de tomar las decisiones sobre un determinado experimento, es necesario seguir un determinado criterio. En el caso de los experimentos subjetivos de audio se ha considerado, a modo de recomendación, el siguiente:

- En primer lugar hay que definir los rasgos del ensayo. Para ello, puede ser de ayuda responder a las preguntas planteadas en el apartado “Identificación de los rasgos del ensayo”.
- El siguiente paso es conocer cuáles son las recomendaciones de los organismos internacionales. Estas instituciones suelen determinar una serie de mínimos que deben seguirse. La mayoría de esas recomendaciones se encuentran englobadas dentro del apartado de “Recomendaciones generales”.
- El último paso es la determinación de los aspectos concretos del estudio. Para ello se recomienda seguir las indicadas en “Recomendaciones específicas” o consultar estudios sobre una temática similar.
- El análisis de los datos es el apartado más abierto de todos. Depende principalmente de la distribución que se presupone que siguen los datos.

En el caso de este estudio, se ha seguido este mismo criterio para la redacción de las recomendaciones.

Capítulo 9

Justificación de las recomendaciones

En este capítulo se explican las decisiones tomadas para las recomendaciones propuestas en el protocolo. Para facilitar el seguimiento de las mismas, se realizan de acuerdo a cada una de las categorías ya explicadas.

En las tablas 7.1 a 7.5 presentadas en el capítulo de “Clasificación de los estudios” se dispone de una vista simplificada de las características analizadas para facilitar la comparación.

Inteligibilidad

En esta categoría se han analizado los estudios [22], [29], [32] y [33]; así como la información relativa de las recomendaciones [1], [2] y [19].

En ellos, el número de participantes oscila entre 8 y 51 personas. Mientras que las recomendaciones determinan que 20 personas expertas son suficientes. Este último valor está aproximadamente en la media de los valores de los estudios, por lo que se opta por este valor, que además refuerza nuestro objetivo de seguir las recomendaciones dentro de lo posible.

El tipo de test utilizado en los estudios varía según el objeto específico del estudio, pero están comprendidos entre los test binomiales y de escalas numéricas que coinciden con las recomendaciones estudiadas.

La duración de las señales utilizadas en los estudios están comprendidas entre los 5 y los 15 segundos. Las recomendaciones estudiadas determinan una duración entre los 10 y los 25 segundos. En este caso, se decide acotar la duración entre 10 y 20 segundos para estar dentro del margen de las recomendaciones.

La duración de la sesión, en los casos en los que se proporciona dicha información, no supera los 5 minutos. Estos valores están dentro de los valores de las recomendaciones que fijan el máximo en torno a los 20-30 minutos aplicando pausas. A su vez, libros consultados como [5] fijan estos mismos límites para no producir fatiga auditiva. Por lo que se decide aplicar las recomendaciones de los organismos internacionales.

Para el análisis de los datos es donde más variaciones se encuentran, puesto que el sistema a utilizar depende de la distribución que sigan los datos. Por ello se plantean recomendaciones en función de las mismas, pero sí que se establece la necesidad de calcular los intervalos de confianza tal y como se refleja en las recomendaciones.

Sensibilidad

Para esta categoría se han considerado los experimentos [9], [10], [21], [24], [25], [28], [31] y [41]; así como las recomendaciones [1], [2] y [19].

En este caso, el número de participantes oscila entre 7 y 23 personas. Esto coincide en parte con las recomendaciones que fijan que el número debe estar entre las 10 y las 20 personas según si son expertas o no (aunque siempre se prefiere que sean expertas).

EL tipo de test utilizado en los estudios se divide principalmente en test binomiales y test de escalas numéricas. Según el objeto del ensayo se utiliza un tipo de test perteneciente a cada uno de ellos como el 2-AFC o MUSHRA, respectivamente.

La duración de las señales en los estudios oscila entre 1 segundo y 15 segundos. En este caso, los datos contrastan con las recomendaciones de la UIT que establecen la duración entre los 10 y los 20 segundos. Por lo que se decide seguir las indicaciones de esta institución frente a los estudios.

La duración de la sesión, en los estudios que proporcionan esta información, varía enormemente desde los 8 hasta los 85 minutos. Pese a que este caso tan extenso se hace con descansos, excede por mucho las recomendaciones de la UIT y de libros como [5] donde fijan el máximo de una sesión en 30 minutos para prevenir la fatiga auditiva.

Para el análisis de datos vuelve a haber mucha variedad. En muchos casos, no se especifica el tipo de análisis realizado, mientras que en los que sí se especifica, el sistema a utilizar depende de la distribución de los datos. Por ello, la recomendación queda referida al tipo de los mismos, haciendo necesario en cualquier caso el cálculo de los intervalos de confianza de acuerdo con las recomendaciones de los organismos internacionales.

Reverberación

En la categoría de reverberación se han considerado los estudios [8], [10], [26] y [30] y las recomendaciones [1], [19] y [2].

El número de participantes por experimento oscila entre los 6 y 26 personas, en su mayoría expertas. Esto entra en consonancia con los valores recomendados por la UIT de entre 10 y 20 personas según su nivel de experiencia, aunque siempre se prefiere que sean expertas.

Los tipos de test utilizados son los binomiales y los de escala numérica. Estos se corresponden con las recomendaciones de la UIT.

La duración de las señales varía entre los 0.6 y los 15 segundos. Las recomendaciones estipulan que la duración no debe exceder los 15 segundos. Por lo que se corrobora este criterio.

Para esta categoría no hay información sobre la duración de las sesiones como para establecer una comparación respecto a las recomendaciones, por lo que se aplica el criterio de estas últimas que fijan que la duración de una sesión no debe exceder los 30 minutos y siendo adecuado la realización de pausas durante la misma.

Al igual que en el resto de las categorías, el análisis de los datos recomendado depende del tipo de distribución de los datos y se toman las recomendaciones de los organismos internacionales seleccionados en este estudio.

Localización

En esta categoría se analizan los estudios [23], [36], [38] y [40]; además de las recomendaciones [1], [2] y [19].

El número de participantes en los experimentos está comprendido entre las 5 y 19 personas, sólo se hay menos de 10 participante en un caso, por lo que se puede suponerse que se sigue la recomendación de la UIT de 10-20 personas. Aunque no hay información concluyente sobre la experiencia de los participantes, se sigue utilizando las recomendaciones de los organismos internacionales.

El tipo de tests utilizados en esta categoría se centran, en su mayoría, en asignar valores numéricos, ya sean continuos o discretos, a una circunferencia alrededor del oyente para determinar la procedencia de un estímulo.

En esta categoría, la duración de las señales es menor que en el resto de casos. En ningún caso superan los 5 segundos.

La duración de las sesiones siguen las recomendaciones de la UIT y los libros de referencia de no sobrepasar los 30 minutos para evitar la fatiga auditiva.

Al igual que en otras categorías, en análisis de los datos depende de la distribución de los mismos. No obstante, en esta categoría puede ser habitual la utilización de sistemas estadísticos esféricos.

Molestia

Para la categoría de la molestia auditiva se han analizado los estudios [35], [39] y [42]. Por otro lado, se consideran las recomendaciones [1], [2] y [19].

Se observa que el número de personas utilizadas en estos ensayos varían entre 32 y 60 participantes. Por lo que se fijan unos ratios de entre 30-40 personas. En este caso concreto, la experiencia no es un factor tan determinante, por lo que pueden utilizarse cualquier tipo de personas.

El tipo de test preferido para esta categoría son las escalas numéricas, por lo que se siguen las recomendaciones de la UIT al respecto.

La duración de las señales cumplen la recomendación de estar entre los 10 y 20 segundos en los casos estudiados.

La duración de las sesiones vienen fijadas, al igual que en las categorías anteriores, con las limitaciones de 30 minutos para evitar la fatiga auditiva que estipulan los libros de referencia y los organismos internacionales.

Por último, el análisis de los datos depende de la distribución supuesta de los mismos, por lo que se plantean soluciones de acuerdo a las recomendaciones.

Capítulo 10

Ejemplo de aplicación del protocolo

En este capítulo se realiza una demostración real del protocolo desarrollado. Para ello, se realiza un estudio subjetivo de audio basado en parte del proyecto de fin de Máster realizado por Guillermo García Barrios [44]. Para la realización de dicha demostración, se siguen los pasos fijados en el protocolo y, al final, se comparan los resultados de ambos estudios.

10.1. Objetivo del estudio

El ensayo consiste en comprobar si la capacidad para distinguir que una pieza está interpretada por uno más cantantes está ligada, preferentemente, a las variaciones de frecuencia (tonalidad) o a la forma de onda (amplitud).

Para ello, se utilizan tres tipos de señales:

- Señales musicales interpretadas por uno, dos, tres o más personas.
- Señales musicales interpretadas por una persona a la que se le aplica un algoritmo que modifica la tonalidad de la voz para que parezcan interpretadas por dos, tres o más personas.
- Señales musicales interpretadas por una persona a la que se le aplica un algoritmo que modifica la amplitud de la voz para que parezcan interpretadas por dos, tres o más personas.

Con ellas se generan piezas musicales con las que el oyente deberá tomar sus decisiones. Estas decisiones estarán hechas en base a dos situaciones en particular.

En la primera situación, el oyente escucha las piezas de forma individual y determina si es capaz de percibir, uno, dos o tres o más oyentes.

En la segunda situación, el participante escuchará diferentes parejas de audios y deberá seleccionar en cuál de las dos opciones percibe un mayor número de cantantes.

10.2. Características del estudio según el protocolo

En este apartado se siguen las recomendaciones establecidas en el protocolo de medidas subjetivas de audio ya definido. Para ello, se comienza respondiendo a las preguntas que plantea:

¿En qué categoría se engloba mi estudio?

Este estudio se centra en analizar la capacidad de los participantes para distinguir si una pieza está interpretada por uno o más interpretes. Por tanto, parece lógico pensar que este estudio está comprendido dentro de la categoría de “Sensibilidad”.

A partir de este momento, las siguientes preguntas se responden atendiendo a las recomendaciones generales y específicas para esta categoría.

¿Qué tipo de test es el más adecuado?

En la categoría en la que se engloba este estudio, se recomienda la utilización de escalas numéricas y, en caso de que se quiera escoger entre dos opciones, utilizar el tipo de test 2-AFC.

El experimento está dividido en 2 partes:

En la primera, el oyente debe escoger cuántos cantantes puede distinguir en su escucha (1, 2 o 3 o más). En este apartado se decide utilizar el test de escalas numéricas para obtener los datos.

Por otro lado, en la segunda parte el participante tiene que decidir qué audio, entre dos opciones, está interpretado por más cantantes. Para esta parte el test a utilizar es el 2-AFC.

¿Participantes expertos o inexpertos?, ¿cuántos necesito?

Para los estudios dentro de la categoría de “Sensibilidad”, se recomienda la participación de un mínimo de 20 personas y, a poder ser, expertas. Del mismo modo, se recomienda la realización de un entrenamiento previo para todos los participantes. El protocolo determina que la duración de la sesión no debe exceder los 30 minutos de duración y deben disponer de descansos entre pruebas. También se fija que los participantes pueden conmutar las señales y repetir la reproducción de las mismas.

¿Qué tipo de señal utilizo?

La respuesta a esta pregunta está limitada ya que las señales a utilizar ya están generadas. La única limitación establecida en el protocolo es que la duración de las señales deben rondar entre los 10 y los 15 segundos. Esto es algo que cumplen aproximadamente los audios utilizados, por lo que no se precisan más consideraciones.

¿Cuál es el tipo de análisis de datos?

Para este estudio se determinan dos tipos de análisis de datos diferentes. Uno para cada parte del ensayo.

En la primera parte se realiza un estudio de las medias de los datos obtenidos, a través del sistema Wilcoxon. De esta forma, se calcula la media de cantantes percibidos para cada uno de los audios y se compara la desviación de la misma con otros audios con respecto a la media para otro audio. Después, se toman diferentes combinaciones de los valores individuales para cada audio y se vuelven a comparar las medias. De esta forma se determina si esas variaciones son significativas o no.

En la segunda parte, al disponer de un sistema de respuestas binario se procede a utilizar los procedimientos recomendados por la norma ISO-10399:2018 [20]. La norma establece que para sistemas binarios se debe consultar la tabla 10.1 y, en función del número de personas que establezcan la diferencia se le adjudica un determinado valor para el la probabilidad α .

n	Alpha-Risk (α)				
	0.2	0.1	0.05	0.01	0.001
6	5	6	6	-	-
7	6	6	7	7	-
8	6	7	7	8	-
9	7	7	8	9	-
10	7	8	9	10	10
11	8	9	9	10	11
12	8	9	10	11	12
13	9	10	10	12	13
14	10	10	11	12	13
15	10	11	12	13	14
16	11	12	12	14	15
17	11	12	13	14	16
18	12	13	13	15	16
19	12	13	13	15	17
20	13	14	15	16	18
21	13	14	15	17	18
22	13	14	15	17	19
23	15	16	16	18	20
24	15	16	17	19	20
25	16	17	18	19	21
...

Tabla 10.1: Valores de *Alpha-Risk* en función del número de personas y el número de respuestas positivas.

10.3. Procedimiento del estudio

Las personas que participan en el estudio firman un documento de consentimiento en el que dan su aprobación para la participación en el estudio.

A continuación, se les muestra en el ordenador una ventana en el que introducen su nombre y aseguran que tienen experiencia musical y que no sufren de ninguna discapacidad auditiva que les impida realizar el test.

Después, se abrirá una ventana en la que se entrena, de formar guiada, al oyente con algunos audios en los que participan uno, dos y tres o más cantantes.

Una vez realizado el entrenamiento, los oyentes realizan la primera parte del experimento. Deben marcar si, cada una de las 18 piezas seleccionadas, está interpretada por 1, 2 o 3 o más cantantes, respectivamente. Cada pieza puede repetirse cuantas veces se quiera, pero será obligatorio escuchar la pieza por completo la primera vez.

Tras acabar las 18 piezas, se procederá a realizar un pequeño descanso antes de comenzar la segunda parte. En esta parte, el sujeto deberá elegir qué pieza, de entre dos, está interpretada por más cantantes. Al igual que en el caso anterior, el oyente puede repetir la reproducción tantas veces como quiera, pero tendrá que escuchar mínimo una vez cada pieza de forma completa. Esta segunda parte concluye tras 18 casos distintos.

Una vez concluída esta parte, se marca el tiempo que ha durado la sesión y el sujeto podrá abandonar la prueba.

10.4. Material del estudio

Para la realización del estudio se ha necesitado el siguiente material:

- Ordenador con sistemas operativo DEBIAN, 4GB de RAM y procesador INTEL i5.

- Auriculares TASCAM TH-02.
- Programa MATLAB R2019b.
- Programa LibreOffice Calc.

10.5. Resultados del estudio

El estudio ha sido realizado por 23 personas (14 mujeres y 9 hombres); de las cuales 17 personas tienen formación musical avanzada. Las otras 6 personas disponen, en su lugar, de experiencia en el campo de la acústica y en la realización de test subjetivos de audio.

Todas las pruebas se realizaron sin problemas reseñables.

A continuación, se comentan los resultados obtenidos para las dos partes del test. Las respuestas íntegras de los participantes pueden consultarse en los anexos 1 y 2.

10.5.1. Resultados de la parte 1

En la primera parte del estudio, los 23 participantes han seleccionado el número de cantantes que perciben en cada uno de los 18 audios que se le han propuesto. Cada uno de estos audios se encuentra incluido dentro de una de las categorías de audio ya presentadas: audios naturales, audios modificados tonalmente y audios modificados en amplitud.

En las figuras 10.1, 10.2 y 10.3 se presentan las respuestas obtenidas para cada tipo de audio en función del número de oyentes teóricos que tenían dichos audios.

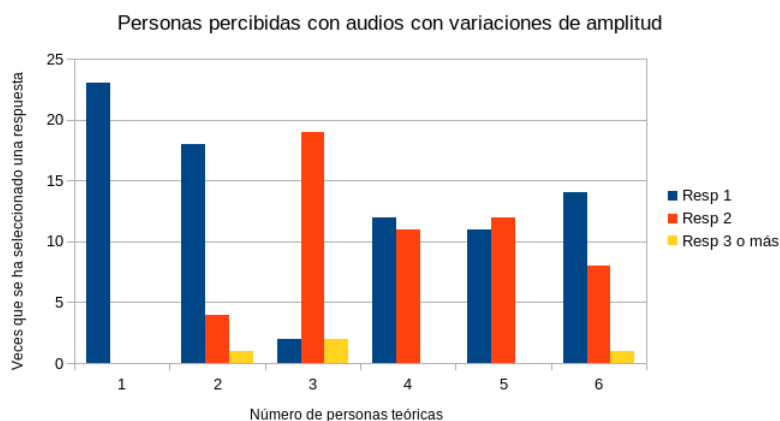


Figura 10.1: Número de oyentes percibidos en función del número de oyentes teóricos para audios modificados en amplitud.

10.5.2. Resultados de la parte 2

En esta parte, los mismos 23 participantes han escogido el audio en el que percibían más cantantes entre dos opciones que se les presentaba. Estos audios están categorizados, de nuevo, dentro de las mismas categorías: audios naturales, audios con modificaciones tonales y audios con modificaciones de amplitud.

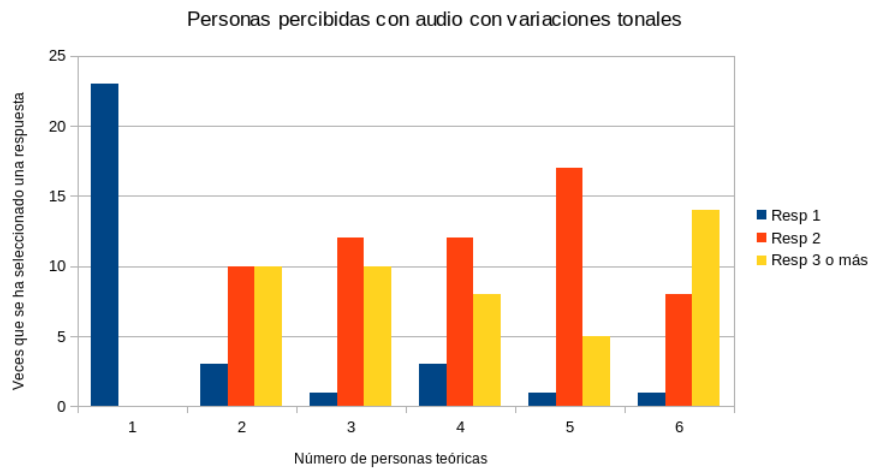


Figura 10.2: Número de oyentes percibidos en función del número de oyentes teóricos para audios modificados tonalmente.

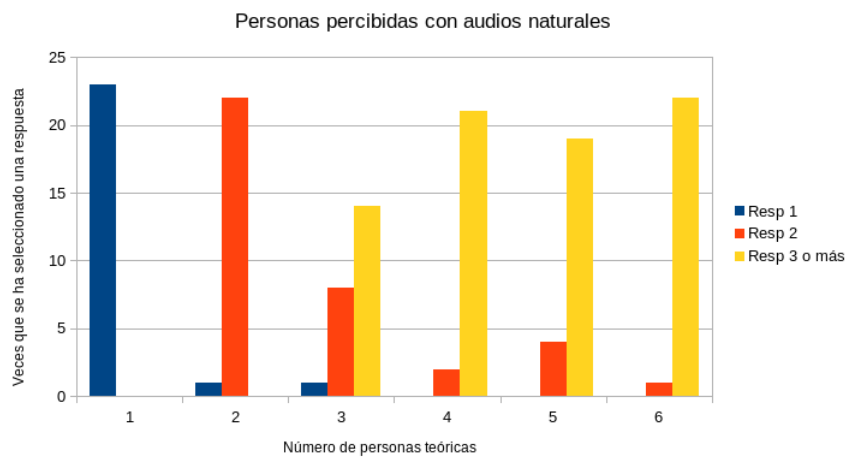


Figura 10.3: Número de oyentes percibidos en función del número de oyentes teóricos para audios naturales.

De esta forma, se pueden dividir los resultados en 3 comparaciones: audios naturales con audios modificados tonalmente, audios modificados tonalmente con audios modificados en amplitud y audios naturales con audios modificados en amplitud. Dentro de cada categoría se recogen el número de veces que se ha seleccionado una opción con respecto de la otra obteniéndose la tabla 10.2.

Amplitud	Tono	Natural	Total	Dudas
22	116	-	138	50
-	46	92	138	69
13	-	125	138	44

Tabla 10.2: Número de respuestas seleccionadas para cada tipo de audio cuando es comparada con otro de los tipos.

Para facilitar la visualización de estos datos se incluye a su vez la figura 10.4 que muestra los mismos resultados en forma de gráficos:

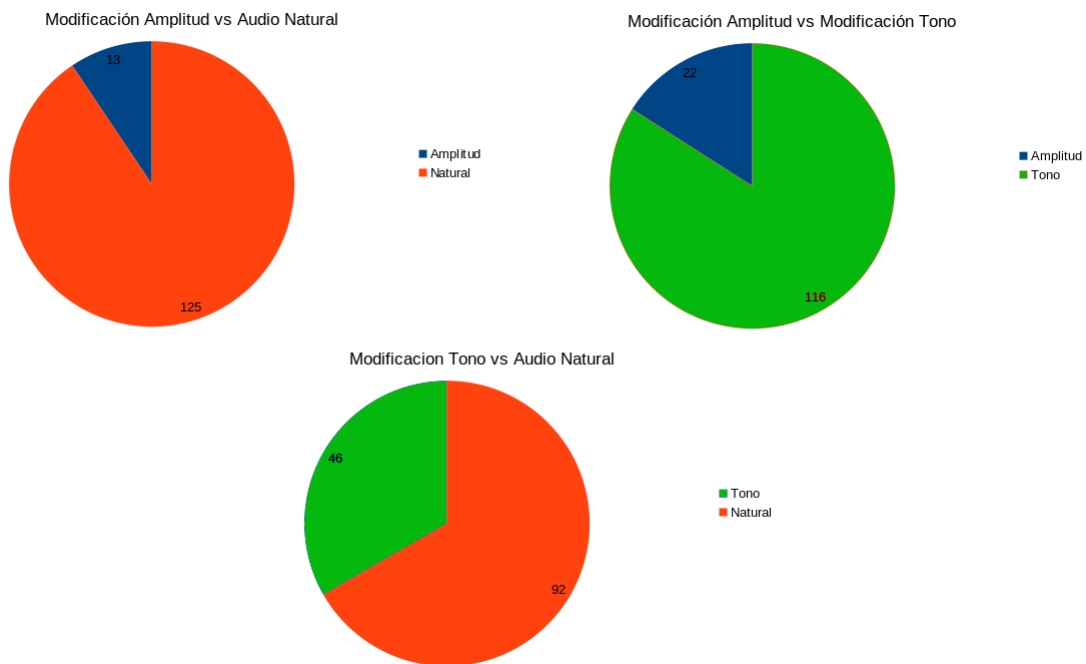


Figura 10.4: Respuestas acumuladas para la percepción de qué tipo de audio dispone de más cantantes al compararse dos tipos de audio.

10.6. Análisis de los resultados

10.6.1. Parte 1

En primer lugar, se realiza el análisis de los datos obtenidos mediante un script de MATLAB que permite realizar de forma automática el método *Wilcoxon* de análisis de medias. Como resultado de dicho análisis se obtiene la tabla 10.3.

En dicha tabla se presentan los datos en función de qué grupos de datos se estén comparando siendo “Tono/Amp” la comparación entre señales modificadas tonalmente y señales modificadas en amplitud;

	Prob. media(1) > media(2)	Prob. media(2) > media(1)	Precisión(1)	Precisión(2)
Tono/Amp1	1.000	1.000	0.000	0.000
Tono/Amp2	0.000	1.000	0.000	0.000
Tono/Amp3	0.009	0.996	0.006	0.004
Tono/Amp4	0.001	0.999	0.002	0.002
Tono/Amp5	0.000	1.000	0.000	0.000
Tono/Amp6	0.000	1.000	0.000	0.000
Nat/Tono1	1.000	1.000	0.000	0.000
Nat/Tono2	1.000	0.000	0.000	0.000
Nat/Tono3	0.252	0.837	0.027	0.023
Nat/Tono4	0.000	1.000	0.000	0.000
Nat/Tono5	0.000	1.000	0.000	0.000
Nat/Tono6	0.009	0.998	0.006	0.003
Nat/Amp1	1.000	1.000	0.000	0.000
Nat/Amp2	0.000	1.000	0.000	0.000
Nat/Amp3	0.000	1.000	0.000	0.000
Nat/Amp4	0.000	1.000	0.000	0.000
Nat/Amp5	0.000	1.000	0.000	0.000
Nat/Amp6	0.000	1.000	0.000	0.000
Nat12	1.000	0.000	0.000	0.000
Nat23	1.000	0.000	0.000	0.000
Nat24	1.000	0.000	0.000	0.000
Nat25	1.000	0.000	0.000	0.000
Tono12	1.000	0.000	0.000	0.000
Tono23	0.550	0.568	0.031	0.031
Amp12	1.000	0.029	0.000	0.011
Amp23	1.000	0.000	0.000	0.000

Tabla 10.3: Resultado del análisis de resultados de la parte 1 por el método Wilcoxon.

“Nat/Tono” la comparación entre señales naturales y señales modificadas tonalmente. “Nat/Amp” se corresponde, por su parte, con la comparación entre señales naturales y señales modificadas en amplitud. El número que les sigue se corresponde con el número de cantantes teóricos que tienen ambos audios. De esta forma, “Nat/Amp2” serían los resultados de la comparación de medias por el método *Wilcoxon* de audios naturales con audios modificados en amplitud con 2 cantantes teóricos.

Por otro lado, también se comparan las medias para diferente número de cantantes si el tipo de audio es el mismo. Estos casos están marcados de la forma “TipoXY”; siendo “Tipo” una de las categorías de audio, y “XY” los dos números de cantantes teóricos analizados.

Es importante remarcar que en este último caso sólo se han analizado comparaciones para casos de 1, 2 y 3 personas, porque interesa ver si los participantes han conseguido diferenciar entre estos casos. Por el tipo de respuestas que pueden dar (1, 2 y 3 o más), no tiene sentido intentar analizar casos más allá con este método.

Conclusiones del análisis de la parte 1

A la vista de los resultados de la tabla 10.3 y las figuras 10.1, 10.2 y 10.3 se puede deducir que los resultados para un cantante teórico son los mismos para todos los tipos de audio. Esto se ve reflejado en el análisis de Wilcoxon con el resultado 1-1. Mostrando que las medias son las mismas para todos los casos y por tanto no existen diferencias.

Para el resto de casos de comparación entre tipos de audio, se buscan aquellos casos en los que la probabilidad sea menor que 0.01, lo que significa que existe una diferencia significativa al 99 %. Esto se cumple para todos los casos supuestos, menos para la comparación “Nat/Tono3”, donde se ha obtenido una probabilidad de 0.25. En este caso, se puede concluir que los oyentes no son capaces de distinguir de forma segura un tipo de audio de otro.

Para el resto de casos se puede concluir que las diferencias son significativas y no son debidas al azar tal y cómo puede comprobarse de las diferencias que se observan las figuras ya presentadas que confirman este análisis.

Analizando las comparaciones entre diferente número de cantantes para el mismo tipo de audio, se observan valores de 1-0, de forma que se confirma que los oyentes son capaces de distinguir entre unos audios y otros. El único caso en el que esto no se produce es para el caso “Tono23” donde con un valor de 0.55 se puede concluir que las diferencias son debidas al azar.

Observando las figuras, se corroboran estas conclusiones ya que los valores para audios con modificaciones de tono para 2 y 3 cantantes son prácticamente iguales.

También se observa que las variaciones tonales son muy efectivas para que los oyentes distingan entre 1 y 2 cantantes, pero no tanto para 2 y 3 o más, aunque puede intuirse una tendencia conforme se aumentan el número de cantantes teóricos, pero no lo suficiente para que sea significativo de forma estadística.

En el caso de los audios modificados en amplitud, se observa que la diferenciación entre cantantes salta de forma clara de 1 a 2 cuando se simulan 3 cantantes. En el caso comparativo de 1 y 2 cantantes las diferencias parecen lo suficientemente significativas desde el punto de vista estadístico, pero no se cumple del todo con una probabilidad menor que 0.01. Esto se ve especialmente en la figura 10.1.

Por último, para los audios naturales se observa que los oyentes son capaces de distinguir perfectamente entre los audios de 1, 2 y 3 cantantes. Incluso esta percepción sigue mejorando conforme más cantantes se añaden. Esto se queda reflejado en los resultados del análisis de *Wilcoxon*, con todos los resultados de la forma 1-0 y lo mismo se observa en la figura 10.3.

10.6.2. Parte 2

En esta parte, no se han tenido en cuenta los resultados para comparaciones entre audios de 1 cantante, porque no tienen sentido lógico si los algoritmos están bien desarrollados. De esta forma, se obtiene la tabla 10.4.

Amplitud	Tono	TOTAL	DUDA
22	93	115	27
Natural	Tono	TOTAL	DUDA
86	29	115	48
Natural	Amplitud	TOTAL	DUDA
108	7	115	21

Tabla 10.4: Número de respuestas para cada tipo de audio cuando es comparado con otro.

El siguiente paso es calcular el número mínimo de respuestas necesarias en una opción para determinar que existe una diferencia significativa entre ambas opciones. Para ello, se recurre a la norma [20] y la tabla 10.1. Como el número de datos es muy alto (115), hay que aplicar la ecuación 10.1 para calcular el número de respuestas necesarias. Esta ecuación tiene la forma:

$$x = (n/2) + z\sqrt{n/4} \quad (10.1)$$

Siendo:

- x : el número mínimo de respuestas necesarias para determinar que existe una diferencia significativa.

- n : el número de respuestas totales.
- z : una constante en función del valor de α .

Aplicando los datos de nuestro estudio se obtienen los siguientes valores de x , representados en la tabla 10.5:

Diferencia para $n=115$				
$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,001$
62	64	66	70	74

Tabla 10.5: Numero de veces que debe ser escogida una opción necesarias para obtener un determinado *Alpha-Risk* teniendo un total de 115 comparaciones.

A continuación, se plantea realizar los mismos cálculos, pero obviando las respuestas marcadas con la opción “con duda” y ver si esto puede afectar a los resultados. Estos cálculos quedan reflejados en la tabla 10.6.

Conclusiones del análisis de la parte 2

A la vista de los resultados obtenidos tras el análisis de los datos, se puede concluir que para cada una de las tres comparaciones supuestas, se obtiene un valor de $\alpha=0.001$. Este valor se consigue de forma holgada, por lo que se puede afirmar que existe una diferencia significativa entre ambos cada uno de los tipos de audio cuando se compara con el otro; siendo los audios naturales con los que mayor número de cantantes se perciben frente a cualquier tipo de audio.

Por otro lado, entre audios modificados tonalmente y modificados en amplitud, los participantes perciben un mayor número de cantantes en el primer tipo de audios también con un valor de $\alpha=0.001$.

Por último, la supresión de las respuestas emitidas “con duda” mantienen estos resultados e incluso los refuerzan.

10.7. Conclusiones del estudio

A la vista del análisis de los datos, se puede concluir que los audios naturales son los que permiten percibir un mayor número de cantantes; tanto si se escuchan por separado, como si se hacen comparados con otros tipos de audio.

Por otro lado, se observa que las modificaciones tonales son más efectivas que las modificaciones en amplitud para que los oyentes perciban un mayor número de cantantes. No obstante, los resultados que se obtienen con esta técnica están lejos de los naturales; especialmente para casos en los que se pretende simular a partir de más de 3 cantantes.

Amplitud	Tono	Total	$\alpha=0.2$	$\alpha=0.1$	$\alpha=0.05$	$\alpha=0.01$	$\alpha=0.001$
8	80	88	48	50	52	55	58
Natural	Tono	Total	$\alpha=0.2$	$\alpha=0.1$	$\alpha=0.05$	$\alpha=0.01$	$\alpha=0.001$
54	13	67	37	39	40	43	46
Natural	Amplitud	Total	$\alpha=0.2$	$\alpha=0.1$	$\alpha=0.05$	$\alpha=0.01$	$\alpha=0.001$
94	0	94	51	53	55	58	62

Tabla 10.6: Número de veces que una opción es escogida en una comparación suprimiendo las respuestas dudosas y cálculo de los valores necesarios para distintos valores de *Alpha-Risk*.

Por último, los audios modificados en amplitud son los que peor resultado tienen para la detección de diferente número de cantantes, quedando estancados para el caso de 2 personas distintas cantando.

Con estas conclusiones, se puede afirmar que, para generar audios artificiales que simulen el canto de varias personas, las modificaciones tonales son un buen primer paso, pero se hace necesario la inclusión de otras técnicas para obtener resultados que puedan competir con los audios naturales.

En futuras investigaciones se podría plantear el efecto de la mezcla de variaciones tonales y amplitud. También se podrían añadir otros efectos subjetivos como la espacialidad de diferentes voces para mejorar el efecto general.

Capítulo 11

Comparación con el estudio original

En este capítulo se comparan los resultados obtenidos en el test subjetivo con los resultados del estudio realizado por Guillermo García Barrios como parte de su proyecto de fin de Máster.

11.1. Características del estudio original

El estudio realizado por Guillermo García Barrios es realizado por 26 personas. 12 de ellas dentro de la categoría de expertas y 14 sin experiencia. En cualquier caso, realizan un entrenamiento previo. La duración de las señales oscila entre los 7 y los 16 segundos aproximadamente. No está determinada la duración de la sesión, aunque aseguran que no excede los 30 minutos de duración.

En el estudio se realizan dos tipos de test: uno de escala numérica y un test de múltiples opciones. El análisis de los datos obtenidos para ambos tests se realiza mediante el estudio de los porcentajes de los resultados. No se realizan otros sistemas estadísticos ni intervalos de confianza.

Con el fin de facilitar la comparación entre las características del experimento original y el del protocolo, se presenta la tabla 11.1.

Estudio	Estudio Original	Estudio con protocolo
Categoría	Sin determinar.	Sensibilidad.
Tipo test	Escala numérica y opción múltiple.	Escalas numéricas y 2-AFC.
Nº Participantes	26	23
Tipo de oyente	Expertos y no expertos.	Expertos.
Duración señal	8-16s	8-16s
Duración sesión	máx 30min.	máx 30min.
Entrenamiento previo	Sí.	Sí.
Análisis de datos	Porcentajes.	Análisis de medias (Wilcoxon) y evaluación de estudios binomiales (<i>Alpha-Risk</i>).
Tipo de estímulo	Señales grabadas y generadas en Matlab	Señales grabadas y generadas en Matlab
Tipo de experimento (si se muestra una sola vez, lo modifica el usuario, etc.)	No se pueden repetir los audios.	Los oyentes pueden conmutar entre señales y repetirlas.

Tabla 11.1: Comparación entre estudio original y siguiendo protocolo

11.2. Comparación de resultados

En ambos estudios se llega a la conclusión de que las modificaciones tonales son más efectivas que las modificaciones en amplitud a la hora de simular el canto de múltiples personas y que sean percibidas por los oyentes. Del mismo modo, se obtienen los mismos resultados sobre la incapacidad de estas últimas para poder simular más de 2 cantantes.

Al mismo tiempo, ambos estudios obtienen resultados similares en la comparación de audios naturales con audios con diferencias tonales. No obstante, en el estudio de Guillermo se considera que existen diferencias notables para los casos 3 y 4 cantantes en esta comparación. En el estudio de Guillermo se considera que el método de modificación de tonos es lo suficientemente parecido a los audios naturales. En el caso de nuestro estudio, sólo se puede aplicar para el caso de 3 cantantes, pero no para más allá, donde el análisis estadístico muestra que las diferencias son demasiado significativas.

Sobre todas estas consideraciones, hay que tener en cuenta que el procedimiento del estudio es distinto. El estudio de Guillermo limitaba las escuchas de cada una de las señales, mientras que el nuestro podían repetirse las escuchas tantas veces como se quisiera.

También hay que tener en cuenta que el análisis de datos realizado por Guillermo es más sencillo, ya que sólo examina el porcentaje de respuestas, mientras que este estudio utiliza metodologías estadísticas más complejas como el método de *Wilcoxon* y la recomendación de la ISO 10399. Estas diferencias probablemente sean las causantes en las pequeñas diferencias entre los resultados obtenidos y las conclusiones

Capítulo 12

Conclusiones

El protocolo desarrollado recoge recomendaciones para la mayoría de tipos de test subjetivos de audio que se realizan en la actualidad y que están basados en test estadísticos.

Del mismo modo, dentro del protocolo se recoge, de forma sencilla, toda aquella información necesaria para el desarrollo de los estudios subjetivos de audio. También se sigue, dentro de lo posible, las recomendaciones establecidas por los organismos internacionales en esta materia.

Este procedimiento es demostrado en test subjetivo realizado como aplicación. En dicho test se consiguen resultados numéricos similares al original, pero se ha conseguido profundizar de forma sencilla en el análisis estadístico, mejorando las conclusiones obtenidas.

Es importante remarcar que, durante la búsqueda estudios previos que se han usado como base en este protocolo, se ha encontrado una gran disparidad de criterios a la hora de realizar estudios dentro del campo de la psicoacústica y, en muchos casos, ni siquiera se especificaba cuáles eran dichos criterios. Esto pone de manifiesto la necesidad de elaboración continuar el trabajo en este tipo de protocolos con el fin de lograr una mayor homogeneidad en los procedimientos que se siguen en este campo.

En futuras investigaciones se podrían incluir nuevas categorías que incluyan aquellos tipos de test subjetivos que pudieran haberse quedado fuera. También se podrían crear subcategorías dentro de las categorías ya creadas para dar recomendaciones un poco más específicas, o incluir más tipos de test estadísticos que sean muy utilizados para el análisis de los datos para futuras consultas.

Apéndice A

Resultados completos de la parte 1 del experimento

Audio	Singers	Type	Answer	Confidence
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/4-2-singers-different-pitch.wav	2	different-pitch	3	2
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	1
audios/4-5-singers-natural.wav	5	natural	3	2
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	2
audios/3-3-singers-different-pitch.wav	3	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	2	1
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/1-5-singers-same-pitch.wav	5	same-pitch	1	2
audios/4-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/1-3-singers-natural.wav	3	natural	3	2
audios/2-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/2-2-singers-same-pitch.wav	2	same-pitch	1	1
audios/4-1-singers-natural.wav	1	natural	1	2
audios/2-4-singers-different-pitch.wav	4	different-pitch	3	2
audios/4-3-singers-different-pitch.wav	3	different-pitch	3	1
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	2	1
audios/1-5-singers-different-pitch.wav	5	different-pitch	3	1
audios/1-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/4-1-singers-same-pitch.wav	1	same-pitch	1	2

Audio	Singers	Type	Answer	Confidence
audios/3-4-singers-natural.wav	4	natural	3	2
audios/3-5-singers-natural.wav	5	natural	3	2
audios/1-4-singers-same-pitch.wav	4	same-pitch	2	2
audios/4-6-singers-natural.wav	6	natural	3	2
audios/3-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/2-2-singers-natural.wav	2	natural	2	2
audios/3-5-singers-same-pitch.wav	5	same-pitch	2	2
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	1
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/2-4-singers-same-pitch.wav	4	same-pitch	1	1
audios/2-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	2	2
audios/3-1-singers-natural.wav	1	natural	1	2
audios/1-4-singers-natural.wav	4	natural	3	2
audios/2-3-singers-different-pitch.wav	3	different-pitch	2	2
audios/4-2-singers-natural.wav	2	natural	2	2
audios/1-4-singers-different-pitch.wav	4	different-pitch	3	2
audios/4-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/3-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	2	1
audios/3-5-singers-natural.wav	5	natural	3	1
audios/1-6-singers-natural.wav	6	natural	3	1
audios/2-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/3-5-singers-same-pitch.wav	5	same-pitch	1	1
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/4-2-singers-different-pitch.wav	2	different-pitch	2	2
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	1
audios/4-4-singers-different-pitch.wav	4	different-pitch	1	1
audios/4-5-singers-natural.wav	5	natural	3	2
audios/3-2-singers-natural.wav	2	natural	1	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	2
audios/3-3-singers-different-pitch.wav	3	different-pitch	2	2
audios/2-3-singers-natural.wav	3	natural	2	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/1-6-singers-natural.wav	6	natural	3	1
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/1-5-singers-same-pitch.wav	5	same-pitch	1	2
audios/4-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/1-3-singers-natural.wav	3	natural	2	2
audios/2-6-singers-different-pitch.wav	6	different-pitch	1	2

Audio	Singers	Type	Answer	Confidence
audios/2-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/4-1-singers-natural.wav	1	natural	1	2
audios/2-4-singers-different-pitch.wav	4	different-pitch	3	1
audios/4-3-singers-different-pitch.wav	3	different-pitch	2	2
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	2	1
audios/1-5-singers-different-pitch.wav	5	different-pitch	3	1
audios/1-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/4-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-4-singers-natural.wav	4	natural	3	2
audios/3-5-singers-natural.wav	5	natural	3	2
audios/1-4-singers-same-pitch.wav	4	same-pitch	1	1
audios/4-6-singers-natural.wav	6	natural	3	2
audios/3-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/2-2-singers-natural.wav	2	natural	2	2
audios/3-5-singers-same-pitch.wav	5	same-pitch	1	1
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/2-4-singers-same-pitch.wav	4	same-pitch	2	2
audios/2-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	3	2
audios/3-1-singers-natural.wav	1	natural	1	2
audios/1-4-singers-natural.wav	4	natural	2	1
audios/2-3-singers-different-pitch.wav	3	different-pitch	2	2
audios/4-2-singers-natural.wav	2	natural	2	2
audios/1-4-singers-different-pitch.wav	4	different-pitch	3	2
audios/4-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/3-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	2	2
audios/3-5-singers-natural.wav	5	natural	3	2
audios/1-6-singers-natural.wav	6	natural	3	1
audios/2-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/3-5-singers-same-pitch.wav	5	same-pitch	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/2-6-singers-same-pitch.wav	6	same-pitch	2	1
audios/2-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	1
audios/2-4-singers-natural.wav	4	natural	3	2
audios/4-1-singers-natural.wav	1	natural	1	2
audios/2-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	2	1
audios/3-4-singers-same-pitch.wav	4	same-pitch	1	1

Audio	Singers	Type	Answer	Confidence
audios/3-3-singers-natural.wav	3	natural	3	2
audios/3-2-singers-different-pitch.wav	2	different-pitch	3	1
audios/3-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/1-3-singers-different-pitch.wav	3	different-pitch	2	1
audios/1-6-singers-natural.wav	6	natural	3	2
audios/4-5-singers-natural.wav	5	natural	3	2
audios/1-5-singers-same-pitch.wav	5	same-pitch	1	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/4-2-singers-different-pitch.wav	2	different-pitch	2	1
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	2	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	1
audios/4-5-singers-natural.wav	5	natural	2	2
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	2	1
audios/3-3-singers-different-pitch.wav	3	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	3	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	2	2
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	1
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	1
audios/4-6-singers-same-pitch.wav	6	same-pitch	2	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/4-2-singers-different-pitch.wav	2	different-pitch	2	2
audios/3-4-singers-natural.wav	4	natural	2	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	1
audios/4-4-singers-different-pitch.wav	4	different-pitch	1	1
audios/4-5-singers-natural.wav	5	natural	3	1
audios/3-2-singers-natural.wav	2	natural	2	1
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	1
audios/3-3-singers-different-pitch.wav	3	different-pitch	2	1
audios/2-3-singers-natural.wav	3	natural	2	1
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/1-6-singers-natural.wav	6	natural	2	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	1	1
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	2
audios/4-6-singers-same-pitch.wav	6	same-pitch	1	1
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/4-2-singers-different-pitch.wav	2	different-pitch	3	1

Audio	Singers	Type	Answer	Confidence
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	2
audios/4-5-singers-natural.wav	5	natural	2	1
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	2	1
audios/3-3-singers-different-pitch.wav	3	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	3	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	2	1
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	2
audios/4-6-singers-same-pitch.wav	6	same-pitch	1	1
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/4-2-singers-different-pitch.wav	2	different-pitch	3	2
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	1	1
audios/4-5-singers-natural.wav	5	natural	2	1
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	2
audios/3-3-singers-different-pitch.wav	3	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	1	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	2	1
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	1
audios/4-6-singers-same-pitch.wav	6	same-pitch	2	1
audios/1-3-singers-natural.wav	3	natural	2	2
audios/2-6-singers-different-pitch.wav	6	different-pitch	2	1
audios/2-2-singers-same-pitch.wav	2	same-pitch	1	1
audios/4-1-singers-natural.wav	1	natural	1	2
audios/2-4-singers-different-pitch.wav	4	different-pitch	2	2
audios/4-3-singers-different-pitch.wav	3	different-pitch	2	2
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/1-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/4-1-singers-same-pitch.wav	1	same-pitch	1	2

Audio	Singers	Type	Answer	Confidence
audios/3-4-singers-natural.wav	4	natural	3	1
audios/3-5-singers-natural.wav	5	natural	3	1
audios/1-4-singers-same-pitch.wav	4	same-pitch	2	1
audios/4-6-singers-natural.wav	6	natural	3	2
audios/3-3-singers-same-pitch.wav	3	same-pitch	1	1
audios/2-2-singers-natural.wav	2	natural	2	1
audios/3-5-singers-same-pitch.wav	5	same-pitch	1	2
audios/1-3-singers-different-pitch.wav	3	different-pitch	2	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/4-2-singers-natural.wav	2	natural	2	2
audios/4-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/3-5-singers-natural.wav	5	natural	3	2
audios/2-4-singers-same-pitch.wav	4	same-pitch	2	2
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/2-6-singers-natural.wav	6	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	2	1
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	1
audios/2-4-singers-natural.wav	4	natural	3	2
audios/3-3-singers-natural.wav	3	natural	3	2
audios/1-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/3-5-singers-same-pitch.wav	5	same-pitch	2	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/3-2-singers-different-pitch.wav	2	different-pitch	3	2
audios/4-5-singers-same-pitch.wav	5	same-pitch	1	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/2-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/4-4-singers-natural.wav	4	natural	3	1
audios/3-4-singers-different-pitch.wav	4	different-pitch	2	2
audios/2-2-singers-natural.wav	2	natural	2	2
audios/4-3-singers-natural.wav	3	natural	3	1
audios/4-1-singers-natural.wav	1	natural	1	2
audios/3-5-singers-natural.wav	5	natural	3	2
audios/3-6-singers-natural.wav	6	natural	3	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-4-singers-same-pitch.wav	4	same-pitch	2	2
audios/2-3-singers-different-pitch.wav	3	different-pitch	2	1
audios/2-2-singers-different-pitch.wav	2	different-pitch	1	1
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/4-5-singers-different-pitch.wav	5	different-pitch	3	1
audios/1-6-singers-different-pitch.wav	6	different-pitch	2	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	3	2
audios/4-2-singers-different-pitch.wav	2	different-pitch	2	2

Audio	Singers	Type	Answer	Confidence
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	3	2
audios/4-5-singers-natural.wav	5	natural	3	1
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	2
audios/3-3-singers-different-pitch.wav	3	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	3	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	1
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	1
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	2
audios/4-6-singers-same-pitch.wav	6	same-pitch	1	1
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/4-2-singers-different-pitch.wav	2	different-pitch	2	1
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	1
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	2
audios/4-5-singers-natural.wav	5	natural	3	2
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	2	1
audios/3-3-singers-different-pitch.wav	3	different-pitch	3	1
audios/2-3-singers-natural.wav	3	natural	2	1
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	2	2
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	3	2
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	1
audios/4-6-singers-same-pitch.wav	6	same-pitch	2	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/4-2-singers-different-pitch.wav	2	different-pitch	3	2
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	2	1
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	2
audios/4-5-singers-natural.wav	5	natural	3	2
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	1
audios/3-3-singers-different-pitch.wav	3	different-pitch	2	2
audios/2-3-singers-natural.wav	3	natural	2	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2

Audio	Singers	Type	Answer	Confidence
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/1-5-singers-same-pitch.wav	5	same-pitch	1	2
audios/4-6-singers-same-pitch.wav	6	same-pitch	2	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	3	2
audios/4-2-singers-different-pitch.wav	2	different-pitch	3	1
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	2
audios/4-5-singers-natural.wav	5	natural	3	2
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	2
audios/3-3-singers-different-pitch.wav	3	different-pitch	2	1
audios/2-3-singers-natural.wav	3	natural	3	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	1
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	1
audios/4-6-singers-same-pitch.wav	6	same-pitch	2	2
audios/1-3-singers-natural.wav	3	natural	3	1
audios/2-6-singers-different-pitch.wav	6	different-pitch	2	1
audios/2-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/4-1-singers-natural.wav	1	natural	1	2
audios/2-4-singers-different-pitch.wav	4	different-pitch	3	2
audios/4-3-singers-different-pitch.wav	3	different-pitch	3	2
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/1-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/4-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-4-singers-natural.wav	4	natural	3	2
audios/3-5-singers-natural.wav	5	natural	3	2
audios/1-4-singers-same-pitch.wav	4	same-pitch	2	2
audios/4-6-singers-natural.wav	6	natural	3	2
audios/3-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/2-2-singers-natural.wav	2	natural	2	2
audios/3-5-singers-same-pitch.wav	5	same-pitch	2	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	2
audios/4-2-singers-different-pitch.wav	2	different-pitch	3	2

Audio	Singers	Type	Answer	Confidence
audios/3-4-singers-natural.wav	4	natural	3	2
audios/1-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	2
audios/4-5-singers-natural.wav	5	natural	3	2
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-same-pitch.wav	4	same-pitch	1	1
audios/3-3-singers-different-pitch.wav	3	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	3	2
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	1
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-1-singers-natural.wav	1	natural	1	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	3	2
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	2
audios/4-6-singers-same-pitch.wav	6	same-pitch	2	2
audios/1-3-singers-natural.wav	3	natural	3	2
audios/2-6-singers-different-pitch.wav	6	different-pitch	2	1
audios/2-2-singers-same-pitch.wav	2	same-pitch	1	2
audios/4-1-singers-natural.wav	1	natural	1	2
audios/2-4-singers-different-pitch.wav	4	different-pitch	3	1
audios/4-3-singers-different-pitch.wav	3	different-pitch	2	1
audios/1-6-singers-same-pitch.wav	6	same-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	2	2
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/1-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/4-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-4-singers-natural.wav	4	natural	3	2
audios/3-5-singers-natural.wav	5	natural	2	1
audios/1-4-singers-same-pitch.wav	4	same-pitch	2	1
audios/4-6-singers-natural.wav	6	natural	3	2
audios/3-3-singers-same-pitch.wav	3	same-pitch	1	2
audios/2-2-singers-natural.wav	2	natural	2	2
audios/3-5-singers-same-pitch.wav	5	same-pitch	1	1
audios/1-6-singers-same-pitch.wav	6	same-pitch	2	1
audios/1-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/2-4-singers-same-pitch.wav	4	same-pitch	2	2
audios/2-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/2-2-singers-different-pitch.wav	2	different-pitch	3	1
audios/3-1-singers-natural.wav	1	natural	1	2
audios/1-4-singers-natural.wav	4	natural	3	2
audios/2-3-singers-different-pitch.wav	3	different-pitch	3	1
audios/4-2-singers-natural.wav	2	natural	2	2
audios/1-4-singers-different-pitch.wav	4	different-pitch	3	2
audios/4-2-singers-same-pitch.wav	2	same-pitch	1	2

Audio	Singers	Type	Answer	Confidence
audios/3-6-singers-different-pitch.wav	6	different-pitch	3	2
audios/2-3-singers-natural.wav	3	natural	3	2
audios/3-5-singers-natural.wav	5	natural	3	2
audios/1-6-singers-natural.wav	6	natural	3	2
audios/2-5-singers-different-pitch.wav	5	different-pitch	2	2
audios/4-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/3-5-singers-same-pitch.wav	5	same-pitch	1	1
audios/1-5-singers-different-pitch.wav	5	different-pitch	2	1
audios/3-1-singers-different-pitch.wav	1	different-pitch	1	2
audios/2-6-singers-same-pitch.wav	6	same-pitch	3	1
audios/2-6-singers-different-pitch.wav	6	different-pitch	2	1
audios/3-2-singers-natural.wav	2	natural	2	2
audios/4-4-singers-different-pitch.wav	4	different-pitch	2	1
audios/2-4-singers-natural.wav	4	natural	3	2
audios/4-1-singers-natural.wav	1	natural	1	2
audios/2-1-singers-same-pitch.wav	1	same-pitch	1	2
audios/3-2-singers-same-pitch.wav	2	same-pitch	1	1
audios/3-4-singers-same-pitch.wav	4	same-pitch	1	2
audios/3-3-singers-natural.wav	3	natural	3	1
audios/3-2-singers-different-pitch.wav	2	different-pitch	2	2
audios/3-3-singers-same-pitch.wav	3	same-pitch	2	1
audios/1-3-singers-different-pitch.wav	3	different-pitch	1	1
audios/1-6-singers-natural.wav	6	natural	3	1
audios/4-5-singers-natural.wav	5	natural	3	1
audios/1-5-singers-same-pitch.wav	5	same-pitch	2	1

Apéndice B

Resultados completos de la parte 2 del experimento

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	2
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	1	1	2	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	1
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-3-singers-natural.wav	audios/1-3-singers-different-pitch.wav	3	1	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-different-pitch.wav	1	2	2	1	1
audios/1-6-singers-same-pitch.wav	audios/1-6-singers-different-pitch.wav	6	2	1	1	2
audios/1-2-singers-different-pitch.wav	audios/1-2-singers-same-pitch.wav	2	2	1	2	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-natural.wav	1	2	3	3	1
audios/3-2-singers-same-pitch.wav	audios/3-2-singers-natural.wav	2	2	3	3	2
audios/1-5-singers-same-pitch.wav	audios/1-5-singers-different-pitch.wav	5	2	1	1	1
audios/1-5-singers-natural.wav	audios/1-5-singers-same-pitch.wav	5	1	3	3	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-different-pitch.wav	3	2	1	1	2
audios/1-4-singers-natural.wav	audios/1-4-singers-same-pitch.wav	4	1	3	3	2
audios/1-4-singers-natural.wav	audios/1-4-singers-different-pitch.wav	4	1	2	3	1
audios/3-6-singers-natural.wav	audios/3-6-singers-different-pitch.wav	6	1	2	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-natural.wav	5	2	2	3	2
audios/3-3-singers-same-pitch.wav	audios/3-3-singers-natural.wav	3	2	3	3	2
audios/2-4-singers-same-pitch.wav	audios/2-4-singers-different-pitch.wav	4	1	1	2	1

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-natural.wav	3	2	3	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-same-pitch.wav	5	1	1	1	1
audios/2-6-singers-natural.wav	audios/2-6-singers-different-pitch.wav	6	1	2	3	1
audios/4-1-singers-natural.wav	audios/4-1-singers-different-pitch.wav	1	1	2	3	1
audios/3-5-singers-natural.wav	audios/3-5-singers-different-pitch.wav	5	1	2	3	1
audios/4-5-singers-same-pitch.wav	audios/4-5-singers-natural.wav	5	2	3	3	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-same-pitch.wav	3	1	1	1	1
audios/2-2-singers-natural.wav	audios/2-2-singers-same-pitch.wav	2	1	3	3	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-natural.wav	1	2	3	3	1
audios/1-1-singers-same-pitch.wav	audios/1-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/3-2-singers-natural.wav	audios/3-2-singers-different-pitch.wav	2	2	2	1	1
audios/3-6-singers-same-pitch.wav	audios/3-6-singers-different-pitch.wav	6	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-4-singers-natural.wav	audios/4-4-singers-same-pitch.wav	4	1	3	3	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-natural.wav	4	2	2	3	1
audios/4-3-singers-natural.wav	audios/4-3-singers-different-pitch.wav	3	2	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	2
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	1	1	2	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	1
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-3-singers-natural.wav	audios/1-3-singers-different-pitch.wav	3	1	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-different-pitch.wav	1	2	2	1	1
audios/1-6-singers-same-pitch.wav	audios/1-6-singers-different-pitch.wav	6	2	1	1	2
audios/1-2-singers-different-pitch.wav	audios/1-2-singers-same-pitch.wav	2	1	1	1	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	1	2	3	1
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-natural.wav	1	1	3	2	1
audios/3-2-singers-same-pitch.wav	audios/3-2-singers-natural.wav	2	2	3	3	2
audios/1-5-singers-same-pitch.wav	audios/1-5-singers-different-pitch.wav	5	2	1	1	1

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/1-5-singers-natural.wav	audios/1-5-singers-same-pitch.wav	5	1	3	3	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-different-pitch.wav	1	1	1	2	1
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-different-pitch.wav	3	1	1	2	1
audios/1-4-singers-natural.wav	audios/1-4-singers-same-pitch.wav	4	1	3	3	2
audios/1-4-singers-natural.wav	audios/1-4-singers-different-pitch.wav	4	1	2	3	1
audios/3-6-singers-natural.wav	audios/3-6-singers-different-pitch.wav	6	1	2	3	1
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-natural.wav	5	2	2	3	2
audios/3-3-singers-same-pitch.wav	audios/3-3-singers-natural.wav	3	2	3	3	2
audios/2-4-singers-same-pitch.wav	audios/2-4-singers-different-pitch.wav	4	1	1	2	1
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-natural.wav	3	2	3	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-same-pitch.wav	5	1	1	1	2
audios/2-6-singers-natural.wav	audios/2-6-singers-different-pitch.wav	6	1	2	3	2
audios/4-1-singers-natural.wav	audios/4-1-singers-different-pitch.wav	1	1	2	3	1
audios/3-5-singers-natural.wav	audios/3-5-singers-different-pitch.wav	5	1	2	3	2
audios/4-5-singers-same-pitch.wav	audios/4-5-singers-natural.wav	5	2	3	3	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-same-pitch.wav	3	2	1	2	1
audios/2-2-singers-natural.wav	audios/2-2-singers-same-pitch.wav	2	2	3	2	1
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-natural.wav	1	2	3	3	1
audios/1-1-singers-same-pitch.wav	audios/1-1-singers-different-pitch.wav	1	1	1	2	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	1
audios/3-2-singers-natural.wav	audios/3-2-singers-different-pitch.wav	2	1	2	3	1

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/3-6-singers-same-pitch.wav	audios/3-6-singers-different-pitch.wav	6	1	1	2	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-4-singers-natural.wav	audios/4-4-singers-same-pitch.wav	4	2	3	2	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	2	1	2	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-natural.wav	4	2	2	3	2
audios/4-3-singers-natural.wav	audios/4-3-singers-different-pitch.wav	3	1	2	3	2
audios/2-1-singers-same-pitch.wav	audios/2-1-singers-different-pitch.wav	1	2	1	1	1
audios/3-6-singers-different-pitch.wav	audios/3-6-singers-same-pitch.wav	6	1	1	1	2
audios/2-3-singers-same-pitch.wav	audios/2-3-singers-natural.wav	3	2	3	3	2
audios/2-2-singers-natural.wav	audios/2-2-singers-different-pitch.wav	2	1	2	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-same-pitch.wav	5	1	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-same-pitch.wav	2	1	3	3	2
audios/4-4-singers-natural.wav	audios/4-4-singers-same-pitch.wav	4	1	3	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2
audios/3-5-singers-same-pitch.wav	audios/3-5-singers-natural.wav	5	2	3	3	2
audios/1-1-singers-different-pitch.wav	audios/1-1-singers-natural.wav	1	1	2	1	1
audios/4-6-singers-natural.wav	audios/4-6-singers-different-pitch.wav	6	1	2	3	2
audios/4-4-singers-different-pitch.wav	audios/4-4-singers-natural.wav	4	2	2	3	2
audios/4-2-singers-different-pitch.wav	audios/4-2-singers-same-pitch.wav	2	1	1	1	2
audios/2-4-singers-different-pitch.wav	audios/2-4-singers-same-pitch.wav	4	1	1	1	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/3-1-singers-natural.wav	audios/3-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-3-singers-different-pitch.wav	audios/1-3-singers-natural.wav	3	2	2	3	2
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	1
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	1
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	2	2	1	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	1	2	1	1
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	2	1	2	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	2	3	2	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	2	2	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	1
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	2	1	2	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	2	1	2	1
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	1	2	1	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	1
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	1
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	2	2	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	1

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	2	2	1	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	1	2	1	1
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	1	2	1	1
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	1
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	2
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	1
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	1	2	1	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	2	3	2	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	1
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	1
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	1
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	1	2	3	1
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	2	2	1	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	2	1	2	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	2	1	2	1
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-5-singers-same-pitch.wav	audios/1-5-singers-natural.wav	5	2	3	3	2
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	1
audios/2-2-singers-different-pitch.wav	audios/2-2-singers-same-pitch.wav	2	1	1	1	2
audios/4-5-singers-same-pitch.wav	audios/4-5-singers-different-pitch.wav	5	2	1	1	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/1-4-singers-natural.wav	audios/1-4-singers-same-pitch.wav	4	1	3	3	2
audios/2-1-singers-same-pitch.wav	audios/2-1-singers-natural.wav	1	1	3	2	1
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-natural.wav	3	2	3	3	2
audios/1-1-singers-different-pitch.wav	audios/1-1-singers-same-pitch.wav	1	1	1	1	1
audios/4-1-singers-different-pitch.wav	audios/4-1-singers-natural.wav	1	1	2	1	2
audios/3-6-singers-different-pitch.wav	audios/3-6-singers-same-pitch.wav	6	2	1	2	1
audios/3-6-singers-natural.wav	audios/3-6-singers-different-pitch.wav	6	1	2	3	2
audios/1-3-singers-different-pitch.wav	audios/1-3-singers-natural.wav	3	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-natural.wav	6	2	3	3	2
audios/3-2-singers-natural.wav	audios/3-2-singers-same-pitch.wav	2	1	3	3	2
audios/2-5-singers-different-pitch.wav	audios/2-5-singers-natural.wav	5	2	2	3	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-different-pitch.wav	4	2	1	1	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/1-4-singers-same-pitch.wav	audios/1-4-singers-natural.wav	4	2	3	3	2
audios/1-1-singers-different-pitch.wav	audios/1-1-singers-natural.wav	1	1	2	1	1
audios/2-3-singers-natural.wav	audios/2-3-singers-different-pitch.wav	3	1	2	3	1
audios/2-1-singers-different-pitch.wav	audios/2-1-singers-same-pitch.wav	1	1	1	1	1
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	2	3	2	1
audios/3-5-singers-same-pitch.wav	audios/3-5-singers-different-pitch.wav	5	2	1	1	1
audios/1-6-singers-different-pitch.wav	audios/1-6-singers-same-pitch.wav	6	1	1	1	2
audios/2-2-singers-different-pitch.wav	audios/2-2-singers-natural.wav	2	2	2	3	1
audios/4-5-singers-different-pitch.wav	audios/4-5-singers-natural.wav	5	2	2	3	1
audios/1-3-singers-different-pitch.wav	audios/1-3-singers-same-pitch.wav	3	1	1	1	1
audios/4-2-singers-different-pitch.wav	audios/4-2-singers-same-pitch.wav	2	1	1	1	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-natural.wav	4	2	2	3	1
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-6-singers-natural.wav	audios/3-6-singers-different-pitch.wav	6	1	2	3	2
audios/1-6-singers-natural.wav	audios/1-6-singers-same-pitch.wav	6	1	3	3	2
audios/2-3-singers-same-pitch.wav	audios/2-3-singers-natural.wav	3	2	3	3	2
audios/4-2-singers-same-pitch.wav	audios/4-2-singers-natural.wav	2	2	3	3	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	2	3	2	1
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	1	3	2	1

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	1	3	2	1
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	1	2	3	1
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	1
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	1	1	2	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	1	2	1	1
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	1
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	1
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	1
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	1
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	2
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	1
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	1
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	1
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	2
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	1
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	2
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	2	2	1	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	1	3	2	1
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	1
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	1	2	1	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	2	2	3	2
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	2	2	3	1
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-3-singers-natural.wav	audios/1-3-singers-different-pitch.wav	3	1	2	3	1
audios/2-1-singers-natural.wav	audios/2-1-singers-different-pitch.wav	1	1	2	3	1
audios/1-6-singers-same-pitch.wav	audios/1-6-singers-different-pitch.wav	6	2	1	1	2
audios/1-2-singers-different-pitch.wav	audios/1-2-singers-same-pitch.wav	2	1	1	1	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-natural.wav	1	2	3	3	1
audios/3-2-singers-same-pitch.wav	audios/3-2-singers-natural.wav	2	2	3	3	2
audios/1-5-singers-same-pitch.wav	audios/1-5-singers-different-pitch.wav	5	2	1	1	2
audios/1-5-singers-natural.wav	audios/1-5-singers-same-pitch.wav	5	1	3	3	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-different-pitch.wav	1	1	1	2	1
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-different-pitch.wav	3	2	1	1	1
audios/1-4-singers-natural.wav	audios/1-4-singers-same-pitch.wav	4	1	3	3	1
audios/1-4-singers-natural.wav	audios/1-4-singers-different-pitch.wav	4	1	2	3	1
audios/3-6-singers-natural.wav	audios/3-6-singers-different-pitch.wav	6	1	2	3	1
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-natural.wav	5	2	2	3	1
audios/3-3-singers-same-pitch.wav	audios/3-3-singers-natural.wav	3	2	3	3	2
audios/2-4-singers-same-pitch.wav	audios/2-4-singers-different-pitch.wav	4	2	1	1	2
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/4-3-singers-natural.wav	audios/4-3-singers-same-pitch.wav	3	1	3	3	2
audios/2-5-singers-same-pitch.wav	audios/2-5-singers-natural.wav	5	2	3	3	2
audios/3-1-singers-different-pitch.wav	audios/3-1-singers-natural.wav	1	2	2	3	1
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-natural.wav	2	2	3	3	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	1
audios/2-4-singers-natural.wav	audios/2-4-singers-different-pitch.wav	4	1	2	3	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-4-singers-same-pitch.wav	audios/3-4-singers-natural.wav	4	2	3	3	2
audios/4-6-singers-different-pitch.wav	audios/4-6-singers-natural.wav	6	1	2	1	1
audios/4-6-singers-same-pitch.wav	audios/4-6-singers-different-pitch.wav	6	2	1	1	2
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-same-pitch.wav	5	1	1	1	2
audios/1-5-singers-different-pitch.wav	audios/1-5-singers-natural.wav	5	1	2	1	1
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	1
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-natural.wav	3	2	2	3	1
audios/2-1-singers-natural.wav	audios/2-1-singers-same-pitch.wav	1	2	3	2	1
audios/1-3-singers-natural.wav	audios/1-3-singers-different-pitch.wav	3	1	2	3	2
audios/2-1-singers-natural.wav	audios/2-1-singers-different-pitch.wav	1	2	2	1	1
audios/1-6-singers-same-pitch.wav	audios/1-6-singers-different-pitch.wav	6	2	1	1	2
audios/1-2-singers-different-pitch.wav	audios/1-2-singers-same-pitch.wav	2	1	1	1	2
audios/4-2-singers-natural.wav	audios/4-2-singers-different-pitch.wav	2	2	2	1	2
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-natural.wav	1	2	3	3	1

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/3-2-singers-same-pitch.wav	audios/3-2-singers-natural.wav	2	2	3	3	2
audios/1-5-singers-same-pitch.wav	audios/1-5-singers-different-pitch.wav	5	2	1	1	1
audios/1-5-singers-natural.wav	audios/1-5-singers-same-pitch.wav	5	1	3	3	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-1-singers-same-pitch.wav	audios/4-1-singers-different-pitch.wav	1	2	1	1	1
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-different-pitch.wav	3	2	1	1	1
audios/1-4-singers-natural.wav	audios/1-4-singers-same-pitch.wav	4	1	3	3	2
audios/1-4-singers-natural.wav	audios/1-4-singers-different-pitch.wav	4	1	2	3	2
audios/3-6-singers-natural.wav	audios/3-6-singers-different-pitch.wav	6	1	2	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-natural.wav	5	2	2	3	2
audios/3-3-singers-same-pitch.wav	audios/3-3-singers-natural.wav	3	2	3	3	2
audios/2-4-singers-same-pitch.wav	audios/2-4-singers-different-pitch.wav	4	2	1	1	2
audios/1-3-singers-same-pitch.wav	audios/1-3-singers-natural.wav	3	2	3	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-same-pitch.wav	5	1	1	1	2
audios/2-6-singers-natural.wav	audios/2-6-singers-different-pitch.wav	6	1	2	3	2
audios/4-1-singers-natural.wav	audios/4-1-singers-different-pitch.wav	1	2	2	1	1
audios/3-5-singers-natural.wav	audios/3-5-singers-different-pitch.wav	5	1	2	3	2
audios/4-5-singers-same-pitch.wav	audios/4-5-singers-natural.wav	5	2	3	3	2
audios/4-3-singers-different-pitch.wav	audios/4-3-singers-same-pitch.wav	3	2	1	2	2
audios/2-2-singers-natural.wav	audios/2-2-singers-same-pitch.wav	2	1	3	3	1
audios/3-1-singers-same-pitch.wav	audios/3-1-singers-natural.wav	1	2	3	3	1
audios/1-1-singers-same-pitch.wav	audios/1-1-singers-different-pitch.wav	1	2	1	1	1

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/1-2-singers-same-pitch.wav	audios/1-2-singers-different-pitch.wav	2	2	1	1	2
audios/3-2-singers-natural.wav	audios/3-2-singers-different-pitch.wav	2	1	2	3	2
audios/3-6-singers-same-pitch.wav	audios/3-6-singers-different-pitch.wav	6	1	1	2	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	2	3	3	2
audios/4-4-singers-natural.wav	audios/4-4-singers-same-pitch.wav	4	1	3	3	2
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-same-pitch.wav	4	1	1	1	1
audios/3-4-singers-different-pitch.wav	audios/3-4-singers-natural.wav	4	2	2	3	2
audios/4-3-singers-natural.wav	audios/4-3-singers-different-pitch.wav	3	1	2	3	2
audios/2-1-singers-same-pitch.wav	audios/2-1-singers-different-pitch.wav	1	1	1	2	1
audios/3-6-singers-different-pitch.wav	audios/3-6-singers-same-pitch.wav	6	1	1	1	2
audios/2-3-singers-same-pitch.wav	audios/2-3-singers-natural.wav	3	2	3	3	2
audios/2-2-singers-natural.wav	audios/2-2-singers-different-pitch.wav	2	2	2	1	1
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-same-pitch.wav	5	1	1	1	2
audios/2-6-singers-same-pitch.wav	audios/2-6-singers-natural.wav	6	1	3	2	1
audios/4-2-singers-natural.wav	audios/4-2-singers-same-pitch.wav	2	1	3	3	2
audios/4-4-singers-natural.wav	audios/4-4-singers-same-pitch.wav	4	1	3	3	2
audios/3-5-singers-different-pitch.wav	audios/3-5-singers-natural.wav	5	2	2	3	2
audios/2-3-singers-different-pitch.wav	audios/2-3-singers-same-pitch.wav	3	1	1	1	2
audios/3-5-singers-same-pitch.wav	audios/3-5-singers-natural.wav	5	2	3	3	2
audios/1-1-singers-different-pitch.wav	audios/1-1-singers-natural.wav	1	1	2	1	1
audios/4-6-singers-natural.wav	audios/4-6-singers-different-pitch.wav	6	1	2	3	1
audios/4-4-singers-different-pitch.wav	audios/4-4-singers-natural.wav	4	2	2	3	2

Audio1	Audio2	Singers	Answer	Mode	Type	Confidence
audios/4-2-singers-different-pitch.wav	audios/4-2-singers-same-pitch.wav	2	1	1	1	2
audios/2-4-singers-different-pitch.wav	audios/2-4-singers-same-pitch.wav	4	1	1	1	2
audios/3-1-singers-natural.wav	audios/3-1-singers-same-pitch.wav	1	1	3	3	1
audios/1-3-singers-different-pitch.wav	audios/1-3-singers-natural.wav	3	2	2	3	2

Apéndice C

Análisis de costes

Recursos utilizados:

- Ordenador portátil con Sistema Operativo “Debian”. 4GB de memoria RAM. Procesador “Intel i5”.
- Auriculares “Tascam TH-02”. Proporcionados por la ETSI de Sistemas de Telecomunicación.
- Software “MATLAB R2019b”. Licencia educativa proporcionada por la Universidad Politécnica de Madrid.
- Software “TexMaker”. Editor de lenguaje LaTeX de código abierto.
- Software “LibreOffice Calc”. Programa de hojas de cálculo de código abierto.
- Servicio de suscripción a bases de datos bibliográficas y de normativas. Suscripción proporcionada por la Universidad Politécnica de Madrid.

Horas de trabajo:

El estudio comenzó a realizarse a principios de septiembre de 2019 y ha requerido un trabajo semi-constante hasta finales de septiembre de 2020. La mayoría de las horas de trabajo se destinó al proceso de búsqueda, análisis y clasificación de los diferentes documentos necesarios para la realización del protocolo. La situación de pandemia global producida por la COVID-19 ha producido retrasos y cambios en la metodología de trabajo, lo que se ha traducido en un incremento de las horas dedicadas.

La toma de datos por parte de los participantes se realizó en un total de 32 horas repartidos en 5 días consecutivos durante el mes de septiembre de 2020. El análisis de los datos se realizó en el periodo de una semana. Finalmente, se ha necesitado alrededor de 2 semanas para revisar los contenidos y corregir erratas o pequeños fallos.

Costes del proyecto

En la tabla C.1 se han representado los costes relacionados con la realización del proyecto. Los valores reflejados para las horas de trabajo y la suscripción de artículos bibliográficos son estimaciones de la cuantía de horas dedicadas al proyecto y el número de artículos consultados. El valor por unidad también está obtenido por el precio medio que suele encontrarse para la contratación de esos servicios. Todos los costes tienen incluidos sus impuestos asociados.

Elemento	Coste/Ud [€]	Unidades
Portátil, 4GB RAM, intel i5	750.00	1
Licencia S.O. Debian	0.00	1
Licencia software código abierto	0.00	3
Licencia software MATLAB	2000.00	1
Auriculares TASCAM TH-02	30.00	1
Suscripción a artículo o norma	27.00	30
Artículos de libre acceso	0.00	20
Hora de trabajo Ing. Junior	9.75	400
-	7490.00	Total

Tabla C.1: Costes asociados a los recursos utilizados en el proyecto

Bibliografía

- [1] Unión Internacional de Telecomunicaciones: *Recomendación UIT-R BS.1116-3: Métodos para la Evaluación Subjetiva de Pequeñas Degradaciones en los Sistemas de Audio*, UIT-R Sector de Radiocomunicaciones de la UIT, 2015.
- [2] Unión Internacional de Telecomunicaciones: *Recomendación UIT-R BS.1534-3: Método para la Evaluación Subjetiva del Nivel de Calidad Intermedia de los Sistemas de Audio*, UIT-R Sector de Radiocomunicaciones de la UIT, 2015.
- [3] European Broadcasting Union: *Assessment Methods for the Subjective Evaluation of the Quality of Sound Programme Material - Music*, Technical Report 3286-E, 1997.
- [4] S. Gelfand *Hearing: An Introduction to Psychological and Psysiological Acoustics*, Marcel Dekker, 2004.
- [5] Soren Bech and Nick Zacharov: *Perceptual Audio Evaluation-Theory, Method and Application*, John Wiley & sons, Ltd, 2006.
- [6] J. Higgins: *Introduction to Modern Nonparametric Statistics*, Duxbury Adadvanced Series, 2004.
- [7] E. Zwicker: *Psychoacoustics: Facts and Models*, Springer-Verlang, 1998.
- [8] G. Soulodre and J. Bradley: *Subjective Evaluation of New Room Acoustic Measures*, The Journal of te Acoustical Society of America Volume, 1995.
- [9] G. Pulvirenti, N. Totaro and E. Parizet: *Simulated Transfer Path Accuracy vs. Sound Perception*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [10] S. Klockgether and S. Van der Par: *Just Noitceable Differences of Spatial Cues in Echoic and Anechoic Acoustical Enviroments*, The Journal of the Acoustical Society of America Volume 140, 2016.
- [11] J. Blauert: *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, 1997.
- [12] D. de la Prida, A. Pedrero et al: *Does the Method Matter?: A Review of the Main Testing Methods for the Subjective Evaluation of Room Acoustics Through Listening Tests*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [13] D. Lu and X. Zeng: *3D Sound Effect Enhancement of Front And Back*, Proceedings of the Inter.Noise Congress, 2019.
- [14] S. Kuwano, H. Ohkawa and K. Inoue: *Differential Limen of Sounds Transmitted from Neighbours*, Proceedings of the Inter.Noise Congress, 2019.

- [15] L. Rinskaya-Korsakova and D. Nechaev: *The Detection of the First and Last Pulses in a Periodic Pulse Train: The Role of Temporal Integration and Masking*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [16] Unión Internacional de Telecomunicaciones: *Recomendación UIT-R BS.1283-1: Guía de las Recomendaciones UIT-R sobre la Evaluación Subjetiva de la Calidad de Sonido*, UIT-R Sector de Radiocomunicaciones de la UIT, 2003.
- [17] Unión Internacional de Telecomunicaciones: *Recomendación UIT-R BS.1285: Métodos de Preselección para la Evaluación Subjetiva de Pequeñas Degradaciones de Audio*, UIT-R Sector de Radiocomunicaciones de la UIT, 1997.
- [18] Unión Internacional de Telecomunicaciones: *Recomendación UIT-R BS.1286: Métodos para la Evaluación Subjetiva de los Sistemas de Audio con Acompañamiento de Imagen*, UIT-R Sector de Radiocomunicaciones de la UIT, 1997.
- [19] Unión Internacional de Telecomunicaciones: *Recomendación UIT-R BS.1284-2: Métodos Generales para la Evaluación Subjetiva de la Calidad de Sonido*, UIT-R Sector de Radiocomunicaciones de la UIT, 2019.
- [20] Asociación Española de Normalización: *UNE-EN ISO 10399:2018: Análisis Sensorial. Metodología. Ensayo dúo-trío. (ISO 10399:2017) (Ratificada por la Asociación Española de Normalización en marzo de 2018.)*, UNE Normalización Española, 2019.
- [21] T. Cox, W. Davies and Y.Lam: *The Sensitivity of Listeners to Early Sound Field Changes in Auditoria*, Acta Acustica United with Acustica Volume 79, 1993.
- [22] J.Egan, E. Carterette and J. Thwing: *Some Factors Affecting Multi-Channel Listening*, The Journal of the Acoustical Society of America Volume 26, 1954.
- [23] S. Carlile, P. Leong and S. Hyams: *The Nature and Distribution of Errors in Sound Localization by Human Listeners*, Elsevier, 1997.
- [24] M. Nowak and P. Kokowsky: *Subjective and Objective Assessments of Noise Barriers in Terms of the Loudness Level*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [25] V. Emiya, E. Vincent et al: *Subjective and Objective Quality Assessment of Audio Source Separation*, IEEE Transactions on Audio, Speech and Language Processing Volume 19, 2011.
- [26] B. Postma and B. Katz: *Perceptive and Objective Evaluation of Calibrated Room Acoustic Simulation Auralizations*, The Journal of the Acoustical Society of America Volume 140, 2016.
- [27] A. Lindau, V. Erbes et al: *A Spatial Audio Quality Inventory (SAQI)*, Acta Acustica United with Acustica Volume 100, 2014.
- [28] L. Kritly, V. Chmelk et al: *Audibility of Spectral Dips and Peaks in Broadband Noise*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [29] J.Bradley, R. Reich and S. Norcross: *A Just Noticeable Difference in C50 for Speech*, Elsevier, 1999.
- [30] P. Zahorik: *Direct-to-Reverberant Energy Ratio Sensitivity*, The Journal of the Acoustical Society of America Volume 112, 2010.

-
- [31] I. Witew, G. Behler and M. Vorländer: *About Just Noticeable Differences for Aspects of Spatial Impressions in Concert Halls*, Acoustical Science & Technology Volume 26, 2005.
- [32] F. Martellotta: *The Just Noticeable Difference of Center Time and Clarity Index in Large Reverberant Spaces*, The Journal of the Acoustical Society of America Volume 128, 2010.
- [33] M. Vigeant and R. Celmer: *Effect of Experimental Design on the Results of Clarity-Index Just Noticeable-Difference Listening Tests*, Proceedings of 20th International COngress of Acoustics (ICA), 2010.
- [34] Z. Shao, W. Zhang and W. Zhu: *Research on the Bandwidth Effect on Binaural Loudness Summation on the Cam Scale*, Proceedings of the Inter.Noise Congress, 2019.
- [35] V. Hongisto and P. Virjonen: *Annoyance Penalty of Amplitude-Modulated Sound*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [36] M. Shiell and E. Formisano: *Acuity of Spatial Stream Segregation Along the Horizontal Azimuth with Non-Individualized Head-Related Transfer Functions*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [37] J. Groose and S. Van der Par: *Improved Binaural Speech Inteligibility by Adding Reverberation to the Target Speaker*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [38] M. Yamada, F. Saze et al: *Feature Analysis of Sound Direction Perception Using Frequency Band-Limited Stimuly: Extension of a Directional Band Model*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [39] J. Lee, S. Chang et al: *Annoyance Modeling of Construction Noise Using Acoustical Features, Noise Sensitivity and Health Condition*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [40] D. Morikawa, D. Kojima and T. Hirahara: *Spatial Sound Segregation in the Monaural Listening Condition*, Proceedings of the 23rd Congress on Acoustics (ICA), 2019.
- [41] J. Schilittenlacher, R. Zhao and C. Moore: *Loudness of Ramped and Damped Sounds that are Temporally Shifted Across Ears*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [42] V. Rajala and V. Hongisto: *Annoyance of Impulsive Sounds - A Psychoacoustic Experiment Involving Synthetic Sounds*, Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.
- [43] S. Bech and N. Zacharov: *Perceptual Audio Evaluation-Theory, Method and Application*, John Wiley & Sons, Ltd, 2006.
- [44] G. Barrios: *Synthesizer of the Choir Effect Applying PSOLA*, Universidad Politécnica de Madrid, 2018.
- [45] B. Walker and G. Kramer: *Ecological Psychoacoustics and Auditory Displays: Hearing, Grouping and Meaning Making*, 1997.