

Capítulo 1

Antecedentes

En este capítulo se pretende realizar un análisis del estado del arte referente a los experimentos psicoacústicos; concretamente aquellos que se centran en la evaluación de diferencias perceptuales en entornos cerrados como auditorios, cines, salas de conferencias, etc.

Para la búsqueda de esta información, se recurre a bases de datos científicas como el sistema *Ingenio UPM* y buscadores de artículos científicos como *Google Scholar*. Con ellos, se han introducido términos de búsqueda como “*Psychoacoustics*”, “*Subjective test acoustics*” o “*Auditorium subjective test*”. Los resultados arrojados por dichas plataformas, se revisan y se lee el resumen o *abstract* de los que se consideran que pueden ser útiles para el proyecto. Si el resumen confirma que la temática concuerda con la del estudio, se procede a la lectura del resto del artículo y se extraen los elementos más importantes. Estos son principalmente: el tipo de análisis estadístico que se ha utilizado, el número de personas que han realizado el test, el tipo de test (tipo de preguntas, formato, duración de los audios utilizados, etc.), entre otros.

Por otro lado, se ha realizado, paralelamente, una búsqueda de fuentes bibliográficas especializados en temas como de detección de señales, análisis estadístico de datos y, más específicamente, sobre modelos Thurstonianos. Para ello, se realizan nuevas consultas a los medios ya presentados a personas con experiencia contrastada en la realización y el estudio de experimentos psicoacústicos.

Utilizando esta metodología, se han consultado alrededor de 50 documentos entre libros de referencia, normativas internacionales, artículos en revistas y congresos, trabajos de fin de grado, etc. De todo este material, han acabado siendo de utilidad 30 fuentes bibliográficas aproximadamente.

1.1. Análisis de la documentación consultada

Al analizar los distintos documentos, se ha observado una gran inconsistencia en las características de los diferentes estudios que se han analizado. Este problema ya se comenta en algunas fuentes como [1]. En este apartado se analizan las características de los proyectos y normativas según varios de sus aspectos más representativos.

1.1.1. Aspectos subjetivos a analizar

Los aspectos subjetivos que se pueden analizar mediante test perceptuales son muy variados. Por un lado algunos experimentos como en [2–5] se centran en estudiar las características relacionadas con la reverberación de las salas (persistencia del sonido, diferencias entre sonidos anecoicos y reverberantes, etc.). Otros estudios se centran en el estudio de la inteligibilidad de diferentes señales acústicas [6–8], mientras otras se centran en temas de localización de sonora [9–12]. También se encuentran aquellos experimentos en los que se pretende estudiar características como las diferencias entre señales, la sonoridad o diferencias de nivel [3, 13–18]. Por último, una de las áreas que está teniendo más importancia en los últimos años se corresponden con el estudio de la molestia acústica [19–21].

Algunos documentos como [1] tratan de ordenar estas fuentes en diferentes categorías y facilitar el análisis de grandes cantidades de bibliografía y extraer sus elementos comunes.

1.1.2. Tipos de test

Por otro lado, se ha observado una gran variedad de formatos para hacer los test. Los más habituales se presentan a continuación.

- **Test de diferencias:** En este tipo de test al oyente se le presentan 2 o más audios y tiene que determinar si son iguales o no. Tiene la ventaja de que son sencillos de implementar y de analizar. En algunos sitios, se le llama “duo” o “binomial” por lo que es importante revisarlos a conciencia para no confundirlos con los “Duo-Trio” o un “ABX”. Algunos de los artículos en los que han aplicado este tipo de test son [2, 5, 6, 8, 16, 17]. Un ejemplo concreto de este tipo de test y que tendrá gran importancia a lo largo del proyecto es el denominado “2-AFC” (*Two-Alternative Forced Choice*). En este test, el participante está “forzado” a escoger entre dos opciones, no pudiendo seleccionarse una opción intermedia, como respuesta a una pregunta relacionada con su exposición a un determinado estímulo [22].
- **Test Duo-Trio:** Para este formato, se presentan tres audios. Uno de ellos es igual a uno de los otros dos que se denomina “Referencia”. El objetivo del oyente es determinar cuál de los audios que se le presenta es igual a dicha referencia. Algunos de los estudios que han utilizado este tipo de test son [23, 24].
- **ABX:** En este tipo de test se presentan tres señales. El oyente tiene que determinar cuál de las dos primeras que se le presentan es igual a la que es presentada en último lugar. El experimento se repite varias veces modificando la última señal para que no sea siempre el mismo audio. Algunos ejemplos donde se aplican este tipo de test se encuentran en [24, 25].
- **Test de escalas numéricas:** El objetivo de este test es que el oyente cuantifique alguna determinada característica de las señales de audio mediante una escala numérica. Algunas variantes permiten que se realice la cuantificación de varios elementos en paralelo o que se pida que se ordenen varios audios simultáneamente en función de la característica que se pretenda evaluar (Test de MUSHRA [26]). Este tipo de test son muy útiles y aparecen como recomendación por parte de normativas internacionales [26–28]. También aparecen en algunos documentos como [4, 12, 13, 15, 20, 21, 29].

Existen otros tipos de test como los triangulares o los A/NotA [24, 30]. No obstante, no se han encontrado experimentos en los que se apliquen dichas metodologías, a parte de ese artículo, por lo que no se han considerado para las siguientes partes del proyecto.

Cada uno de los tipos de test mencionados arriba tienen sus propias particularidades, sus ventajas y sus inconvenientes tanto para su realización como para el análisis de sus resultados que se comentarán posteriormente en el capítulo de “Realización de los test subjetivos de audio”.

1.1.3. Características de los participantes

Si se analizan las características de las personas que participan en los test perceptuales, se observa una gran variedad tanto en el número como en su experiencia.

Por un lado, se tienen estudios que utilizan un número muy pequeño de personas, diez o menos, como es el caso de [3, 5, 14, 29]. Por otro lado, algunos casos presentan números mucho mayores llegando a utilizar casi cien personas, como ocurre en [31]. También se observa que el número de participantes varía en función del aspecto subjetivo que pretenda analizarse [1]; de esta forma, los estudios que estudian la molestia acústica son los que utilizan un mayor número de participantes (en torno a los 30-40 según [1]), mientras que el resto utilizan unas 20 personas aproximadamente, que se corresponde con las recomendaciones de las normativas [26–28].

Otro aspecto importante muy relacionado con el número de participantes consiste en la utilización de participantes con o sin experiencia en este tipo de experimentos psicoacústicos. Normativas como [26–28] recomiendan la utilización de participantes con experiencia cuando los aspectos a analizar son bastante técnicos o requieren de personas con oídos especializados en la detección de ciertos eventos (como, por ejemplo, educación musical avanzada). No obstante, aceptan la presencia de participantes

sin experiencia si se aumenta el número total de personas que participan en el experimento o si se les aplica algún tipo de entrenamiento previo, incluso se recomienda realizarlo para participantes con experiencia.

A pesar de estas recomendaciones, la realidad es bastante diversa y se encuentran estudios como [2, 4, 6–10, 15, 17, 21, 23–25, 29, 32] en los que sí se aplican entrenamientos previos, mientras que en otros estudios como [12, 16, 18, 30] no se realizan. También se han encontrado varios estudios en los que no se reflejaba si se habían realizado estos entrenamientos o no [3, 5, 11, 13, 14, 20, 33].

1.1.4. Interacción del participante con los estímulos auditivos

A nivel de interacción por parte de la persona participante, existe división de opiniones sobre la posibilidad del usuario de poder controlar la reproducción de los estímulos [2, 7, 8, 12, 15] frente a [3, 5, 6, 17, 19, 21] que no permiten a los participantes esa opción. No obstante, normas como [26–28, 34–36] recomiendan que los participantes tengan, siempre que sea posible, la capacidad de interactuar con los estímulos de forma directa y repetir su reproducción si así lo desean. En ningún momento se hace mención sobre el tiempo máximo que un usuario debe pasar con cada estímulo.

1.1.5. Duración del experimento

En cuanto a la duración de los estímulos, se han observado que los valores suelen rondar de 1 a 5 segundos [3, 5, 7, 10–16, 18, 32]. No obstante, también se han encontrado estudios donde la duración se encuentra entre los 8–10 segundos [7, 8, 20], obteniéndose casos con un máximo de 15 segundos como en [2, 17]. Esto concuerda con las recomendaciones de [26–28, 34–36] al respecto donde se indica que la duración de dichos estímulos debe ser lo suficientemente reducida para evitar que los participantes se acostumbren a las señales [37]. Unido a esto, la duración total de la sesión, a pesar de ser el apartado del que menos información se aporta, se recomienda que no exceda de los 30 minutos, siendo necesaria la inclusión de descansos de la misma duración, si se tiene que alargar [26–28, 37–39]. Esto ocurre en casos como [16] donde se tiene una duración de casi 90 minutos. En ningún momento se especifica si el entrenamiento previo computa dentro de la duración del experimento, aunque en nuestra experiencia particular, se suele incluir.

1.1.6. Análisis de los datos

En cuanto al análisis de los datos, históricamente se han utilizado procedimientos como los análisis de la varianza (ANOVA) [4, 10, 13, 15, 16, 19], o normas como [40]. No obstante, en estudios más recientes como [23, 24] aparecen los modelos thurstonianos como una alternativa sencilla de aplicar y con cualidades que pueden resultar útiles para ciertos tipos de estudio. Algunos de estos ejemplos se explican con más profundidad en el apartado “1.2.2 Modelos Thurstonianos”.

A la vista de todo lo expuesto, queda patente la gran variedad de sistemas que se siguen en la actualidad para realizar los test perceptuales. Por ello, para nuestro test es necesario identificar sus particularidades para poder determinar las características que mejor se ajusten al mismo. Esto se trabaja en profundidad en los apartados “4.1.1 Características del test previo” y “4.1.2 Características de los test finales” siguiendo la metodología del capítulo “Toma de datos”.

1.2. Conceptos teóricos para el análisis estadístico

1.2.1. UNE-EN ISO 10399

La norma UNE-EN ISO 10399: “Análisis sensorial. Metodología. Ensayo Duo-Trio.” [40] es un documento que permite analizar la probabilidad de que eventos perceptuales sean percibidos como iguales o diferentes según el número de respuestas definidas como correctas o erróneas al realizar experimentos basados en test duo-trio. Esta norma no se centra en señales acústicas, sino que puede utilizarse para experimentos subjetivos relacionados con olores, sabores, imágenes, etc.

Al comienzo del documento se definen diferentes términos que aparecen de forma constante a lo largo de la norma. Para nuestro caso particular, los términos más relevantes son:

- alpha-risk o α -risk: es la probabilidad, o riesgo, de afirmar que existe una diferencia perceptual cuando en realidad no existe.
- beta-risk o β -risk: es la probabilidad, o riesgos, de afirmar que no existe una diferencia perceptual cuando en realidad sí existe.
- diferencia perceptual: situación en la que dos o más muestras pueden ser distinguidas por sus propiedades sensitivas (a través del oído, tacto, gusto, vista, etc.)
- similaridad perceptual: situación en la que las diferencias entre muestras son tan pequeñas que no pueden distinguirse entre sí de forma sensitiva.

Para el cálculo de las probabilidades α -risk y β -risk, la norma proporciona sendas tablas. En este proyecto sólo se precisa el cálculo de α que se encuentran en el Anexo B (la tabla B.1). También se proporciona la ecuación 1.1 donde se puede calcular el mínimo de respuestas “correctas” para que se obtenga un determinado valor de α -risk. Con estas herramientas se puede comprobar si los resultados de un test arrojan valores para α y β se encuentran dentro de unos determinados valores (20 %, 10 %, 5 %, 1 % y 0.1 %)

$$x = (n/2) + z * \sqrt{n/4} \quad (1.1)$$

Donde:

- x: número de respuestas correctas mínimas necesarias para que se obtenga un determinado α -risk.
- n: número de respuestas totales.
- z: variable que toma un valor en función de α -risk:
 - 0.84 para $\alpha = 0,20$.
 - 1.28 para $\alpha = 0,10$.
 - 1.64 para $\alpha = 0,05$.
 - 2.33 para $\alpha = 0,01$.
 - 3.09 para $\alpha = 0,001$.

Un ejemplo de aplicación para conocer el coeficiente α -risk sería el siguiente: Suponemos que tenemos un caso de estudio entre dos señales acústicas distintas entre sí. Se ha realizado un test perceptual y se han obtenido 100 respuestas, 87 de las cuales opinan que las dos señales son diferentes.

En primer lugar, se acude a la tabla del Anexo B para comprobar si están los datos para ese número de respuestas. En este caso, se comprueba que, desgraciadamente, la tabla sólo arroja datos hasta las 88 respuestas, por lo que se procede a resolver la ecuación 1.1. Sustituyendo las variables por los datos que se tienen, se obtienen 5 valores diferentes de x para cada uno de los posibles valores de α :

- $x = 55$ para $\alpha = 0,20$.
- $x = 57$ para $\alpha = 0,10$.
- $x = 59$ para $\alpha = 0,05$.
- $x = 62$ para $\alpha = 0,01$.
- $x = 66$ para $\alpha = 0,001$.

Es importante remarcar que los resultados siempre se redondean hacia arriba, puesto que es imposible tener fracciones de respuestas.

Una vez obtenidos los valores, el valor de α se corresponde con el del caso más cercano que sea inferior a nuestro número de respuestas diferentes. En nuestro caso, $\alpha < 0,001$ ya que 66 es el número más cercano a 87 (y además se cumple que es menor).

Como ya se explicó anteriormente, el coeficiente α -risk muestra la probabilidad de que dos eventos sean diferentes cuando, en realidad, son iguales. Por este motivo, cuanto más pequeño es el valor del coeficiente, estadísticamente más representativa es la diferencia perceptual. Por este motivo, en estudios como el nuestro, se busca obtener valores pequeños.

Puede ocurrir que se obtengan situaciones donde el número de respuestas en las que se han indicado que son diferentes sea menor que para el caso de $\alpha = 0,20$ (en el ejemplo anterior, si se hubiera obtenido menos de 55 respuestas marcadas como diferentes). Si este fuera el caso, no se podría concluir que estadísticamente la percepción de dichos estímulos es diferente. En todo caso, habría que concluir que las posibles diferencias son debidas, mayoritariamente, al azar.

1.2.2. Modelos Thurstonianos

Los *modelos Thurstonianos*¹ [22,23,41] son modelos estadísticos en los que se utilizan variables de distribuciones normales y que se utilizan en gran medida en estudios de discriminación sensorial.

En el caso concreto de la psicoacústica, se utilizan generalmente para obtener un valor “ d' ” que da información ordenada y cuantitativa sobre una determinada percepción subjetiva. Además de este valor, se calcula a su vez la desviación estandar “ σ ”. Con estos dos valores, se puede aproximar las respuestas de un test subjetivo como una sucesión de distribuciones gaussianas en las que en función del valor “ d' ” y “ σ ” están más o menos superpuestas. Esta superposición da información sobre la probabilidad de que ambos estímulos puedan ser distinguibles o no entre sí y cuánto. Esto se observa más fácilmente en la figura 1.1 obtenida en [22].

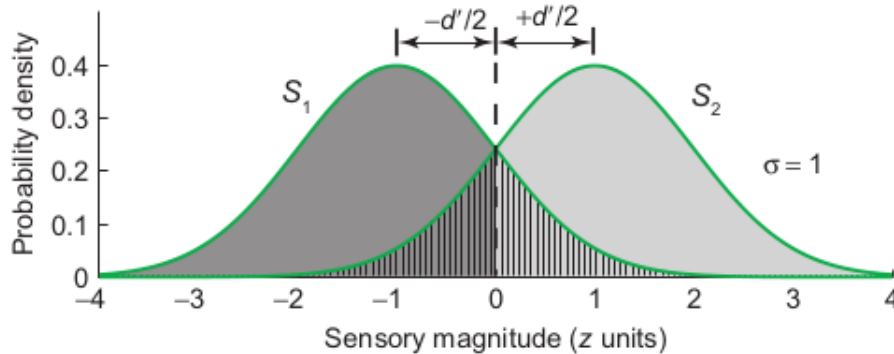


Figura 1.1: Ejemplo de cálculo de d' utilizando modelos thurstonianos. Fte: [22]

Utilizando la figura 1.1 como ejemplo, se puede suponer que S_1 y S_2 son las distribuciones obtenidas de las respuestas de un test en las que S_1 son las respuestas en las que dos estímulos han sido identificados como iguales, mientras que S_2 se corresponde con las respuestas cuando los dos estímulos se identifican como diferentes. Cuanto más se solapen ambas distribuciones, más difícil será para los participantes distinguirlas. Gráficamente resulta fácil comprobar que las variables que determinan este solapamiento es principalmente d' , sin dejar de lado la importancia que tiene la desviación estándar σ .

Además de lo ya expuesto, este tipo de análisis es especialmente interesante porque nos permite ordenar los valores obtenidos para d' de forma que las diferencias entre ellos nos da información cuantitativa sobre cómo de diferentes o de similares son cada una de las distribuciones. Con la facilidad añadida que tiene este sistema para representar gráficamente mediante las técnicas habituales

¹Nombrados así por Louis Leon Thurstone.

como diagramas de barras, entre otros. También tiene la ventaja de que el valor del coeficiente d' es independiente del tipo de test que se realice, incluso es posible compararlos entre sí [24].

Bibliografía

- [1] Tejada, V. de: *Desarrollo de un protocolo para la realización de experimentos de percepción subjetiva*. Universidad Politécnica de Madrid, 2020. <https://oa.upm.es/66850>.
- [2] Soulodre, G. y J. Bradley: *Subjective Evaluation of New Room Acoustic Measures*. Journal of the Acoustical Society of America, 1995.
- [3] Klockgether, S. y S. van der Par: *Just Noticeable Differences of spatial cues in echoic and anechoic acoustical environments*. Journal of the Acoustical Society of America, 2016.
- [4] Postma, B. y B. Katz: *Perceptive and Objective evaluation of Calibrated Room Acoustic Simulation Auralizations*. Journal of the Acoustical Society of America, 2016.
- [5] Zahorik, P.: *Direct-to-Reverberant Energy Ratio Sensitivity*. Journal of the Acoustical Society of America, 2002.
- [6] Bradley, J., R. Reich y S. Norcross: *A Just Noticeable Difference in C50 for Speech*. Elsevier, 1999.
- [7] Martellotta, F.: *The Just Noticeable Difference of Center Time and Clarity Index in Large Reverberant Spaces*. Journal of the Acoustical Society of America, 2010.
- [8] Vigeant, M. y R. Celmer: *Effect of Experimental Design on the Results of Clarity-Index Just-Noticeable-Difference Listening Tests*. En *Proceedings of 20th International Congress of Acoustics, ICA 2010*, 2010.
- [9] Carlile, S., P. Leong y S. Hyams: *The nature of distribution errors in sound localization by human listeners*. Elsevier, 1997.
- [10] Shiell, M. y E. Formisano: *Acuity of Spatial Stream Segregation Along the Horizontal Azimuth with Non-Individualized Head-Related Transfer Functions*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [11] Yamada, M., F. Saze, T. Horiuchi y K. Okubo: *Feature Analysis of Sound Direction Perception Using Frequency Band-Limited Stimuli: Extension of a Directional Band Model*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [12] Morikawa, D., D. Kojima y T. Hirahara: *Spatial Sound Segregation in the Monaural Listening Condition*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [13] Pulvirenti, G., N. Totaro y E. Parizet: *Simulated Transfer Path Accuracy vs Sound Perception*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [14] Nowak, M. y P. Kokowsky: *Subjective and Objective Assessments of Noise Barriers in Terms of the Loudness Level*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [15] Emiya, V., E. Vincent, N. Harlander y V. Hohmann: *Subjective and Objective Quality Assessment of Audio Source Separation*. IEEE Transactions on Audio, Speech and Language Processing, 2011.

- [16] Kritly, L., V. Chmelik, Y. Sluyts, C. Glorieux y M. Rychtáriková: *Audibility of Spectral Dips and Peaks in Broadband Noise*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [17] Witew, I., G. Behler y M. Vorländer: *About Just Noticeable Differences for Aspects of Spatial Impressions in Concert Halls*. Acoustical Science and Technology, 2005.
- [18] Schlittenlacher, J., R. Zhao y B. Moore: *Loudness of Ramped and Damped Sounds that are Temporally Shifted Across Ears*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [19] Hongisto, V. y P. Virjonen: *Annoyance Penalty of Amplitude-Modulated Sound*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [20] Lee, J., S. Chang y J. Jang: *Annoyance of Construction Noise Using Acoustical Features, Noise Sensitivity and Health Condition*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [21] Rajala, V. y V. Hongisto: *Annoyance of Impulsive Sounds - a Psychoacoustic Experiment Involving Synthetic Sounds*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [22] Kingdom, F. y N. Prins: *Psychophysics: a practical introduction*. Elsevier, 2010.
- [23] Prida, Daniel de la y Antonio Pedrero: *Methodology for the subjective evaluation of airborne sound insulation through 2-AC and Thurstonian models*. Applied Acoustics, 157, 2019.
- [24] Prida, Daniel de la y Antonio Pedrero: *Listening Tests in Room Acoustics: Comparison of overall differences regarding operational power*. Applied Acoustics, 182, 2021.
- [25] Braun, V., Rogeaux M. et al.: *Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water*. Food Quality and Preference, 2004.
- [26] UIT, Sector de Radiocomunicaciones de la: *Recomendación UIT-R BS.1534-3: Método para la Evaluación Subjetiva del Nivel de Calidad Intermedia de los Sistemas de Audio*. Informe técnico, Unión Internacional de Telecomunicaciones, 2015.
- [27] UIT, Sector de Radiocomunicaciones de la: *Recomendación UIT-R BS.1116-3: Métodos para la Evaluación Subjetiva de Pequeñas Degradaciones en los Sistemas de Audio*. Informe técnico, Unión Internacional de Telecomunicaciones, 2015.
- [28] UIT, Sector de Radiocomunicaciones de la: *Recomendación UIT-R BS.1284-2: Métodos Generales para la Evaluación Subjetiva de la Calidad de Sonido*. Informe técnico, Unión Internacional de Telecomunicaciones, 2019.
- [29] Shao, Z., W. Zhang y W. Zhu: *Research on the Bandwidth Effect on Binaural Loudness Summation on the Cam Scale*. En *Proceedings of the Iner.Noise Congress 2019*, 2019.
- [30] Brockhoff, P. y R. Christensen: *Thurstonian models for sensory discrimination tests as generalized linear models*. Food Quality and Preference, 21, 2010.
- [31] Egan, J., E. Carterette y E. Thwing: *Some factors affecting multi-channel listening*. The Journal of the Acoustical Society of America, 1954.
- [32] Groose, J. y S. van de Par: *Improved Binaural Speech Intelligibility by Adding Reverberation to the Target Speaker*. En *Proceedings of the 23rd International Congress of Acoustics, ICA 2019*, 2019.
- [33] Christensen, H. y P. Brockhoff: *Estimation and inference in the same-different test*. Food Quality and Preference, 20, 2009.

- [34] Union, European Broadcasting: *Assessment Methods for the Subjective Evaluation of the Quality of Sound Programme Material - Music*. Informe técnico, European Broadcasting Union, 1997.
- [35] UIT, Sector de Radiocomunicaciones de la: *Unión Internacional de Telecomunicaciones: Recomendación UIT-R BS.1285: Métodos de Preselección para la Evaluación Subjetiva de Pequeñas Degradaciones de Audio*. Informe técnico, Unión Internacional de Telecomunicaciones, 1997.
- [36] UIT, Sector de Radiocomunicaciones de la: *Unión Internacional de Telecomunicaciones: Recomendación UIT-R BS.1286: Métodos para la Evaluación Subjetiva de los Sistemas de Audio con Acompañamiento de Imagen*. Informe técnico, Unión Internacional de Telecomunicaciones, 1997.
- [37] Gelfand, Stanley A: *Hearing : an introduction to psychological and physiological acoustics*. Marcel Dekker, New York, 4th ed. edición, 2004, ISBN 978-0-8247-5652-9.
- [38] Zwicker, E: *Psychoacoustics : facts and models*. Springer series in information sciences 22. Springer-Verlag, Berlin, 1998, ISBN 3-540-65063-6.
- [39] Blauert, Jens: *Spatial hearing : the psychophysics of human sound localization*. MIT Press, Cambridge, Massachusetts, rev. ed. edición, 1997, ISBN 0-262-02413-6.
- [40] (ISO), International Standarization Organization: *Análisis sensorial. Metodología. Ensayo dúo-trío*. UNE-EN, 2018.
- [41] Mcmillan, N. y C. Douglas: *Detection theory: a user's guide*. 2005.