

Subjective Evaluation in Acoustics through Listening Tests: Difference Testing in Sound Insulation, Room Acoustics and Soundscapes

Tribunal:

Presidente: Rubén Fraile Muñoz, Universidad Politécnica de Madrid.

Secretaria: Elena Blanco Martín, Universidad Politécnica de Madrid.

Vocal: María Machimbarrena Gutiérrez, Universidad de Valladolid

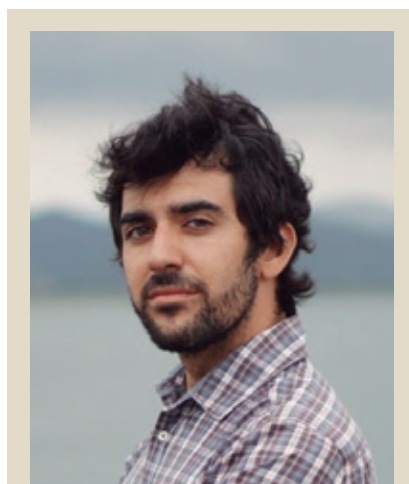
Vocal: Jaime Ramis Soriano, Universidad de Alicante

Vocal: Patrizio Fausti, Università degli Studi di Ferrara

Resumen:

En las diferentes áreas de la acústica se utilizan a día de hoy diversas métricas e indicadores para caracterizar el comportamiento sonoro de los entornos y su impacto sobre la vida de los seres humanos. Aunque los parámetros utilizados en las diversas áreas han sido descritos a lo largo de las décadas tratando de tener en cuenta la percepción humana del sonido, en los últimos años se está produciendo de manera más evidente un cambio paradigmático centrado en desarrollar mejores parámetros, así como en mejorar los parámetros ya existentes, con el fin de asegurar que estos sean todo lo representativos posible de la percepción subjetiva del sonido.

Para ello, se están llevado a cabo pruebas subjetivas en las que se consulta a muestras de la población acerca de su percepción subjetiva de diferentes estímulos sonoros. La finalidad de estas pruebas, que se pueden llevar a cabo en entornos reales (encuestas in situ) o en laboratorio (*listening tests*), es conocer, a partir de sus respuestas, cómo se comporta el sistema auditivo humano y



Autor:

Daniel de la Prida Caballero

Director/es:

**Antonio Pedrero González,
Luis Antonio Azpicueta Ruiz**

Fecha de defensa:

24 de junio de 2021

Lugar:

Universidad Politécnica de Madrid

e-mail:

d.delaprida@alumnos.upm.es

así valorar la adecuación de los indicadores actuales y realizar mejoras en caso necesario.

El principal obstáculo para que estas pruebas permitan alcanzar de manera satisfactoria este fin es la alta heterogeneidad de metodologías utilizadas para realizarlas, lo que lleva a que diversas investigaciones estén sujetas a diferentes fuentes de sesgo y métodos de análisis, dificultando la comparación de sus resultados. Este hecho resta relevancia a sus conclusiones y dificulta la realización de análisis conjuntos en base a los resultados de varias pruebas, lo que permitiría obtener mejoras de los in-

dicadores representativas de todas las situaciones acústicas valoradas por todas ellas.

Esta tesis doctoral tiene como principal objetivo la realización de contribuciones metodológicas que faciliten la definición de metodologías homogéneas, precisas y de mínimo sesgo, con herramientas de análisis altamente informativas para la realización de pruebas perceptuales en entornos de laboratorio (*listening tests*) en las áreas de aislamiento acústico, acústica de recintos y paisaje sonoro. Estas contribuciones metodológicas se identificaron tras el estudio pormenorizado de los tipos de *listening tests* utilizados a día de hoy en cada una de las áreas, evaluando sus ventajas, inconvenientes y principales fuentes de sesgo, así como estudiando en profundidad el estado del arte en otras áreas de la evaluación sensorial donde las pruebas perceptuales tienen un bagaje mayor y donde la calidad de las diferentes metodologías se encuentra, a día de hoy, más minuciosamente estudiada.

Así, se realizaron contribuciones metodológicas específicas para el diseño, realización y análisis de *listening tests* especialmente apropiadas para los fines más relevantes en cada una de las áreas. Estas contribuciones, basadas en métodos comparativos de testeo de las diferencias (*difference testing*), no solo se centraron en una homogeneización metodológica en cada una de las tres áreas, sino también en una homogeneización transversal, que pueda facilitar la descripción de un marco metodológico común basado

en *difference testing* que permita la utilización de métodos de análisis altamente informativos y apenas utilizados a día de hoy en acústica como son los modelos Thurstonianos.

Estas contribuciones fueron evaluadas mediante la realización de tres *listening tests*, uno en cada una de las áreas de estudio. Así, en el área de aislamiento acústico, se llevó a cabo un *listening test* siguiendo un protocolo 2-AC atributivo de testeo de las diferencias con una amplia muestra de 119 participantes, con la finalidad de evaluar la molestia percibida por estos ante diferentes tipos de ruido urbano bajo varias condiciones de aislamiento acústico de fachadas. Los resultados de esta prueba permitieron no sólo evaluar la representatividad de las magnitudes globales de aislamiento acústico (SNQs) actuales, sino también proponer una mejora de éstas mediante la aplicación de un algoritmo de optimización automático basado en Programación Cuadrática Secuencial (SQP), utilizando como parámetros de entrada los resultados de la prueba analizados mediante modelos Thurstonianos. La aplicación del algoritmo de optimización permitió obtener un espectro de referencia para el cálculo de un SNQ más representativo de sonidos urbanos ($L_{opt-urban}$) que cualquiera de los existentes a día de hoy.

Por otro lado, en el área de acústica de recintos se llevó a cabo un *listening test* con la finalidad de evaluar el poder operacional de siete protocolos de testeo de las diferencias generales, algunos de ellos comúnmente utilizados para realizar *listening tests* en esta área y otros potencialmente aplicables a acústica pero, hasta la fecha, habitualmente aplicados en otros campos de la evaluación sensorial. Este *listening test*, en el que una amplia muestra de 134 participantes evaluó mediante estos protocolos la dife-

rencia perceptual entre diversas realizaciones confundibles de un mismo recinto, permitió la comparación del poder operacional de los diversos protocolos mediante el uso de modelos Thurstonianos. Esta comparación permitió determinar y proponer como recomendado aquel protocolo para el que se obtuvo un mejor poder operacional y una menor influencia de las diversas fuentes de sesgo.

Por último, en el área de paisaje sonoro, se desarrollaron dos tareas noveles de *listening test* también basadas en testeo de las diferencias para la determinación de la similitud perceptual entre distintos paisajes sonoros. La utilidad de estas tareas se comprobó mediante la realización de un *listening test* sobre una pequeña población de 17 participantes que comparó paisajes sonoros urbanos con diferentes características arquitectónicas. Estos paisajes sonoros también fueron clasificados objetivamente mediante algoritmos de clasificación automáticos (*K-means* y clasificación jerárquica) en base a percentiles de los indicadores psicoacústicos *Loudness* ($N5$, $N50$ y $N95$), *Sharpness* ($S5$, $S50$ y $S95$), *Roughness* ($R10$, $R50$ y $R95$) y *Fluctuation Strength* ($FS10$, $FS50$ y $FS95$). Posteriormente, los resultados de esta clasificación se compararon con los obtenidos mediante el *listening test*, encontrándose una buena correlación, en la mayoría de las ocasiones, y poniéndose de relieve no solo la aplicabilidad de las tareas de *listening test* desarrolladas sino también la utilidad de los indicadores psicoacústicos para valorar la similitud perceptual de paisajes sonoros urbanos con características arquitectónicas similares.

Abstract:

In the different areas of acoustics, various measures and indicators are used nowadays to assess the acoustic behaviour of environments and

their impact on the lives of human beings. Although the metrics used in the different areas have been described over the decades attempting to take into account the human perception of sound, in recent years a paradigm shift is becoming more evident, focusing on developing better parameters, as well as improving existing ones, in order to ensure that they are as representative as possible of the subjective perception of sound.

To this end, subjective tests are being carried out in which samples of the population are consulted about their subjective perception of different sound stimuli. The purpose of these tests, which can be carried out in real environments (in situ surveys) or in a laboratory (listening tests), is to find out from their responses how the human auditory system behaves in order to assess the adequacy of current indicators and to make improvements when necessary.

The main obstacle for these tests to satisfactorily achieve this goal is the high heterogeneity of methodologies used to perform them, which leads to different investigations being subject to different sources of bias and analysis methods, hindering the comparison of their results. This fact reduces the relevance of their conclusions and makes it difficult to carry out meta-analyses based on the results of different tests, which would enable to obtain improvements in the indicators, making them representative of all the acoustic situations assessed by the different perceptual tests.

The main objective of this thesis is to provide methodological contributions that facilitate the definition of homogeneous, accurate and minimally biased methodologies with highly informative analysis tools for the performance of perceptual tests in laboratory environments (listening tests) in the areas of sound insulation, room acoustics and sound-

scapes. These methodological contributions were identified after a detailed study of the types of listening tests currently employed in each of these areas, evaluating their advantages, disadvantages and main sources of bias, as well as studying in depth the state of the art in other areas of sensory evaluation where these types of tests have a greater background and where the quality of the different methodologies has, to date, been more thoroughly studied.

Accordingly, specific methodological contributions were described for the design, performance and analysis of listening tests particularly appropriate for the most relevant purposes in each of the areas. These contributions, based on comparative difference testing methods, not only focused on a methodological homogenization in each of the three areas, but also on a transversal homogenization, which could facilitate the description of a common methodological framework based on difference testing that would allow the use of highly informative analysis methods barely employed so far in acoustics such as the Thurstonian models.

These contributions were evaluated through the performance of three listening tests, one in each of the study areas. In the area of sound insulation, a listening test following an attribute-related 2-AC difference testing protocol was carried out on

a large sample of 119 participants with the aim of assessing the annoyance perceived by them in the presence of different types of urban noise under different façade sound insulation conditions. The results of this test allowed not only to evaluate the representativeness of the current Single-Number Quantities of sound insulation (SNQs), but also to propose an improvement of these through the application of an automatic optimization algorithm based on Sequential Quadratic Programming (SQP), using as input parameters the results of the test analyzed by means of Thurstonian models. The application of the optimization algorithm enabled to obtain a reference spectrum for the calculation of a more representative SNQ of urban sounds ($L_{opt-urban}$) than any of the existing ones.

On the other hand, in the area of room acoustics, a listening test was carried out in order to evaluate the operational power of seven overall difference testing protocols, some of them commonly used for listening tests in this area and others potentially applicable to room acoustics but, to date, usually applied in other fields of sensory evaluation. This listening test, in which a large sample of 134 participants evaluated, using these protocols, the perceptual difference between different confusable auralizations of the same room, allowed a comparison of the opera-

tional power of the various protocols using Thurstonian models. This comparison allowed to identify and recommend the protocol with the best operational power and the least influence of the various sources of bias.

Finally, in the area of soundscapes, two novel listening test tasks were developed, also based on difference testing, for the determination of perceptual similarity between different soundscapes. The usefulness of these tasks was tested by means of a listening test conducted on a small population of 17 participants who compared urban soundscapes with different architectural features. These soundscapes were also objectively classified using automatic clustering algorithms (i.e., K-means and hierarchical clustering) based on percentiles of the psychoacoustic indicators Loudness ($N5$, $N50$ and $N95$), Sharpness ($S5$, $S50$ and $S95$), Roughness ($R10$, $R50$ and $R95$) and Fluctuation Strength ($FS10$, $FS50$ and $FS95$). Subsequently, the results of this clustering were compared with those obtained by the listening test, finding a good correlation, in most cases, and highlighting not only the applicability of the listening test tasks developed but also the usefulness of psychoacoustic indicators to assess the perceptual similarity of urban soundscapes with similar architectural characteristics.