

LOS LISTENING TEST COMO MÉTODO DE VALORACIÓN SUBJETIVA E IMPLEMENTACIÓN DE SOFTWARE DE EVALUACIÓN.

PACS 43.55.Rg

De la Prida, Daniel; Rodrigues A. Monteiro, Carolina; Machimbarrena, María
Universidad de Valladolid. Escuela de Arquitectura. Departamento de Física Aplicada
Av. Salamanca s/n 47014
Valladolid
España
Email: danidelaprida@gmail.com

ABSTRACT

Listening test have proved to be effective as a method to rate the subjective perception, although their preparation and performance can be tedious and there are many possible sources of bias. These biases are mainly caused by the human interaction of both the researcher and the participant. The aim of this research is the development of matlab-assisted applications to ease the preparation, tests performance and the analysis of the results, reducing as well the bias arising from human interaction. A test methodology is developed and real tests are performed in order to check its effectiveness.

RESUMEN

Los listening test han probado su eficacia como método de evaluación de la percepción subjetiva. Aunque eficaces, pueden ser tediosos en su realización y numerosas las fuentes de error, principalmente por la interacción humana tanto del participante como del examinador. El objetivo de esta investigación es el desarrollo de aplicaciones informáticas en entorno matlab, que faciliten las tareas de preparación y realización de los test y el análisis de los resultados, minimizando además los errores derivados de la interacción humana. Se define una metodología de ensayo y se realizan ensayos reales para comprobar su eficacia.

INTRODUCCIÓN

La presente comunicación describe los fundamentos básicos relativos a los listening test y las motivaciones que llevan a la implementación de aplicaciones informáticas para automatizar su proceso. Se presenta una biblioteca de aplicaciones informáticas desarrollada a tal fin, destacando sus aspectos más importantes. Por último, se presenta un estudio piloto, enmarcado en el ámbito del aislamiento acústico, con el fin de comprobar su efectividad y que permite extraer interesantes conclusiones con respecto al descriptor de aislamiento acústico R_{living} , equivalente a $R_w+C_{50-5000}$.

LOS LISTENING TEST Y LA VALORACIÓN SUBJETIVA

Los test de percepción o test sensoriales han sido utilizados de manera habitual en campos como el de la alimentación, con el fin de determinar las diferencias que perciben los seres humanos ante variaciones en una determinada característica de los productos, aportando ventajas competitivas a aquellos que son percibidos como mejores con respecto a esa característica [1].

En el campo de la acústica, los test de percepción son una herramienta joven y valiosa pero que precisa aún de abundante investigación y desarrollo, con el fin de obtener metodologías de uso estandarizadas. En este campo los test de percepción toman el nombre de listening test.

¿Qué es un listening test?

Un listening test es un test de percepción en el que el órgano sensorial involucrado es el oído.

Se presentan varios estímulos sonoros a un participante y se le pide, mediante un método de consulta determinado, que aporte su juicio con respecto al grado de adecuación que presentan éstos, en relación a una determinada característica acústica. Las características estudiadas más comúnmente son la sonoridad y la molestia.

Aunque existen, en los test sensoriales, numerosos métodos de consulta de los usuarios, en el caso de los listening test se utilizan principalmente dos métodos [1–4]:

1. Método de comparación por pares: Se presentan al participante, de manera automática o manual, dos estímulos sonoros y se le pide que seleccione aquel que represente mejor la característica consultada. Un hipotético caso podría ser aquel en el que se presentan al participante dos estímulos sonoros con distinto nivel y se le pregunta por cuál de los dos estímulos es más sonoro. El participante debería seleccionar aquel que él perciba con más nivel.
2. Método de escalado: Se presenta al participante un estímulo sonoro y se le pide que evalúe la relación que tiene ese estímulo con una determinada característica, mediante la selección de un valor en una escala. Un caso hipotético podría ser aquel en el que se presenta al usuario un estímulo muy estridente y se le pide que valore, en una escala de “nada molesto” a “extremadamente molesto”, la molestia que le causa escuchar el estímulo presentado. Este método suele ser muy utilizado para categorizar la molestia.

Tras realizar un listening test, se pueden extraer los resultados aportados por los participantes, ante distintos estímulos, y obtener conclusiones al respecto. Sin embargo, para que estas conclusiones sean estadísticamente significativas, el test se deberá realizar por un conjunto de participantes lo suficientemente grande y definido, como para que sea representativo de la población a la que está destinado.

¿Para qué sirve un listening test?

Los usos que pueden tomar los listening test son muy variados, pero éstos siempre están relacionados con la percepción subjetiva, que presentan los usuarios, ante determinados sucesos acústicos. Debido a la forma que tienen los seres humanos de abordar la categorización de ciertos eventos físicos, existen algunos usos predominantes de los listening test, cuyas características y motivación serán expuestas a continuación.

Los sucesos o eventos físicos suelen categorizarse mediante descriptores objetivos. Estos descriptores pueden estar o no relacionados con la interacción que estos sucesos físicos tienen con los seres humanos. Cuando estos descriptores objetivos tratan además de categorizar la percepción subjetiva que presentan los seres humanos ante un suceso, el descriptor ha de ser altamente representativo de la percepción subjetiva. Dado que la comprensión de algo tan etéreo puede ser complicada, se presenta un ejemplo que, aunque no está relacionado con la acústica, facilita el entendimiento.

La temperatura y la “sensación térmica” poseen descriptores objetivos relacionados con un determinado suceso físico. Sin embargo, mientras que la temperatura no está relacionada con la percepción subjetiva de los seres humanos, la sensación térmica sí lo está. Si el descriptor objetivo, que representa la “sensación térmica”, no representase fielmente la percepción de los seres humanos, de nada serviría tener tal descriptor.

Los listening test permiten monitorizar la percepción subjetiva que presenta un determinado grupo de participantes ante un determinado estímulo sonoro y observar si el descriptor objetivo que categoriza tal suceso se corresponde ella. Permiten extraer conclusiones y posibles puntos de mejora del descriptor objetivo, fomentando su adaptación en caso de que no sea representativo de la percepción subjetiva.

¿Qué se ha de tener en cuenta a la hora de realizar un listening test?

Aunque la idea de desarrollo de un listening test puede parecer sencilla, lo cierto es que su proceso de desarrollo no es trivial. Este proceso se divide en tres etapas:

1. Preparación del listening test: El experimentador ha de decidir qué hipótesis quiere probar, qué estímulos, participantes y entorno va a utilizar para tal fin (y obtenerlos) y qué diseño de listening test es el más adecuado para la hipótesis que desea probar.
2. Desempeño del listening test: Los participantes realizan el test, aportando su juicio con respecto a las preguntas preparadas por el experimentador. Los juicios se recolectan mediante encuestas en papel o programas informáticos implementados a tal efecto.
3. Análisis de los resultados: Una vez que el grupo de participantes es suficientemente grande, el experimentador deberá extraer los resultados de cada participante y analizar estadísticamente todos los resultados de manera conjunta. Así se podrá determinar si la hipótesis definida puede ser probada.

La realización de estas tres etapas puede ser un proceso largo, tedioso y en algunas ocasiones extremadamente complicado.

Además, el hecho de que estos test se basen en la interacción con los seres humanos, los hace especialmente sensibles al sesgo. Principalmente existen, en el campo de los listening test, tres fuentes de sesgo [1,5]:

1. Sesgo del experimentador: Parcialidad, selección de hipótesis errónea...

2. Sesgo del elemento de consulta: Ambigüedad de las preguntas, entorno no real, estímulos de corta duración, interfaz poco intuitiva...
3. Sesgo del participante: Aprendizaje, fatiga, detección de patrón, falta de motivación...

La utilización de aplicaciones informáticas dedicadas durante todo el progreso de un listening test permite, por un lado, reducir la dedicación y esfuerzo tanto del experimentador como del participante y por otro lado minimizar el efecto de todas las fuentes de sesgo, si durante su desarrollo éstas se tienen en cuenta. Por ello, se decidió llevar a cabo una biblioteca de aplicaciones informáticas que permitiera automatizar el desarrollo de listening test y que tuviese por finalidad principal el análisis de descriptores objetivos utilizados para categorizar el aislamiento acústico, aunque aportando además gran flexibilidad a la hora de llevar a cabo listening test muy variados.

Esta motivación principal se debe a que no existe un descriptor objetivo estandarizado, de uso común en todos los países, para evaluar el aislamiento acústico, tal y como detalla [6]. Para evitarlo, varios grupos de investigación tratan de encontrar un descriptor único e investigan si la inclusión de componentes de baja frecuencia en este descriptor aportaría una mejor adecuación con respecto a la percepción de los usuarios.

La biblioteca de aplicaciones permitirá comprobar si estos nuevos descriptores objetivos del aislamiento acústico, se relacionan mejor con la percepción subjetiva de los usuarios.

DESARROLLO DE APLICACIONES INFORMÁTICAS

La biblioteca de aplicaciones se ha desarrollado en MATLAB R2013a, y se han utilizado comandos VBA de Excel para realizar ciertas conexiones necesarias entre ambas aplicaciones.

Descripción de la biblioteca de aplicaciones

La biblioteca de aplicaciones está compuesta por seis aplicaciones que se interrelacionan entre sí (*Imagen 1*), para cubrir todas las etapas de desarrollo de un listening test. A continuación se enumeran:

1. Aplicación de filtrado de los estímulos.
2. Aplicación de configuración del test (su usuario será el experimentador).
3. Aplicación de entrenamiento del test (su usuario será el participante).
4. Aplicación de recolección de respuestas o listening test (su usuario será el participante).
5. Aplicación de análisis estadístico de los resultados.
6. Aplicación de realización automática de audiometrías. (No finalizada)

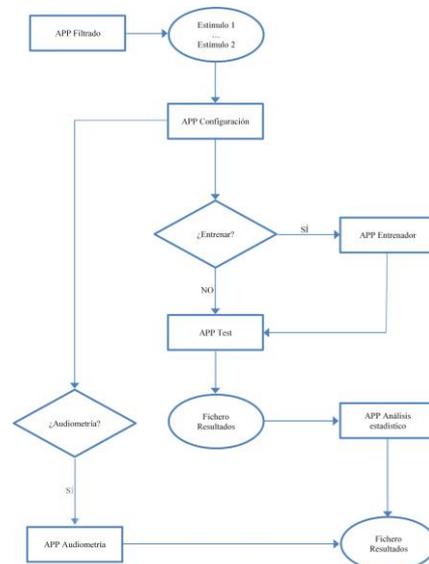


Imagen 1: Diagrama de relación entre las distintas aplicaciones.

Aplicación de filtrado de estímulos sonoros

Esta aplicación permite al experimentador filtrar, de manera sucesiva y automática, varios sonidos base por varias curvas de respuesta en frecuencia. Además, la aplicación muestra información tanto gráfica como tabulada relacionada con las curvas de respuesta en frecuencia utilizadas. Dado que el presente proyecto se encuentra enmarcado en el ámbito del aislamiento

acústico, las curvas de respuesta en frecuencia serán aquellas representativas del aislamiento acústico de determinados cerramientos.

Los cerramientos, que pueden ser cargados de manera manual o automática, son almacenados, si el experimentador lo desea, en una base de datos privada tanto local como accesible desde cualquier parte del mundo mediante protocolo ftp.

Debido a los hechos presentados anteriormente, esta aplicación simplifica considerablemente tanto el proceso de selección como el de filtrado de los estímulos que serán presentados a los participantes.

Se analizaron multitud de algoritmos de filtrado y finalmente se seleccionó un filtro FIR de magnitud arbitraria, al que se le realiza un filtrado bidireccional, con el fin de obtener una distorsión de fase nula. Mediante este algoritmo de filtrado se obtienen estímulos cuya correlación con estímulos sonoros filtrados con Audacity® es superior al 95%.

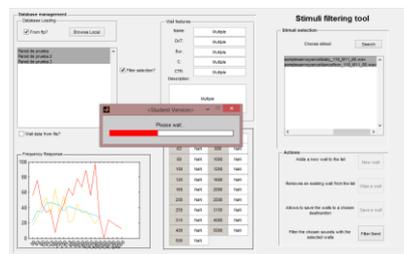


Imagen 2: Interfaz gráfica de la aplicación de filtrado. La barra de progreso informa al experimentador del % de filtrado completado.

Aplicación de configuración del test

Permite al experimentador diseñar y configurar el test que se presentará a los participantes. Se puede seleccionar qué método de consulta, preguntas y estímulos se utilizarán durante el test. Permite seleccionar un botón de reproducción de estímulos automático o dos botones manuales, para el método de comparación por pares, así como escalas verbales, numéricas e icónicas, para el método de escalado. Incluye, además, un tono de calibración que permite calibrar el nivel dentro de la aplicación.

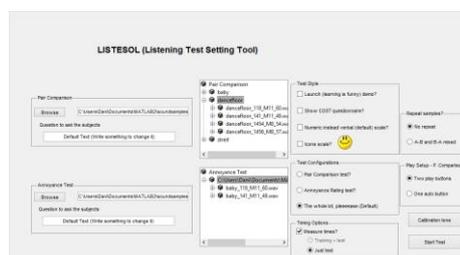


Imagen 3: Interfaz gráfica de la aplicación de configuración del test.

Aplicaciones de listening test y entrenamiento

La aplicación de listening test permite a los participantes interactuar con los estímulos y presentar sus respuestas. Puede presentar hasta seis pantallas distintas, que van variando a medida que se progresa, en función del diseño realizado por el experimentador. La primera de las pantallas presenta el test al participante y le detalla cuál va a ser el proceso, evitando de esta manera la interacción directa entre participante y experimentador. La segunda se

corresponde con el método de consulta de comparación por pares (*Imagen 4*), que mostrará el diseño, estímulos y pregunta seleccionados en la aplicación de configuración. De la misma manera, la tercera de las pantallas (*Imagen 5*), presentará el método de escalado con una escala verbal, numérica o icónica en función de la configuración realizada por el experimentador. A continuación es posible presentar una pantalla que muestra dos preguntas del cuestionario desarrollado por el proyecto COST Action TU0901 [7]. La quinta pantalla presenta al participante preguntas sociodemográficas (edad, género, nacionalidad y ocupación) muy útiles para experimentador, con fines estadísticos. Por último se le presenta al usuario una pantalla de despedida indicándole que el test ha finalizado. En ese momento, MATLAB administra los resultados y actualiza una base de datos Excel que contiene los resultados de todos los test realizados hasta el momento.

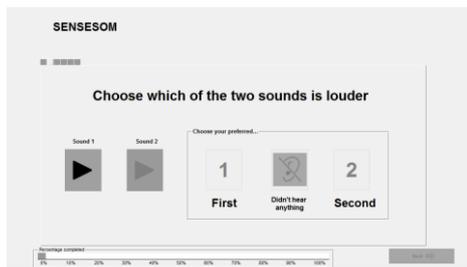


Imagen 4: Pantalla del método de comparación por pares con dos botones.



Imagen 5: Pantalla del método de escalado con escala numérica.

Con respecto a la aplicación de entrenamiento, ésta se configura de la misma manera que la de test y se presentará, en caso de que el experimentador lo haya seleccionado, antes de que aparezca la aplicación de listening test, para que el participante se familiarice con el proceso. La aplicación de entrenamiento guiará al participante mediante iconos coloreados y ventanas emergentes con información acerca de los pasos que debe llevar a cabo.

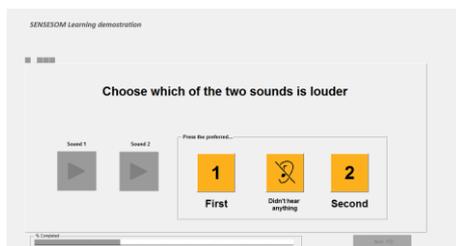


Imagen 6: Iconos coloreados para guiar al usuario a tomar un primer contacto.

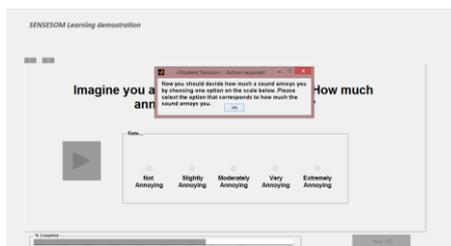


Imagen 7: Ventana emergente con instrucciones para guiar al usuario.

Aplicación de análisis estadístico

Esta aplicación no posee interfaz de usuario. Analiza estadísticamente la base de datos Excel que genera la aplicación de listening test y presenta resultados tanto tabulados como gráficos. El análisis estadístico se lleva a cabo por tres métodos, para comprobar si todos ellos presentan los mismos resultados:

1. Ranking no riguroso basado en el conteo de selecciones de cada estímulo.
2. Ranking obtenido mediante el método estadístico de análisis de la varianza (ANOVA), en función del conteo de selecciones.
3. Análisis estadísticamente riguroso basado en el análisis pormenorizado de los datos por parejas, siguiendo el procedimiento detallado en la norma UNE-EN ISO 5495:2009 [4].

ESTUDIO PILOTO

Los listening test se han utilizado en el campo del aislamiento acústico principalmente para observar si la correlación entre algunos descriptores acústicos del aislamiento y la percepción subjetiva de los usuarios es suficiente. Gran cantidad de estudios [3,8–12] han tratado estos temas, así como aportado diferentes variaciones metodológicas con el fin de generar una metodología común. De la misma manera, proyectos como Nordtest [2] y COST Action TU0901 “Building acoustics throughout Europe Volume 1: Towards a common framework in building acoustics throughout Europe.” [13] han aportado ciertas metodologías específicas para el desarrollo de listening test en el campo de la acústica y el proyecto COST, del aislamiento acústico en particular.

Se efectuó entre los días 30 de junio de 2014 y 11 de julio de 2014 una estancia de investigación en la Universidad Católica de Lovaina, con el fin de poner a punto las aplicaciones informáticas desarrolladas, así como de realizar un pequeño estudio piloto, en el ámbito del aislamiento acústico, que permita comprobar la efectividad de la biblioteca de aplicaciones. La Dra. Monika Rychtáriková fue la anfitriona en aquella institución y guió el proceso de desarrollo del estudio piloto.

Motivación, hipótesis y estímulos

El estudio piloto desarrollado se basa en uno de los experimentos previos [14] de la Dra. Rychtáriková, tratando de comprobar si los resultados obtenidos mediante esta nueva biblioteca de aplicaciones se corresponden con los de aquel estudio.

Para ello se utilizan tanto la misma hipótesis como los mismos cerramientos utilizados en aquel estudio [14]. Sin embargo, el estudio piloto solo utiliza como señal de excitación ruido rosa.

Se trata de probar si ante un mismo valor del descriptor objetivo de aislamiento R_{living} [15] (equivalente al $R_w+C_{50-5000\text{Hz}}$), un cerramiento ligero se comporta, acústicamente hablando, de la misma manera que uno pesado, en cuanto a la percepción subjetiva de usuarios. Para ello se utilizan cinco parejas de cerramientos (*Imagen 7*) con distintos valores de R_{living} . Cada una de estas parejas estará compuesta de un cerramiento pesado y uno ligero con el mismo valor de R_{living} . El método de consulta utilizado será el de comparación por pares, ya que se consultará a los participantes acerca de la sonoridad de los estímulos.

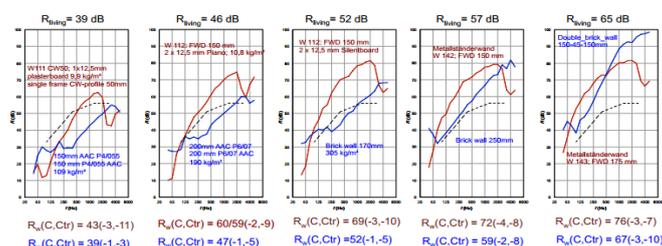


Imagen 7: Respuesta en frecuencia de los cerramientos utilizados.

Entorno y medios utilizados

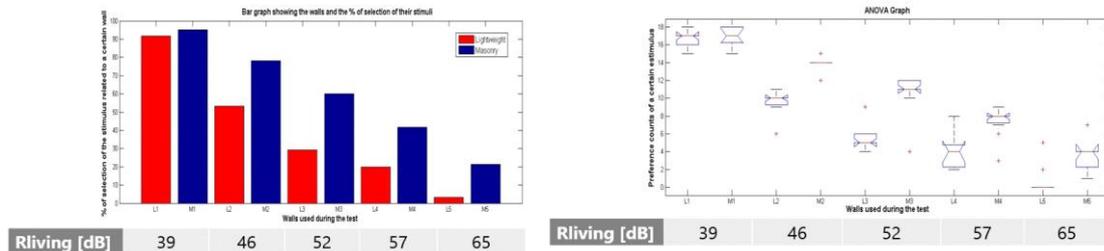
Para llevar a cabo el listening test se hizo uso de la cámara semianecoica del Departamento de Física Aplicada de la Universidad de Lovaina, que presenta un $R_{\text{fondo}} < 15$ dB [16].

Además, se han utilizado elementos electroacústicos de la casa Head Acoustics (Etapa de amplificación HPS IV, auriculares HA y cabeza artificial HMS III) para reproducir los estímulos y calibrar el sistema.

Resultados del estudio piloto

El test se llevó a cabo por 14 personas, de seis nacionalidad distintas, entre los 23 y los 49 años, de los cuales 3 fueron mujeres y el resto hombres.

Los resultados obtenidos mediante los dos primeros métodos de análisis se presentan en las *Imagen 8* e *Imagen 9*:



Los resultados presentados permiten demostrar que para los cerramientos utilizados y una excitación de tipo ruido rosa, el descriptor objetivo R_{living} no representa la sensación subjetiva de los participantes, subestimando las capacidades acústicas de los cerramientos ligeros, cuyos estímulos asociados son siempre seleccionados en menor proporción que los asociados a los cerramientos pesados, para mismo valor de R_{living} .

El tercero de los métodos de análisis utiliza las tablas de la norma UNE-EN ISO 5495:2009 [4] y permite llegar al mismo resultado que los dos métodos anteriores, con los valores de significancia estadística mostrados en la *Tabla 1*.

Tabla 1: Conteo de selecciones por parejas y valor de significancia con el que se puede probar la hipótesis.

	Masonry	Lightweight	Null	Total	α
$R_{living} = 39$ [dB]	18	7	0	25	0.10
$R_{living} = 46$ [dB]	25	0	0	25	0.001
$R_{living} = 52$ [dB]	23	2	0	25	0.001
$R_{living} = 57$ [dB]	22	3	0	25	0.001
$R_{living} = 65$ [dB]	21	1	3	22	0.001

CONCLUSIONES

Los listening test son una herramienta muy útil para determinar si existe correlación entre la percepción subjetiva y el descriptor objetivo de un suceso. Se ha de investigar en el desarrollo de una metodología estandarizada para cada ámbito del campo acústico. El R_{living} no parece representar fielmente la sensación subjetiva de los usuarios, por lo que se deberá investigar en descriptores objetivos del aislamiento acústico que sí representen correctamente la sensación subjetiva de los usuarios.

AGRADECIMIENTOS

Se agradece al Departamento de Física Acústica y Térmica de la Universidad Católica de Lovaina, y en particular a la Dra. Monika Rychtáriková, el Dr. Christ Glorieux y Daniel Urbán, por la prestación de sus medios y la ayuda a la hora de llevar a cabo el estudio piloto.

BIBLIOGRAFÍA

- [1] Meilgaard MC, Carr BT, Civille GV. Sensory Evaluation Techniques. 4th Editio. CRC Press; 2006.
- [2] NT ACOU 111 Acoustics : Human sound perception – Guidelines for listening tests. vol. 1. Espoo: Nordtest; 2002.
- [3] Poulsen T. Psychoacoustic Measuring Methods. Version 2.3. 2007.
- [4] UNE-EN ISO 5495:2009 - Análisis sensorial - Metodología - Prueba de comparación por parejas. Asociación Española de Normalización y Certificación; 2009.
- [5] Lazar DJ, Feng DJH, Hochheiser DH. Research Methods in Human-Computer Interaction. John Wiley & Sons; 2010.
- [6] Machimbarrena M, Fausti P. En busca de un marco común europeo en materia de aislamiento acústico en viviendas: Intercambio de experiencias y dificultad de consenso. Tec. 2013, vol. 0901, Valladolid: Sociedad Española de Acústica; 2013, p. 33.
- [7] COST Action TU0901: Integrating and Harmonizing Sound Insulation Aspects in Sustainable Urban Housing Constructions. <http://www.costtu0901.eu/> n.d.
- [8] Rychtáriková M. Listening Matters in Building and Room Acoustics. In: European Acoustics Association, editor. Euronoise 2012, Prague: European Acoustics Association; 2012, p. 24–33.
- [9] Hongisto V, Oliva D, Keränen J. Disturbance caused by airborne living sounds heard through walls – preliminary results of a laboratory experiment. Internoise 2013, Innsbruck: Institute of Noise Control Engineering; 2013.
- [10] Bradley JS. Evaluation of revised ISO airborne sound insulation ratings. NRC Constr 2012.
- [11] Pedersen TH, Antunes S, Rasmussen B. Online listening tests on sound insulation of walls – A feasibility study. Euronoise 2012, Prague: European Acoustics Association; 2012, p. 6.
- [12] Ordoñez R, Vicentini C, Markovic M, Fausti P. Objective and subjective evaluation of façade sound insulation. Internoise 2013, Innsbruck: Institute of Noise Control Engineering; 2013.
- [13] COST Action TU0901. Building acoustics throughout Europe Volume 1: Towards a common framework in building acoustics throughout Europe. DiScript Preimpresion, S. L.; 2014.
- [14] Rychtankova M. Influence of temporal and spectral features of neighbour's noise on perception of its loudness. Internoise 2013, Innsbruck: Institute of Noise Control Engineering; 2013.
- [15] Scholl W. Revision of ISO 717: Future Single-Number Quantities for Sound Insulation in Buildings? Forum Acusticum 2011, Aalborg: European Acoustics Association; 2011.
- [16] Muellner H. Methodology for laboratory listening tests with focus on building acoustics requirements - Appropriate test statistical applications. Leuven: 2013.