atf 2017
5th conference on applied technology, 14.-16. September, Zagreb, Croatia

THE CHALLENGES IN PREPARING THE STIMULI TO BE USED IN SUBJECTIVE EVALUATION OF IMPACT SOUND INSULATION

Marko Horvat[1], Jakub Benklewski[2], Kristian Jambrošić[1], Herbert Müllner[2], Monika Rychtáriková[3], and Rudolf Exel[4]

[1] University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb Croatia, Email: marko.horvat@fer.hr , kristian.jambrosic@fer.hr

[2] TGM Vienna, Department of Acoustics and Building Physics, Wexstraße 19-23, A-1200 Vienna, Austria, Email: jbenklewski@tgm.ac.at , hmuellner@tgm.ac.at

[3] KU Leuven, Faculty of Architecture, Leuven, Belgium, Email: monika.rychtarikova@luleuven.be

[5] Unternehmensberatung Rudolf Exel, A-7423 Grafenschachen 343, Austria, Email: beratung@exel.at

ABSTRACT

This paper offers an overview of the general guidelines related to the preparation of the sound stimuli to be used in listening tests devised for subjective evaluation of impact sound insulation properties of building elements. Different approaches are presented to both collecting the sound samples and to converting them into a form that can be used for carrying out the listening tests. The advantages and the limitations of the presented approaches are discussed.

Besides general guidelines, an example is given on the preparation procedure of impact stimuli in the form of footsteps sound as a result of walking on a typical floor, as heard in the room below. The variables taken into account are floor construction, type of footwear worn by the walker, the weight of the walker, and walking tempo. For the recording part of the procedure, the specifics are given on the recording setup and the acoustical conditions in the recording room. The conversion of raw sound samples into a form suitable for reproduction over headphones and loudspeakers is described as well. Preliminary analysis of the collected sound samples is presented, regarding the overall level and loudness, as well as the spectral content, with respect to the considered variables.

INTRODUCTION

In any kind of subjective or objective assessment made through listening tests, the process of collecting the sound samples to be presented to a jury of listeners is a crucial part of listening test design. The samples need to be collected in a way that will maintain the highest audio quality. Furthermore, it is often required that raw samples be prepared through conversion to a form that can be utilized in listening tests.

The common trait of virtually all listening tests is that the sound samples have a normal or even elevated level, depending on the phenomenon being examined. However, the listening tests related to subjective assessment of sound insulation, both airborne and impact, need to overcome the issue of dealing with low-level sound samples. These sound samples need to represent realistic sounds found in an average home or workplace that can be heard from neighbouring housing units or from outdoors. Given that the sound insulation is at least adequate, such sounds should be either inaudible or barely audible.

To obtain a fully realistic impression of the acoustic situation in an average home that needs to be recreated for the purpose of listening tests aimed at assessing the sound insulation in a subjective way, the sound samples need to have the direction of arrival assigned along with the realistically set level. Therefore, spatial audio reproduction techniques have to be used, with the sound samples properly processed and prepared, in order to meet these requirements.

Some research on the methodology of the listening tests related to evaluation of sound insulation has been done [1, 2, 3, 4]. The research presented in this paper is an extension of these efforts and is carried out within the scope of the H2020-MSCA-RISE-2015:papabuild project [5]. It is focused on yielding a set of guidelines related to collecting and preparation of sound samples to be used in listening tests specifically designed for subjective assessment of impact sound insulation. These guidelines stem from experimental research conducted in order to address all aforementioned requirements that need to be met.

The paper gives an example of a sample collecting process in the form of directly recorded footsteps on a floor partition built in a laboratory that could represent a real partition between two apartments.

GENERAL GUIDELINES

The general guidelines on the procedures of preparing sound stimuli that can be used in subjective evaluations of sound insulation can roughly be divided into three categories.

The first one deals with maintaining the actual sound pressure level of the prepared stimuli throughout the recording, processing and reproduction stages. To achieve this, it is necessary to have both the recording and the reproduction system calibrated. The easiest way to do this is simply to measure the sound pressure level of the sounds during recording, and to repeat the procedure during reproduction, thus yielding the value of the required level correction. A calibration sound such as pink noise or a 1-kHz tone can be used instead of the sound stimuli being recorded and later reproduced. A frequency correction must be applied to compensate for non-flat frequency responses of both the recording microphone and the headphones or loudspeakers used for reproducing the sound stimuli.

The second category of guidelines deals with noise, which represents a significant factor of influence when low-level sounds are concerned, and can cause irreversible damage to the stimuli to be prepared, thus making it utterly useless. It is crucial to use low-noise microphones for recording of the stimuli. Only then can noise be suppressed enough in raw recorded sounds not to represent a problem in the processing and reproduction stage. As certain steps in the processing stage add noise as well, care should be taken to amplify the raw recorded sounds as much as possible before forwarding them to the processing stage, while maintain the noise in the recording itself at acceptable level.

The third category of guidelines deals with achieving the correct direction of arrival of the reproduced sound, from the listening point of the listener who is involved in the listening tests. Specifically, in listening tests that deal with assessments of impact sound insulation, one of the likely directions of arrival is from above. To make the listening test and the sound scenery believable, the listener must perceive the reproduced sound as if it was indeed coming from above. The direction of arrival can either be preserved during recording, or be assigned in the processing stage.

The next section describes the specifics of the preparation of the sounds of footsteps as the impact stimuli, and gives the details on the steps that need to be taken to meet these general guidelines. As the sounds that will serve as the impact stimuli are of very low level, many options that are normally available cannot be used. These limitations are presented in detail.

AN EXAMPLE OF THE PREPARATION PROCESS – SOUNDS OF FOOTSTEPS AS THE STIMULI

To test the possibilities of creating appropriate stimuli to be used for subjective assessment of impact sound insulation, a recording session was carried out in laboratory conditions. Specifically, the sounds of footsteps on a floor were recorded as a typical type of impact stimuli found in multi-storey housing buildings and as a common source of noise coming from the floor above. The floors included in this session were built in the measurement facility. They were composed of a standard load-bearing part with the floating floor added on top of it as the improvement against impact noise, and were ready for the final layer to be installed (parquet, ceramic tiles, etc.).

One lightweight and one heavy floor were prebuilt in the measurement facility and as such have served as representatives for both widespread types of floor constructions. The objective assessment of their impact sound insulation properties go in favour of the heavy floor. The single-number value of $L_{n,w}$ was found to be 39 dB for the heavy floor, whereas the value of the same parameter for the lightweight floor was measured and evaluated at 56 dB. This suggests that the level of the recorded footsteps sounds will be considerably higher for the lightweight floor, whereas for the heavy floor these sounds will be barely audible. These conclusions are supported by the results of the analysis of recorded sounds in terms of their level, both unweighted and A-weighted, as will be shown further down.

To obtain the desired sounds of footsteps, several people were included in the experiment with the task of walking on the floor, thus producing the required footsteps, the sounds of which were recorded in the recording room below. To help the walkers maintain a stable walking tempo, they were asked to synchronize their steps with the sound of a metronome realized through a smartphone application. The walking path was drawn on the floor itself to help the walkers maintain a stable tempo. The requirement put on the path was that its curvature should be as small as possible and change as little as possible. Therefore, in case of the lightweight floor the path was truly circular and it was made to utilize almost the entire surface of the floor. The path on the heavy floor was made mildly elliptic due to the overall shape of the room, but also having in mind the desire to use as much surface area of the floor as possible.

The research encompassed several different variables, apart from the type of floor construction. It was assumed that the mass of the walker will play a certain role in the level and spectrum of the resulting sound of their footsteps, although mass is only one factor in the overall kinetic energy transferred from the walker to the floor on impact of their feet. It is foreseeable that the velocity of the foot on contact with the floor rather than mass is the prevailing factor when it comes to footsteps of children. However, due to ethical reasons, as well as the questionable ability of children to maintain a stable walking tempo, they were not included as walkers in this particular experiment. Specifically, three adult walkers took part in the experiment, and each of them was chosen to represent a certain weight/mass category, designated as light, medium and heavy. The female walker with a mass of 60 kg was chosen to represent the light category. A male walker weighing 85 kg represented the medium category, while another male walker weighing 115 kg represented the heavyweight category.

To take into account the fact that people wear different kinds of footwear at home, the walkers were asked to walk in three different conditions: wearing sneakers, wearing casual shoes, but also wearing no footwear, just having socks on their feet.

The walking tempo was chosen as a variable as well, in the sense that the walkers were asked to do their walk in three different tempos: 45, 75 and 105 steps/min. The slowest tempo of 45 steps/min was chosen due to the constraints made by the acoustical properties of the recording room, i.e. its reverberation time. The goal here was to obtain a recording that will provide the sounds of footsteps as individual sound events. The tempo of 75 steps/min was recognized as normal pace for an adult, and is close to their heart rate. The recordings made at this tempo can still be used for extraction of single sound events, but can also be used as a continuous stimuli. Finally, the fastest tempo of 105 steps/min was recognized as fast pace, and these recordings can only be used as continuous stimuli.

RECORDING

To investigate the limitations of the recording equipment, it was decided to employ several different independent recording systems.

Binaural recordings intended for direct reproduction over headphones were made using two different systems, the Squadriga recorder with a pair of microphones with the correction that is independent of direction, and the standard Kemar dummy head. A 1-st order Ambisonics recording using a TetraMic microphone was made as well, to be directly reproduced over a loudspeaker-based spatial audio reproduction system based on Ambisonics coding and decoding. All three recordings were made at the recording height of 1.5 m. Finally, a simple monaural recording was made using a dedicated low-noise measurement microphone B&K 4179 placed 1 m below the ceiling, i.e. close to the ceiling as the sound source, in order to capture as much direct sound as possible.

Binaural and Ambisonics recordings already contain directional information, thereby offering the possibility of correct localization of the source in direct reproduction. Monaural recording, however, contains only the content of the sound, but it has no directional information, which must be reassigned in the processing stage. Therefore, the monaural recording can only serve as the input to the processing stage.

By definition, the recording room in the measurement facility has the reverberation time between 1 and 2 seconds. This much reverberation is not acceptable for recording purposes. Therefore, the acoustical conditions in the recording room were corrected by means of additional absorption, achieved mainly by placing 10-cm thick plates of glass wool on the floor and the walls of the recording room. The quantity was determined with the requirement to reduce the reverberation time to less than 0.5 seconds at middle frequencies. Additional cube-shaped absorbers were placed in the corners of the room to help even out the reverberation time and make it more stable with frequency. The absorbers were placed in the room in a way that would suppress the occurrence of pronounced room modes and to reduce the flutter-echo as much as possible.

PROCESSING AND PREPARATION

The processing and preparation stage involves the conversion of raw samples into a form suitable for reproduction.

In case of binaural recordings a minimum amount of preparation is needed, in the form of a frequency-dependent amplitude correction that takes into account the position of the microphones during recording (at the eardrum, in the ear canal, outside the ear). Ambisonics recordings need to be decoded for the specific loudspeaker configuration used for reproduction purposes. The decoder itself can be put into the processing chain within the DAW software, and the samples can then be processed and reproduced in real time. Another possibility is to decode the recordings in advance and then use the DAW software solely for reproduction over the loudspeaker system.

After the completion of the recording session, all recordings were analysed, and a test listening session was made to determine if these recordings meet the requirements stated above. It was determined that binaural recordings reproduced over headphones cannot provide for adequate and correct localization when the sound is coming from above, as is the case here. Limited success was achieved with recordings made with Squadriga recorder, while the dummy head recordings proved to be useless. On the other hand, Ambisonics recordings reproduced over loudspeakers do provide correct localization, in this case from above. However, none of these recording systems were initially designed as low-noise systems, and their microphones all have the noise floor too high. As a consequence, all the recordings, binaural and Ambisonics, suffer from excessive noise, and are not usable for reproduction to listeners in listening tests.

When it comes to meeting the low-noise requirement, only the recordings made with a dedicated low-noise measurement microphone fully satisfy this requirement due to the very low noise floor of the microphone itself. As the microphone itself is an ordinary pressure microphone with an omnidirectional polar pattern, no information on the direction of arrival can be preserved in the recording stage. Therefore, the sound of the recorded footsteps has to be given the correct direction in the processing stage, if it is to be used in listening tests. Two different ways of achieving this were tested.

One approach was to create a simulated room of a size typical for living quarters and with appropriately short reverberation time using the ODEON room acoustics simulation software. The correct direction of arrival of the impact sound can be assigned in a simple manner by placing the sound source in the simulation directly above the receiver point, right under the ceiling in the simulated room. A more advanced approach is to define multiple sources in a layout that forms a predefined path such as a circle, a figure-of-eight or some other form, or it can be a bit more randomized. Both ways can serve as a simulation of a person walking on a floor above, in a regular or irregular pattern. To achieve the desired sound of continuous moving footsteps, single sound events, i.e. individual footsteps sounds have to be cut out of the recording and each of them has to be assigned to each sound source with a correct delay to create the illusion of continuous walking. It is not recommendable to use the same sound for all the sources, as the resulting sound sounds too perfect and uniform. Instead, as many individual sounds as possible should be cut out from the original recording, and a database of such sounds should be formed. Individual sounds from the database should be assigned to sound sources in a random pattern.

The output from ODEON can be encoded both for binaural and Ambisonics reproduction. Both was tested through listening and it was found that loudspeaker-based Ambisonics reproduction is superior to binaural reproduction over headphones, when it comes to achieving a perceptually adequate and believable test scenario. The issue that needs to be addressed is poor handling of low-level source sounds in ODEON. Noise is added in the process of convolution of these sounds with the appropriate impulse responses, if the source sounds have too low a level. Therefore, the recorded sounds must be amplified as much as possible, while maintaining the inherent noise inaudible, before they are cut and sent to ODEON as individual source sounds. Another issue that requires attention is the problem of triple reverberation; the reverberation of the recording room contained in the recording itself, the reverberation of the simulated room in ODEON, and the reverberation of the listening room where the reproduction/listening takes place. The recorded reverberation cannot be avoided, other than treating the recording room with excessive amounts of absorbing material, which is not desirable. The reverberation of the simulated room can be finely tuned and controlled by choosing the appropriate materials for the surfaces of the room. The reverberation of the listening room is a potential problem only for loudspeaker-based reproduction, whereas the room does not have an influence if headphones are used. However, special-purpose listening rooms such as ones that host Ambisonics-based loudspeaker systems are usually well dampened, and their reverberation time is usually kept under 0.2 seconds in a broad range of frequencies, with a possible upward deviation at low frequencies.

The second way of creating scenarios to be used in listening tests is to use the recorded sounds directly in the DAW software that is used for reproduction anyway. The recorded mono track needs to be assigned a 3D panning plugin capable of panning the apparent source to the desired spatial coordinates. Like in ODEON, the panning plugins are able to create sounds appropriate for both binaural and Ambisonics reproduction. A clear advantage over the simulations in ODEON is that in this case no acoustic simulations are needed, thus removing a time-consuming step in the preparation process that cannot be done in real time. The processing itself is simpler in this case and can be done on-the-fly, while the sound is being reproduced. The problem of reverberation being introduced multiple times is now reduced only to recorded reverberation and the one added by the listening room, which means that most of the reverberation is contained in the recording itself. The panning plugins can create an illusion of a single direction of arrival, corresponding to a single source defined in ODEON simulations. However, the parameters used to define the direction of arrival in these plugins can be changed manually, thus allowing the apparent source to move in virtually any path in a 3D space. An advanced approach that ensures the repeatability in this processing stage is to automatize the changes of these parameters, thus enabling the creation of a stable source path that does not change, unless a different automation procedure is defined. A great advantage is the fact that the source path is practically continuous, as opposed to the simulations in ODEON in which the source path is made of discrete steps, using many sources. In this case, the apparent source is moving automatically as the reproduction advances with time. The entire original recording can simply be reproduced through a plugin (and the Ambisonics decoder in case of Ambisonics-based loudspeaker reproduction), without the need to cut out individual sounds of the footsteps, thus removing another time-consuming part of the processing stage.

ANALYSIS OF THE COLLECTED SAMPLES

All the collected sound samples were analysed to determine their overall level and loudness, as well as their spectral content, with respect to all the considered variables.

As stated earlier, the walking tempo of 75 steps/minute was recognized as normal pace for an adult that is most likely to be found in a normal living situation, and the recordings made with the walkers walking at this tempo can be utilized in both approaches to processing the recordings that were described earlier. Therefore, to maintain brevity, only the results of the analysis made on sound samples recorded for this walking tempo are shown below.

Table 1. The overall level, unweighted and A-weighted, of the sounds of the footsteps produced by three walkers of different mass, wearing different footwear on a lightweight and a heavy floor, walking at 75 steps/minute

Walker	Footwear		Lightweight floor	Massive floor
light	Shoes	L (dB)	75,9	47,8
		L (dBA)	28,1	14,8
	Sneakers	L (dB)	78,6	43,3
		L (dBA)	25,4	9,7
	Socks	L (dB)	81,5	58,0
		L (dBA)	30,4	24,1
medium	Shoes	L (dB)	77,5	49,0
		L (dBA)	23,5	16,0
	Sneakers	L (dB)	77,1	51,9
		L (dBA)	24,2	17,6
	Socks	L (dB)	80,6	60,8
		L (dBA)	29,1	27,0
heavy	Shoes	L (dB)	82,0	49,8
		L (dBA)	28,9	14,1
	Sneakers	L (dB)	83,2	49,7
		L (dBA)	28,5	13,5
	Socks	L (dB)	83,5	63,4
		L (dBA)	33,4	29,5

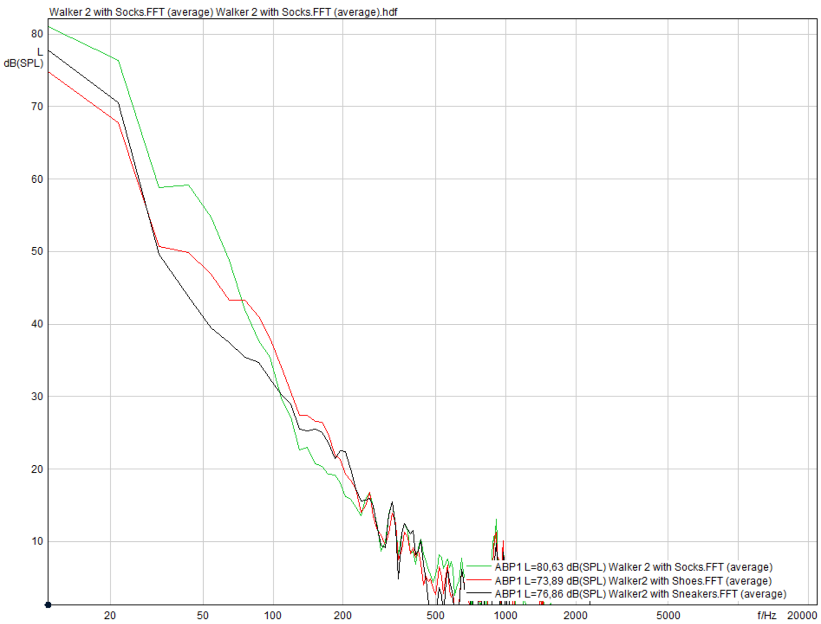

Figure 1. The spectra of footsteps sounds produced by the light walker on the lightweight floor, wearing only socks (green curve), casual shoes (red), and sneakers (black), walking at 75 steps/minute

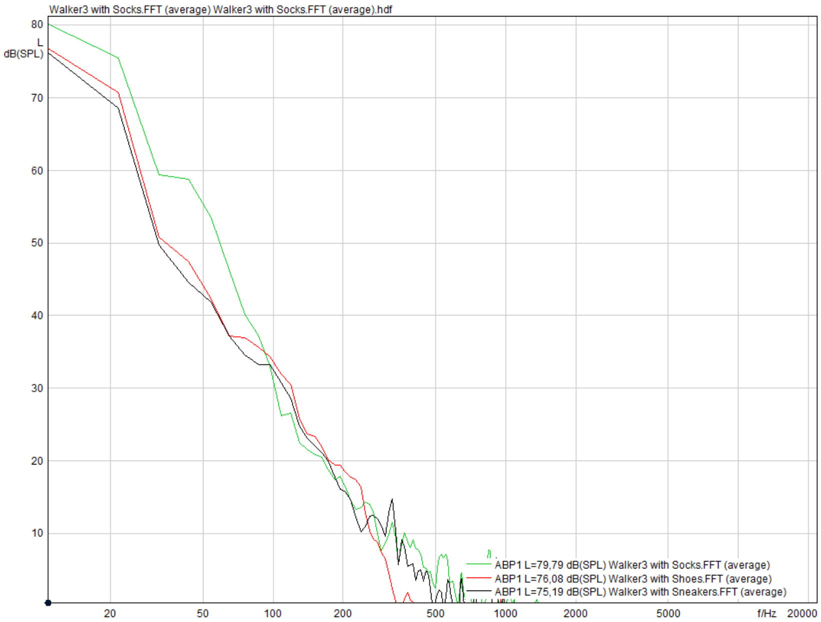

Figure 2. The spectra of footsteps sounds produced by the medium walker on the lightweight floor, wearing only socks (green curve), casual shoes (red), and sneakers (black), walking at 75 steps/minute

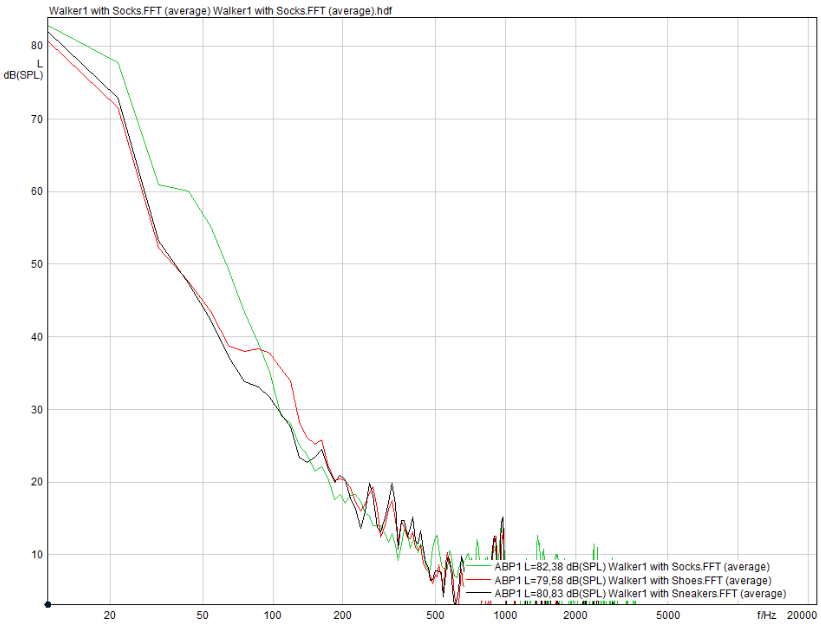

Figure 3. The spectra of footsteps sounds produced by the heavy walker on the lightweight floor, wearing only socks (green curve), casual shoes (red), and sneakers (black), walking at 75 steps/minute

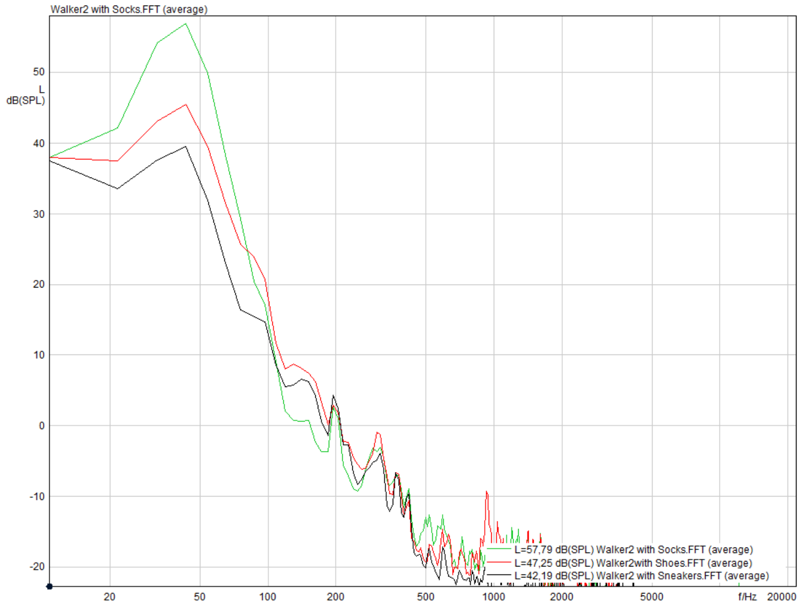

Figure 4. The spectra of footsteps sounds produced by the light walker on the massive floor, wearing only socks (green curve), casual shoes (red), and sneakers (black), walking at 75 steps/minute

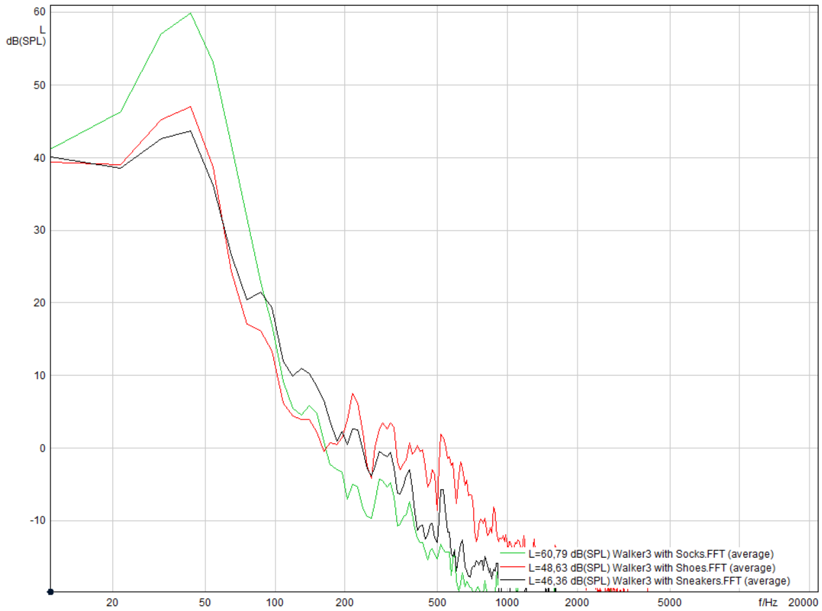

Figure 5. The spectra of footsteps sounds produced by the medium walker on the massive floor, wearing only socks (green curve), casual shoes (red), and sneakers (black), walking at 75 steps/minute

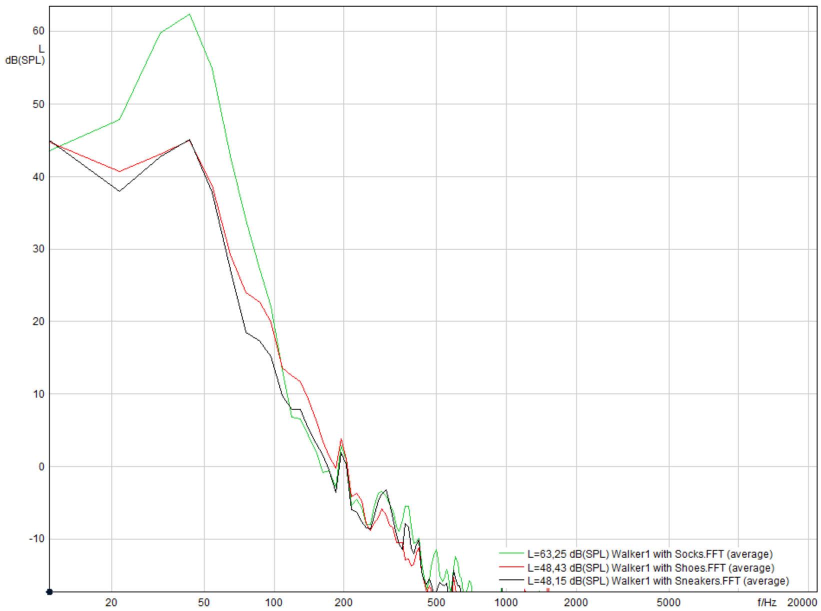

Figure 6. The spectra of footsteps sounds produced by the heavy walker on the massive floor, wearing only socks (green curve), casual shoes (red), and sneakers (black), walking at 75 steps/minute

The single-number values of the overall unweighted and A-weighted sound pressure shown in Table 1 reveal that the performance of the heavy floor is considerably better than the performance of the lightweight floor, as expected and proved by classic tapping machine measurements. However, the results suggest that the performance of the two floors depends on the type of footwear as well, and the difference in performance of the two floors is the smallest when the walker is not wearing any footwear, but walks only in socks, which is a common situation in an average home. Moreover, the difference in the sound pressure level of footsteps generated by any of the three walkers with and without footwear is much smaller for the lightweight floor, for which this difference does not exceed 6 dB, than for the heavy floor, for which the highest observed difference is 15 dB. In both cases the walkers produce the highest-level sounds with no footwear on their feet. As for the walkers themselves, their mass is only partially related to the sound pressure level of the sound of the footsteps they produce. However, the final result will also depend on the walking style of each walker, taking into account the velocity at the contact of the foot with the floor, as well as the type of contact of the foot and the floor. This contact can be soft or hard, depending on whether the initial contact point is the heel, the fingers or something in between.

The spectra shown in Figures 1 to 6 were made for each walker wearing different footwear on a lightweight and a heavy floor. For both floors the dominant spectral content is below 100 Hz, but the difference between the overall spectral shape obtained for the lightweight and the heavy floor is notable. While the spectral content of footsteps on the lightweight floor reaches its highest values in the infrasonic frequency region, the one obtained for the heavy floor reaches its maximum at about 50 Hz. The difference between walking in footwear and without it is clearly visible in all the spectra shown above. As stated before, removing the footwear has much more impact on the heavy floor than on the lightweight one.

CONCLUSIONS

The general guidelines on collecting and preparing the stimuli to be used in listening tests related to subjective assessment of impact sound insulation are rather straightforward and easy to satisfy when dealing with sounds of normal level.

When it comes to low-level sounds, which are readily dealt with in assessments of any kind of sound insulation, the stated guidelines impose many limitations. These limitations reduce the number and variety of the possible courses of action that can be taken when collecting and preparing the sound stimuli to be used in actual listening tests.

The conducted research shows that direct recording is the preferred way to obtaining usable stimuli for assessing impact sound insulation. Only true low-noise measurement microphones can be used in the recording stage. As no spatial information is maintained in this case, the correct direction of arrival of sound in relation to the listener needs to be recreated in the processing stage. This can be achieved either through simulations in dedicated room acoustics simulation software capable of auralisation, or in digital audio workstation software via appropriate plugins designed to deal with spatial audio. The latter option is the preferred choice, as it is simpler and can be done in real time, without the danger of introducing additional noise into the final sound stimuli.

Possible alternative ways of preparing the sound stimuli for this particular purpose remain the topic of future work.

ACKNOWLEDGEMENT

This research has been made within the scope of the H2020-MSCA-RISE-2015:papabuild project (grant agreement No. 690970).

REFERENCES

[1] M. Rychtáriková, M. Horvat (2013): Developing a Methodology for Performing Listening Tests Related to Building Acoustics // COST Action TU0901 – Building acoustics throughout Europe. Volume 1: Towards a common framework in building acoustics throughout Europe, Spain: DiScript Preimpresion, S. L., 2014, pg. 133-156.

[2] M. Rychtáriková, M. Horvat (2013): Developing a Methodology for Performing Listening Tests Related to Building Acoustics // COST Action TU0901 – Building acoustics throughout Europe. Volume 1: Towards a common framework in building acoustics throughout Europe, Spain: DiScript Preimpresion, S. L., 2014, pg. 121-142.

[3] M. Horvat, K. Jambrošić, H. Domitrović (2012): Examination of Required Signal-to-Noise Margin in Laboratory Subjective Evaluation of Sound Insulation. Proceedings of the 5th Congress of the Alps Adria Acoustics Association, 13 pgs.

[4] M. Horvat, H. Domitrović, K. Jambrošić (2012): Suitability of 3D Sound Reproduction and the Influence of Background Noise on Subjective Assessment of Sound Insulation. Proceedings of Euronoise 2012, pg. 1225-1230.

[5] www.papabuild.eu , last accessed on 18 October 2017