DOI: 10.61782/fa.2023.0385

PERCEPTION OF ROOM MODE DAMPING FROM A MUSICAL PERSPECTIVE

David Jun[1,2*] **Christ Glorieux**[3] **Zuzana Fisarova**[2]
Josef Plasek[2] **Monika Rychtarikova**[1]
[1] KU Leuven, Department of Architecture, Campus Brussels and Ghent, Belgium
[2] Brno University of Technology, Faculty of Civil Engineering, Brno, Czech Republic
[3] KU Leuven, Department of Physics and Astronomy, Heverlee, Belgium

ABSTRACT

Undamped acoustic modes inside rooms - especially in small ones - are often responsible for a decreased acoustic quality, in particular when the rooms' main function is "critical listening" and/or "playing musical instruments". Mainly in the latter case, treatment with broadband absorption is not always efficient and yields either uneven or too short reverberation times. Rehearsal rooms are typical cases where the "liveness" of the space needs to be guaranteed and putting a broadband absorber would make it "too dry". Rehearsal room design can be seen as a search for an optimal compromise between sound strength and reverberation time across the audible part of the spectrum. This contribution presents results from listening test experiments, in which the perception of modes was investigated, in particular with respect to what extent differences in axial room mode damping are audible.

Keywords: *room modes, listening test, modal decay.*

1. INTRODUCTION

The acoustic design of a small rehearsal room usually consists of defining adequate values for the volume, aspect ratio, reverberation time, room strength and background noise [1–3]. Figure 1 depicts the relation between two important room acoustic quantities, the strength and the reverberation time, and possible problems caused by too high or too low values. A complicating factor is that the reverberation time and the strength are frequency dependent, in particular in the case of small rooms, in which the spectral dependences are strongly affected by the existence of separated room modes. Practice rooms for individuals and even small ensembles usually have properties that cause the Schroeder frequency (the empirical lower frequency limit for the applicability of the statistical acoustics approaches) to be inside the frequency range of sound produced by many types of musical instruments and the human voice. This implies an imbalanced frequency response and decay characteristics caused by the room modes.

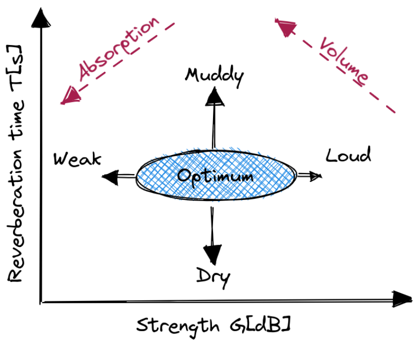

Figure 1. Relations between room reverberation and strength. Rehearsal room design can be viewed as a search for an optimum proportion between room volume and absorption (adapted from [1]).

Many studies have been performed to describe the human perception of resonances, mainly in critical lis-

*Corresponding author: David.Jun@vut.cz.

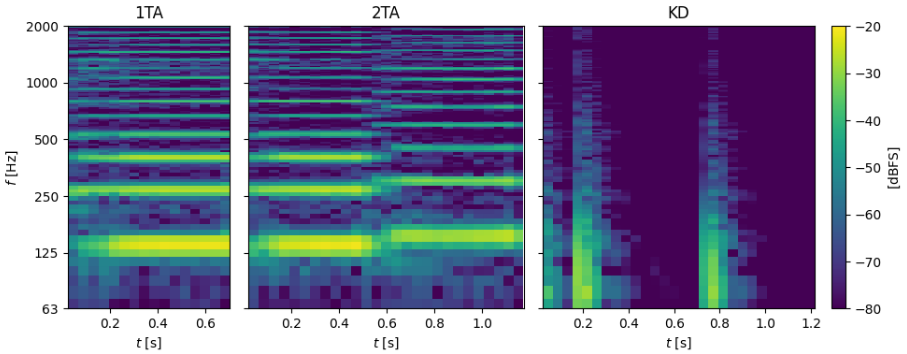

Figure 2. Spectrograms of the anechoic stimuli used for convolution in the ABX listening test: one (1TA) and two (2TA) tone stimuli played with a bow on a double bass, and a simple rhythmic pattern played on a kick drum (KD).

tening environments. Toole and Olive [4] studied perception thresholds for delayed and undelayed resonance peaks added to a signal path of a loudspeaker system. The experiments focused on pulses and broadband noise. The authors mention that continuous broadband sound appears to be advantageous for the detection of resonances and that the detection is in general less difficult for wide (low-Q) resonances of the same peak amplitude than for narrow (high-Q) ones. For more transient sounds, adding reverberation can be beneficial for the detection of a lower-Q resonance. Some further tests were performed by Olive et al. [5], who compared the detection of both peaks and dips and showed that (a) pulses perform similarly for high-Q peaks and notches and (b) pink noise performs the best for peak detection of various Q values, but is not effective in revealing high-Q notches. Kritly et al. studied the audibility of coincidence dips and structural resonance dips in sound isolation spectra [6] and of spectral dips and peaks in broadband sound [7, 8]. Avis et al. [9] extended the research towards room acoustics, performing listening tests aiming at the room mode detection thresholds based on the constant resonance Q of all the modes below the Schroeder frequency. They suggested a limit value of Q=16, but at the same time mentioned that the modal decay may be a more appropriate descriptor. In one of the latest works in this area by Fazenda et al. [10], the authors elaborated on the dependence of absolute mode detection thresholds on the modal decay. Tests were done for both musical an artificial stimuli and the followed approach included a combination of a low-frequency reverberant sound and a high-frequency anechoic sound. Work by Rizzi et al. [11] aimed at the transient behaviour of a room based on the so-called ”Overshoot Response”, i.e. the room response to short sounds, which often does not lead to a steady acoustic field development.

In this work, we extend the above research to the reverberant environment of a shoe-box type practice room for music rehearsal, of which the acoustic response is crucial for the users. This contribution addresses mainly two questions: (a) What are the perceivable changes in room mode damping in a practice room for individuals? (b) How is the discrimination performance affected by the nature of the instrument sound? In the following, we elaborate on the simulation and auralization of sounds in rooms with different modal characteristics, and on the results of listening tests that were aimed at assessing information on the audibility of the differences.

2. METHOD

A music rehearsal room was assumed with dimensions 2.63×3.40×3.40 m ($V = 30.4$ m^3, which is somewhat smaller than the ISO 23591 [3] recommended minimum for this kind of rooms, 35 m^3). The room parameters were chosen to mimic the common situation of the use of a room that has been ad hoc converted into a (non-ideal) practice room, and also for 2.63 m dimension to correspond with the second lowest room mode frequency at 131 Hz (which also corresponds to the fundamental frequency of the C_3 note). According to ISO 23591, the optimal reverberation time based on is $T_{opt} = 0.4$ s, which corresponds to the Schroeder fre-

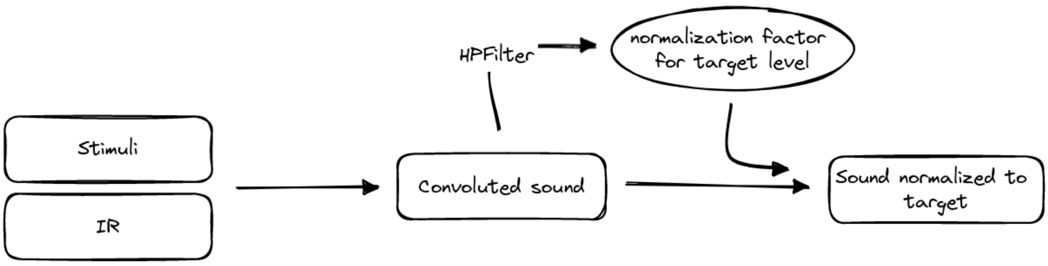

Figure 3. Processing steps applied on the stimuli during the listening tests. After the convolution, the normalization factor is deduced from the high-pass filtered sound. The original convolved sound is then multiplied by this factor.

quency $f_S = 2000\sqrt{T/V} = 229$Hz. Similarly to the work by Fazenda [10], this study quantifies mode damping by modal decay. Any modal resonance can be described based on its quality factor Q $[-]$

$$Q = \frac{f}{\Delta f}, \tag{1}$$

where f [Hz] is the central frequency of the resonance and Δf [Hz] is its spectral bandwidth described by the -3 dB decrease of SPL when compared to the resonance peak value.

This modal Q value can be then directly related to the modal decay T_{60} [s] by [12]

$$T_{60} \approx \frac{2.2Q}{f} \approx \frac{2.2\pi}{\delta}, \tag{2}$$

where $\delta_n = \Delta f/\pi$ $[\mathrm{rad}^{-1}]$ is a damping factor [13]. Also similarly to the work of Fazenda et al. [10], the room impulse response was calculated following Kuttruff in [13], with some changes:

$$p_\omega(\omega, r) = iQ_v c^2 \rho_0 \sum_n \frac{p_n(r)p_n(r_0)}{(\omega^2 - \omega_n^2 - 2i\delta_n\omega)\Gamma_n V}, \tag{3}$$

where r [m] and r_0 [m] are the receiver and source positions respectively, Q_v $[\mathrm{m}^3\mathrm{s}^{-1}]$ is the volume velocity of the source, c $[\mathrm{ms}^{-1}]$ is the speed of sound in air, ρ_0 $[\mathrm{kgm}^{-3}]$ is the density of air, ω $[\mathrm{rad}^{-1}]$ is the angular frequency and V $[\mathrm{m}^3]$ the room volume. The quantity K_n $[\mathrm{m}^3]$ in Kuttruff's notation has been replaced by the quantities $\Gamma_n V$, which determine the amplitude factors of the axial, tangential and oblique modes, as also used in [14]. In this work, different from Kuttruff, we were interested in a source with a flat spectrum and therefore the ω dependency was skipped in the part before the summation term. Using this analytical solution, the reverberation time of each mode could be individually manipulated. The targeted T_{60} in Eq. 2 was different between the low and high frequency part of the spectrum: (a) A constant $T_{60,HF} = T_{opt}$ was set for frequencies higher than the 175 Hz crossover frequency. (b) For the modes below the crossover frequency, different values of $T_{60,LF}$ were tested: 0.48 s, 0.56 s, 0.64 s and 0.72 s. The spectrum was calculated for modes lower than approximately 12 kHz and in a next step, the inverse Fourier transform of Eq.3 was calculated numerically and the resulting impulse response was convolved with different anechoic sounds. Three different kinds of stimuli were generated: *1 tone arco* (1TA), *2 tones arco* (2TA) and *kick drum* (KD). Sound 1TA consisted of a single C_3 note played on a double bass in an arco style (with bow), 2TA consisted of two successive notes C_3 and D_3 played in arco style and KD was a simple kick drum pattern of three successive hits. The first two sounds were generated using a sampled instrument, the drum sound was synthesized. The spectrograms of the anechoic stimuli are shown in Figure 2.

In order to achieve a valuable adequate listening test comparison between different variants of $T_{60,LF}$, which differed not only in modal decay times but also in amplitude, the HF regions of the convolved sounds to be compared were energetically aligned. The adopted procedure is schematically shown in the Figure 3. After convolution, a high-pass Butterworth filter at 200 Hz was applied in order to keep only the HF region, which was not affected by differences in low frequency modal decay times. For that region, a target peak level was chosen and a corresponding normalization factor was calculated. Next, this normalization factor was applied to the convolved sound. The target peak was chosen to -24 to -26 dBFS for all

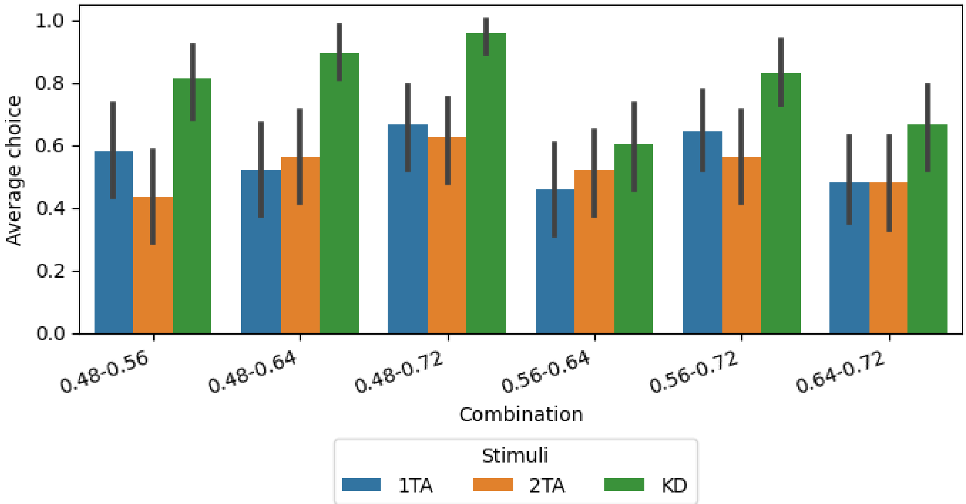

Figure 4. Average results and their standard deviation for each of the tested combinations and stimuli (1TA – one tone arco, 2TA – two tones arco, KD – kick drum).

stimuli, which corresponded to RMS SPL values of approximately 69 dB for both *arco* stimuli and 81 dB for *kick drum* stimuli.

As mentioned already, comparative listening tests were performed on the generated convolved stimuli in order to get insight in the audibility of differences in low frequency modal decay times. The ABX type listening test protocol was used: three stimuli **A**, **B** and **Reference (X)** were presented to the listener, who was then deciding, which of the **A** or **B** was the same as the **X**. The first playback was strictly in the order **A-B-X**, but after that, the listener was allowed to replay any of the material. Each stimuli pair was presented in four combinations (A-B-A, A-B-B, B-A-A, B-A-B), in a random order. For each stimulus, the number of combinations N presented to the listener can be calculated as:

$$N = C(5,2) \cdot n = \frac{4!}{2!(4-2)!} \cdot 4 = 24 \qquad (4)$$

where C is the number of unique pairs of different $T_{60,LF}$ variants and n is the number of repetitions of each pair within that test. Since the focus of our interest was on the lower frequencies, a 20dB limit for only frequencies $\leq$ 1000 Hz was chosen for passing the pure-tone audiometry test.

3. RESULTS

The test was performed in the semi-anechoic chamber in the acoustic laboratory at KU leuven using the HEAD acoustics® HPS IV listening unit. Since this test represents a preliminary study on the topic, also the responses of two co-authors were part of the evaluation. A total of 13 listeners participated on the test, including several active musicians. For a better understanding, some terms should be defined. *Combination* here refers to the unique pair of RIRs (with the specific $T_{60,LF}$ values) used for convolution with the anechoic stimuli. *Stimuli*, on the other hand, refers to the original sound before convolution. The difference in the LF decay times can be calculated as $\Delta T_{60,LF} = |T_{60,LF,1} - T_{60,LF,2}|$, where $T_{60,LF,1}$ and $T_{60,LF,2}$ are the LF mode decay times presented in a combination.

Figure 4 shows the average results for each combination together with their standard deviation. There are evident differences between stimuli (ANOVA repeated measures shows a significant difference between the stimuli with $p = 7.8 \cdot 10^{-7}$). The further pairwise T-test (with a Bonferroni correction applied) reveals (Table 1) that significant differences can be found between the arco (1TA, 2TA) and drum stimuli (KD), but not between the two arco stimuli.

In view of that, the results were further analysed separately for arco stimuli (TA) and kick drum (KD). In

Table 1. Pairwise T-test results for different stimuli (1TA – one tone arco, 2TA – two tones arco, KD – kick drum).

Stimuli 1	Stimuli 2	p_{corr}
1TA	2TA	1
1TA	KD	**<0.001**
2TA	KD	**<0.001**

case of TA, no significant differences in performance were found between different combinations. Furthermore, the t-test showed that the performance was higher than 0.5 only for combinations 0.48-0.72 (p=0.002) and 0.56-0.72 (p=0.020). In the case of KD stimuli, all the combinations showed performance higher than 0.5, with one exception of a combination 0.56-0.64, for which only a trend was present (p=0.075). Also in case of KD, several significant differences in performance were found between different combinations, but none of them between combinations with a shared lower value of $T_{60,LF}$ or with a constant $\Delta T_{60,LF}$, as can be seen in Table 2.

Table 2. Pairwise T-test for different combinations, only for kick drum stimuli (KD).

Combination 1	Combination 2	p_{corr} (KD)
0.48-0.64	0.56-0.64	**0.036**
0.48-0.64	0.64-0.72	*0.071*
0.48-0.72	0.56-0.64	**0.049**
0.48-0.72	0.64-0.72	**0.018**
0.56-0.72	0.64-0.72	*0.069*

4. DISCUSSION AND CONCLUSION

Interestingly, the comments, which were provided by the participants after the test, contained a wide range of facets, ranging from "timbral changes", "muddiness", "clarity", "differences in reverberation", "masking towards higher frequencies", or even "slight changes in pitch". Even though the focus was widely spread, it can be concluded from the listening test results that the kick drum sound (and supposingly other highly transient sounds) serves better than the steady tonal signal when searching for the differences in room mode damping. Hence, temporal differences seem easier to evaluate than spectral differences. This finding is in agreement with previous studies and particularly with Olive et al. [5], who found that detecting high-Q resonances in case of pulses as stimuli was done more adequately than when broadband noise signals were used as stimulus. The upcoming research will further focus on the audible differences in room mode damping in a semi-reverberant environment. Attention will be paid to the different musical instrument characteristics as well as to more narrow band changes in room reverberation.

5. ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon Europe research & innovation programme under the HORIZON-MSCA-2021-DN-01 grant agreement No. 101072598 – "ActaReBuild". The contribution was produced also with support for specific university research at Brno University of Technology, FAST-J-22-7880 (2022) and FAST-J-23-8284 (2023). The main author's mobility enabling for this research was supported with "Rozvojový projekt 1.17 Mobility studentů VUT" and "Program pro podporu strategického řízení 2023".

6. REFERENCES

[1] J. Rindel, "New Norwegian standard on the acoustics of rooms for music rehearsal and performance," (Krakow), Sept. 2014.

[2] J. Olsen and J. Rindel, "The acoustics of rooms for music rehearsal and performance—The Norwegian approach," *The Journal of the Acoustical Society of America*, vol. 141, pp. 3597–3597, May 2017.

[3] "ISO 23591:2021 - Acoustic quality criteria for music rehearsal rooms and spaces," 2021.

[4] F. E. Toole and S. E. Olive, "The Modification of Timbre by Resonances: Perception and measurement," *AES: Journal of the Audio Engineering Society*, vol. 36, no. 3, pp. 122–142, 1988.

[5] S. E. Olive, P. L. Schuck, J. G. Ryan, S. L. Sally, and M. E. Bonneville, "The Detection Thresholds of Resonances at Low Frequencies," *Journal of the Audio Engineering Society*, vol. 45, pp. 116–128, Mar. 1997.

[6] L. Kritly, K. Jambrosić, M. Rychtáriková, and C. Glorieux, "Audibility of coincidence dips and structural resonance dips in acoustic isolation curves," in *Proceedings of the ATF conference*, (Leuven, Belgium), Oct. 2018.

[7] L. Kritly, C. Glorieux, V. Chmelík, M. Rychtarikova, and Y. Sluyts, "Audibility of Spectral Dips and Peaks in Broadband Noise," (Aachen, Germany), Deutsche Gesellschaft für Akustik, 2019. Number: RWTH-CONV-239064.

[8] L. Kritly, L. Zelem, V. Chmelík, C. Glorieux, and M. Rychtarikova, "Preliminary study on the estimation of just noticeable differences of spectral dips and peaks by adaptative method," in *Proceedings of the International Conference on Noise and Vibration Engineering (ISMA2020) and the International Conference on Uncertainty in Structural Dynamics (USD2020)*, Department of Mechanical Engineering, KU Leuven, Nov. 2020. ISSN: 9781713827054.

[9] M. R. Avis, B. M. Fazenda, and W. J. Davies, "Thresholds of Detection for Changes to the Q Factor of Low-Frequency Modes in Listening Environments," *Journal of the Audio Engineering Society*, vol. 55, pp. 611–622, July 2007.

[10] B. M. Fazenda, M. Stephenson, and A. Goldberg, "Perceptual thresholds for the effects of room modes as a function of modal decay," *The Journal of the Acoustical Society of America*, vol. 137, pp. 1088–1098, Mar. 2015.

[11] L. Rizzi, F. Ascari, G. Ghelfi, and M. Ferroni, "Perception of Low Frequency Transient Acoustic Phenomena in Small Rooms for Music," Audio Engineering Society, May 2016.

[12] M. Kleiner and J. Tichy, *Acoustics of Small Rooms*. Boca Raton: CRC Press, 1st edition ed., Apr. 2014.

[13] H. Kuttruff, *Room acoustics*. London, [England] ; New York, NY: Spon Press, 4th ed ed., 2000.

[14] F. P. Mechel, *Formulas of acoustics*, vol. 57. New York, NY: Springer, 2nd ed ed., 2009. ISSN: 07362501.