

The Influence of the Measurement Chain on the Impulse Response of a Reverberation Room and Its Application Listening Tests

P. MAŁECKI*, M. ZASTAWNIK, J. WICIAK AND T. KAMISIŃSKI

Department of Mechanics and Vibroacoustics, AGH University of Science and Technology

al. A. Mickiewicza 30, 30-059 Kraków, Poland

There are many objective parameters of the room impulse response which we can calculate that describe its acoustic field. The room impulse response can also be used to provide listening tests using the convolution technique. The aim of this work is to consider the influence of the impulse response measurement chain on objective numerical parameters and on the subjective perception of listeners. This article shows the impulse response measurement results obtained in the reverberation room using not only commonly known measurement chain elements but other elements, too. Standard components such as an omni-directional source and measurement microphone, as well as a consumer active loudspeaker and various audio recording stereo microphone pairs were used. The results are compared taking into consideration calculated numerical parameters and the application of calculated impulse responses to subjective listening tests. To determine the applicability of impulse responses to psychoacoustic listening tests, a preliminary listening test using headphones was carried out on a selected group of listeners. The audio test material consisted of samples of choral music and speech recorded in an anechoic room and convolved with previously analyzed impulse responses. This article shows how big the differences of subjective listener feelings are for different acoustic transducers.

PACS: 43.55.Hy, 43.55.Mc, 43.66.Lj

1. Introduction

Nowadays there are reasonably accurate methods to measure, design and predict sound fields. It depends on the acoustic engineer how a given interior will sound, but there is very often the question of what the desired values of reverberation time and other room acoustics parameters are. There are many preferred values, as they are known, for specific target use of a room and its volume and these are designed for a specific kind of music, speech or both. There are also several approaches to evaluating the most desired acoustic properties. Beranek's *Concert Halls and Opera Houses* [1] well-known book contains the measurement results of many famous interiors along with the opinions of musicians and other experts. Ando's *Concert Hall Acoustics* [2] contains the results of acoustic experts' surveys carried out in different concert halls. The opinions were formulated on the basis of the same music samples which had been previously recorded in an anechoic room and played back in these halls by means of a loudspeaker. There are also methods which do not require such effort and can be applied to other structures. Tests using the room impulse response and convolution technique are now common place [3, 4], but still there

is no unified or normalized methodology in this kind of research. The aim of this work is to consider the influence of the impulse response (IR) measurement chain on objective numerical parameters and on the subjective perception of listeners.

2. The concept of the experiment

The chief aim of this study was to record impulse responses for the same room, place and time, using the same measurement signal, software and AD/DA converter, but with a different sound source and receiver, and then to calculate and compare numerous estimators as well as perform listening tests using the measured IRs convolved with anechoic samples of music and speech.

2.1. Experiment setup

The measurements were performed in a medium sized Catholic church in Czarnochowice near Cracow (Poland). Built on a rectangular plan (Fig. 1), the church is fairly modern and has a volume of 1700 m³ and a floor area of 290 m². It is furnished with wooden pews. There are not many sculptures or other ornaments. The walls and ceiling are finished with tile plaster and the floor is made of marble.

The measurements were performed using two different sound sources:

* corresponding author; e-mail: pawel.malecki@agh.edu.pl

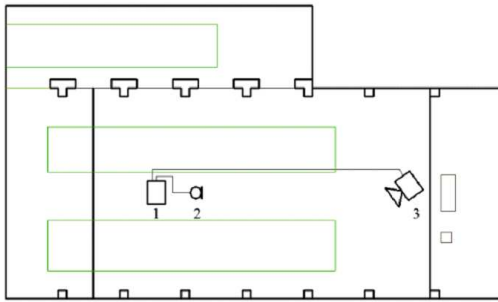


Fig. 1. Church plan with measurement points (1 — PC+RME Fireface800+EASERA software, 2 — microphones, 3 — loudspeakers).

- 12 ways, 8" omni-directional loudspeaker set (W),
- JBL EON 515 active loudspeaker (JBL).

The abbreviations in brackets are used in the figures and graphs that follow. The first source meets ISO3882 requirements [5] so it is commonly used to measure reverberation time and/or other acoustic parameters. On the other hand, its construction and the small diameter of the loudspeaker membrane produce a non-flat frequency response. The commercial JBL loudspeaker has a much flatter frequency response but its characteristic is not omni-directional over the entire frequency band.

These sets of microphones were also used:

- 2× B&K 4134 in AB stereo setup (B&K),
- Behringer EMC8000 (BEHR),
- 2× Rode NT5 in ORTF stereo setup (ORTF),
- 2× Soundman OKM II (BINA).

As with loudspeakers, typical measurement microphones were used. The first pair were high quality B&K omni-directional with a very flat frequency response and high sensitivity. The second microphone listed above was the best-known cheapest measurement microphone. The pair of Rode NT5 microphones is commonly used by audio professionals in stereo recordings. The Soundman OKM II is a pair of binaural microphones. The recordings were made on a human head instead of a dummy head.

The measurements were performed with EASERA software using log sweep-sine. The AD/DA and the microphone preamplifier was RME Fireface800.

2.2. Listening tests

The second part of the experiment involved the human factor. Subjective feeling of space or timbre is very individual and depends on many circumstances. The main goal of this part was to find out the listeners' opinion about the recorded IRs convolved with anechoic music samples, and to ascertain how big the differences are

between listeners opinions and which IRs are preferred. Samples of choir music and speech were recorded in an anechoic room of the AGH University. There are several methods of estimating human preferences. An ABX listening test [6] requires a reference signal which is not available for this kind of material. For very small differences between signals, the paired comparison method can be used [7]. There are also tests where listeners assess every single sample and there are two options for this test: with or without a grade scale. What is known as the absolute magnitude test does not include any fixed scale. Listeners are free to mark particular samples in any way they feel in this kind of test. Absolute magnitude tests have better "dynamics" of results compared to traditional scale tests [8] because there is no possibility of marking all uncertain samples as "average" or "medium". Because of these arguments the absolute magnitude tests are performed in order to estimate the prepared samples. The test was performed using closed studio headphones: Beyerdynamic DT770 PRO. All listeners received a test sheet and were instructed how to fill it in. There were two samples of choir music (14 and 28 s long) differing in complexity, tempo and dynamics and one sample of male speech (11 s). This was a double blind test with the following loudspeakers and microphones combinations played in a random order:

- (a) W + B&K (mono);
- (b) W + B&K;
- (c) W + BEHR;
- (d) W + ORTF;
- (e) W + BINA;
- (f) JBL + B&K (mono);
- (g) JBL + B&K;
- (h) JBL + BEHR;
- (i) JBL + ORTF;
- (j) JBL + BINA;
- (k) EASE (IR-generated using the auralization module of the EASE software, based on a geometrical model of the church where IRs measurements were performed).

The abbreviations are explained in Sect. 2.1. The test was performed on a group of 20 listeners of different acoustic experience. There were professional musicians and sound engineers as well as amateurs and average music listeners. There were also differences in religion (there were members of the Roman Catholic, Orthodox, and Evangelical Churches). This was intentional and important as Orthodox Church choral music samples were used.

3. Comparison of the calculated numerical acoustic parameters

The main and most important parameter is reverberation time. Figure 2a and b shows that the differences between particular combinations of equipment are not significant. There are some meaningful differences at low frequencies for binaural (BINA) microphones and this is caused by their low sensitivity.

The most popular acoustic parameters were calculated and showed in Table, such as clarity index for music (C80) and speech (D50), center time (T_s), support (ST)

and rapid speech transmission index (RaSTI). All of this parameters are well described in [5, 9].

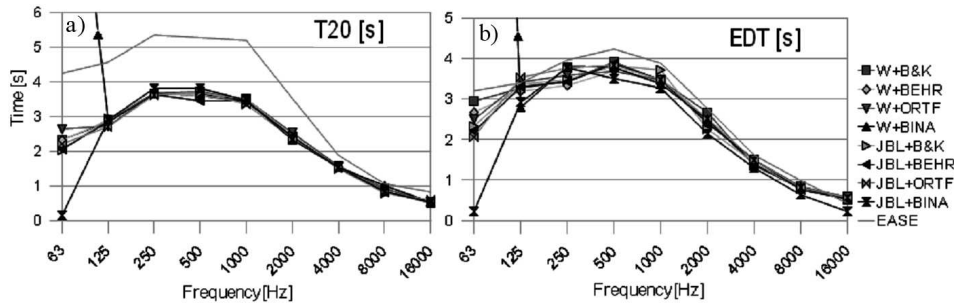


Fig. 2. (a) T20 and (b) EDT reverberation frequency characteristic of all IRs combinations.

Numerical acoustic parameters calculated for IRs combinations.

TABLE

	W+B&K	W+BEHR	W+ORTF	W+BINA	JBL+B&K	JBL+BEHR	JBL+ORTF	JBL+BINA	EASE
T20 [s]	3.51	3.54	3.61	3.55	3.59	3.45	3.54	3.64	5.23
EDT [s]	3.69	3.59	3.57	3.58	3.77	3.67	3.65	3.39	4.06
C80 [dB]	-4.3	-4.5	-3.5	-3.8	-4.6	-4.0	-3.6	-2.7	-4.1
D50 [-]	0.19	0.17	0.21	0.19	0.18	0.20	0.22	0.27	0.19
T_s [ms]	246	244	235	235	246	247	235	216	270
ST early [dB]	7.9	7.1	6.5	10.6	6.3	4.7	4.2	8.1	5.0
RaSTI [-]	0.35	0.35	0.36	0.35	0.34	0.37	0.38	0.40	0.34

The most relevant results were obtained using an omnidirectional loudspeaker and a high quality measuring microphone because this combination fully follows ISO3882 requirements [5]. Other measurement results do not differ by more than 10% for reverberation time. Reverberation parameters calculated from the IR of the geometrical model (Sect. 2.1, k) are much higher in this case which is the result of adjusting the model using the Sabine formula instead of the ray-tracing technique. On the other hand, most of the parameters calculated from artificial IR (except room temperature (RT)) are very similar to the reference measurement.

The largest variation in results is for the magnitude spectrum of particular IRs (Fig. 3). This is the result of the non-flat frequency response of loudspeakers. A review of the specifications provided by the manufacturers of the electro-acoustic transducers used here shows that the microphone characteristics are very flat. Therefore they do not have any particular influence on the magnitude of the recorded IRs.

4. Listening test results

It is difficult to interpret the results of the listening tests definitively. The graphs in Fig. 4 indicate some

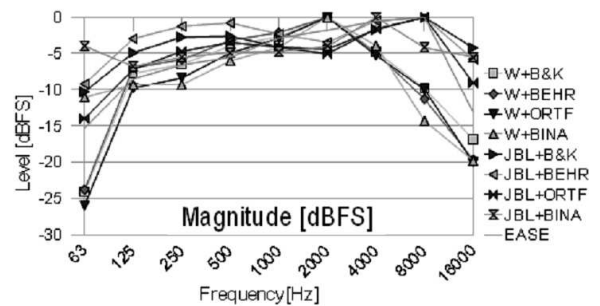


Fig. 3. Magnitude spectrum of particular IRs.

differences but mostly between loudspeakers, which was easy to predict. It is significant that standard deviations are so large. It is notable that subjective sense of space and timbre are strongly correlated in all cases. As stated in Sect. 2.2, the order of hardware combinations was random for every music sample so it is important to emphasize the similarity among the envelope shapes in all graphs. However, there are no statistical grounds to conclude which selected combinations are mostly preferred by listeners, but the conclusion can be drawn that

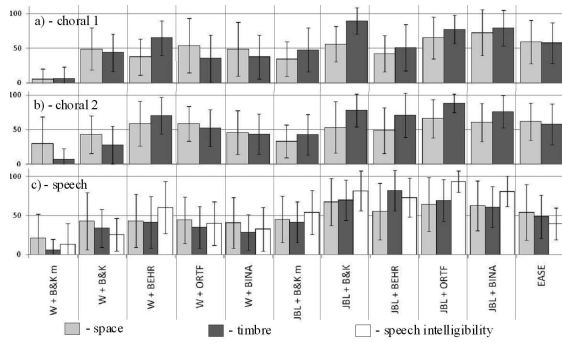


Fig. 4. Listeners percentage preferences and standard deviation in results for particular samples and combinations of equipment: (a) choral 1; (b) choral 2; (c) speech.

there are meaningful differences and that they are not totally random. It can be seen that the first column (which follows the ISO 3382 standard [5]) was marked the worst of all combinations and that the difference is quite large.

Listeners were also questioned regarding speech intelligibility and the results were also correlated with other survey parameters (sense of space and timbre). Furthermore, these results were the most statistically clear of all the results, the deviation of particular results was not so large and the differences in preferences were the biggest.

5. Conclusion and future work

The most important conclusion from this experiment is that the IR measurement chain (in this case the electro-acoustic transducers) has a significant impact on the subjective sense of space and timbre in contrast to numerical parameters which are not as sensitive to these kinds of changes. The results obtained do not answer the question regarding which combination gives the most realistic

convolution and concerns impressions in real conditions (in the church). This problem requires further experiment or an additional type of experiment. It can be noticed on the base of Fig. 4 that the most preferable are stereo microphones pairs so the spatial techniques should be the aim of future research as well as considering the B-format and other surround techniques for room acoustic applications.

There are no international or universal standards to measure impulse response in listening tests so there is a need to standardize this test procedure.

References

- [1] L. Beranek, *Concert and Opera Halls*, Acoustical Society of America, 1996.
- [2] Y. Ando, *Concert Hall Acoustics*, Springer Verlag, Berlin 1985.
- [3] M.R. Schroeder, D. Gottlob, K.F. Siebrasse, *J. Acoust. Soc. Am.* **56**, 1195 (1974).
- [4] F. Martellotta, in: *Proc. 118th AES Convention, Barcelona (Spain)*, 2005, Paper 6460.
- [5] ISO 3382: *Acoustics — Measurement of the reverberation time of rooms with reference to other acoustical parameters*, 2001.
- [6] ITU-R BS.1284: *General methods for the subjective assessment of sound quality*, 2003.
- [7] ITU-R BS.1116: *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*, 1997.
- [8] A. Miskiewicz, A. Rakowski, in: *Proc. 16th ICSV, Kraków (Poland)*, 2009, Paper 606.
- [9] W. Ahnert, W. Schmidt, *Fundamentals to perform acoustical measurements*, Appendix to EASERA, 2005.