inter-noise 2008 From Silence to Harmony
37th International Congress and Exposition on Noise Control Engineering
26-29 October 2008·Shanghai·China

The Effect of Room Acoustics on the Perceived Acoustics of Reproduced Sound

C.C.J.M. Hak[a], R.H.C. Wenmaekers[b]
[a] Eindhoven University of Technology, Den Dolech 2, BWK/BPS, 5600 MB Eindhoven, Netherlands
[b] LeVeL Acoustics BV, De Rondom 10, 5612 AP Eindhoven, Netherlands

ABSTRACT

Recordings of music, speech or other sounds have acoustical properties, such as reverberation, definition and clarity. For educational, engineering or entertainment purposes these recordings are often played back through loudspeakers in listening rooms, such as classrooms, recording studios or cinemas. The acoustical properties of a listening room affect the perceived acoustics of the recording (room in room acoustics), which may make an accurate reproduction or demonstration of the recording's acoustical properties impossible. To find the practical impact of room in room acoustics, combinations have been investigated of room acoustic impulse responses using convolution techniques. The resulting change in an acoustical property is assumed to be accurate if it does not exceed the JND (Just Noticeable Difference). It is found that to accurately reproduce a steady sound energy decay rate (related to the reverberation time), the playback room is allowed to have as much as half this decay rate, under diffuse sound field conditions. For energy modulations (related to speech intelligibility) this decay rate should be more than four times smaller. Finally, initial energy ratios (related to definition and clarity) require auditive judgement in the direct soundfield.

1 INTRODUCTION

It often happens to lecturers in sound/acoustics or speakers at a congress on room acoustics that part of the planned sound effects cannot be demonstrated well because of unsuitable acoustics of the presentation room [lit 11]. Indeed, room acoustic sound recordings cannot be simply demonstrated in every room. From experience it is clear that a recorded short reverberation time can only be heard in a room having a reverberation time shorter than the one in which the recording was made. If all acoustic demonstrations would be given through professional headphones or nearfield monitors, sound recordings would sound as they are meant to. The smallest details and the finest nuances with regard to colouring, definition and stereo image could then be judged and criticized. However, usually the listening or playback room in combination with the used sound system affects the recorded acoustics. This happens in class rooms, congress halls, cinemas and even in sound recording studios.

Using formerly measured impulse responses and the convolution function in the acoustic measurement program DIRAC, a first step is made in investigating the impact of the reproduction room acoustics on recorded acoustics. In this case the impact on reverberation, speech intelligibility and clarity has been investigated. From the results, criteria for the

[a] Email address: c.c.j.m.hak@tue.nl
[b] Email address: r.h.c.wenmaekers@tue.nl

listening room's reverberation time have been derived for the proper playback of each of these parameters, starting from a diffuse soundfield and the JND (Just Noticable Difference) as allowable error.

2 PROCEDURE

From an extensive set of measured impulse responses, a selection is made (Chapter 3). From each pair (h_1, h_2) out of this selection the first is considered as a recorded impulse response and the other as a listening room impulse response. Each pair components are mutually convolved (see Chapter 4) to obtain the impulse response h_{12}, heard when playing back the recorded impulse response in the listening room. h_{12}, thus representing the h_1 affected by h_2, is then compared with h_1, with respect to the reverberation time T [lit 4], the Modulation Transfer Index MTI [lit 5,6] and the Clarity C80 [lit 4]).

3 IMPULSE RESPONSES AND MEASUREMENTS

The subset selection from the original set of impulse responses is based on the measurement quality, the measurement equipment, the rooms in which the measurements are performed and the positions of the sound source and the measurement microphone. Finally 11 impulse responses have been selected from which the 500, 1000 and 2000 Hz octave bands have been used.

For each selected impulse response the same measurement equipment has been used, consisting of the following components:

- *software:* DIRAC (B&K/Acoustics Engineering - Type BZ5449) running on a laptop PC;
- *input/output:* USB audio device (Acoustics Engineering - Triton);
- *power amplifier:* (Acoustics Engineering - Amphion);
- *sound source:* omnidirectional (B&K - Type 4292);
- *microphone:* omnidirectional, sound level meter (Rion - NL21);
- *signal:* synchronous or asynchronous ([lit 9,10])

All impulse responses are obtained from diffuse sound field measurements using deconvolution techniques [lit 1, 7] with MLS and e-sweeps [lit 3], resulting in INR values > 50 dB [lit 8]. Some properties of the selected impulse responses are shown in Table 1. The values are the averages over the 500, 1000 and 2000 Hz octave bands [lit 4].

Table 1: Properties of used room impulse responses.

Room	INR [dB]	T20 [s]*	C80 [dB]*	RASTI [-]**
Ice chapel	62	0,20	25,63	0,89
Lecture room	66	0,63	7,23	0,70
Auditorium 1	59	0,85	4,38	0,61
Conference hall	56	1,09	2,45	0,61
Measurement room 1	80	1,11	2,05	0,54
Chamber music hall	63	1,21	4,93	0,58
Concert hall 1	60	1,28	1,74	0,56
Concert hall 2	54	2,00	2,08	0,50
Concert hall 3	53	2,92	-2,34	0,37
Auditorium 2	68	4,90	-6,23	0,27
Measurement room 2	73	5,30	-5,84	0,30

* Average over 500-1000 Hz
** Calculated from MTI 500 and MTI 2000 Hz [lit 2, 5, 6]

4 CONVOLUTION

The convolution $\boldsymbol{y}$ of signal $\boldsymbol{s}$ and system impulse response $\boldsymbol{h}$ is written as $\boldsymbol{s} * \boldsymbol{h}$ and defined as:

$$\boldsymbol{y}(\boldsymbol{t}) = \boldsymbol{s}(\boldsymbol{t}) * \boldsymbol{h}(\boldsymbol{t}) \tag{1}$$

or

$$\boldsymbol{y}(\boldsymbol{t}) = (\boldsymbol{s} * \boldsymbol{h})(\boldsymbol{t}) = \int_{-\infty}^{\infty} \boldsymbol{s}(\boldsymbol{t}) \cdot \boldsymbol{h}(\boldsymbol{t} - \boldsymbol{\tau}) d\boldsymbol{\tau} \tag{2}$$

In words: the convolution is defined as the integral of the product of two functions s and h after one is reversed and shifted. From a room acoustical point of view ***s(t)*** is a sound that is recorded in an anechoic room (dry recording) and played back in a standard room, ***h(t)*** the impulse response of the standard, more or less reverberant room and ***y(t)*** the convolved sound as it is heared in that standard room. Therefore, an impulse, for instance a hand clap, recorded in an anechoic room, played back in a reverberant room, is heared as an impulse response of that reverberant room. A recorded impulse in the reverberant room that is played back in the anechoic room is again heared as the impulse response of the reverberant room. In both cases the derived room acoustic parameter values will be the same. When both the recording room and the playback (listening) room are reverberant, smoothing of the sound occurs. Therefore, in some cases it is impossible to judge the original recordings in detail. The room acoustics in the sound recording that we want to demonstrate will be affected by the acoustics of the listening room. With a double convolution by which an impulse response from one room is convolved with a dry recording and afterwards the result is convolved with the impulse response of another room, it is possible to hear how a recording, made in a reverberant room, sounds when played in another reverberant room. The result is usually an unwanted smoothed sound signal. By using a pure impulse (Dirac delta function) instead of a normal sound signal to be convolved with both room impulse responses (eq 3 and 4) we can examine what one room does with the other concerning the values for the room acoustic parameters (eq 5). So it is possible to derive a 'room in room' acoustic parameter value from the smoothed impulse response (Figure 1).

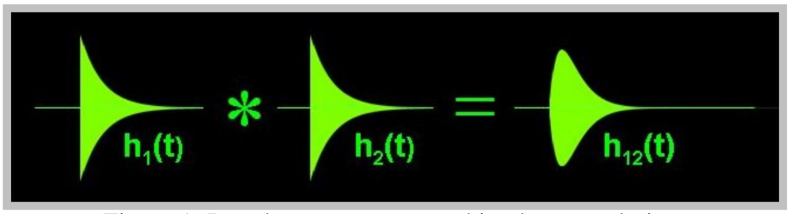

Figure 1: Impulse response smoothing by convolution.

Mathematically:

$$h(t) * \delta(t) = \delta(t) * h(t) = h(t) \tag{3}$$

Where:
h(t) = room impulse response
δ(t) = Dirac delta function (ideal impulse)

$$h_{12}(t) = \delta(t) * h_1(t) * h_2(t) = h_1(t) * h_2(t) \tag{4}$$

Where:
$h_{12}(t)$ = 'total' impulse response room1 ∗ room2
$h_1(t)$ = impulse response room 1
$h_2(t)$ = impulse response room 2

Substituting equation (4) into equation (1) results in:

$$y_{12}(t) = s(t) * h_{12}(t) \tag{5}$$

Where:
$y_{12}(t)$ = convolution of a random sound signal with the 'total' impulse response
s(t) = random sound signal

5 RESULTS

In Figure 2 through 4 the results of the convolutions are depicted as scatter diagrams. Each graph shows the difference between 2 values of a parameter, one calculated from h_{12} and one from h_1. Using the reverberation time as the base parameter, on the x-as the ratio $T20(h_1)/T20(h_2)$ of the reverberation time calculated from h_1 (= $T_{recorded\ room}$) and the reverberation time calculated from h_2 (= $T_{listening\ room}$) is given. For symmetry reasons ($h_{12} = h_1 {*} h_2 = h_2 {*} h_1$), only impulse response pairs with $T20(h_1) > T20(h_2)$ have been depicted.

The differences are calculated for three acoustical parameters. One is a decay related parameter T20, which is the reverberation time calculated over a decay range from -5 to -25 dB. The second one is a modulation related parameter MTI, a value between 0 and 1, used to calculate the speech intelligibility. The third is an energy distribution related parameter C80, which is the Clarity and determined as the ratio of the impulse response energy of the first 80 ms to the energy after 80 ms.

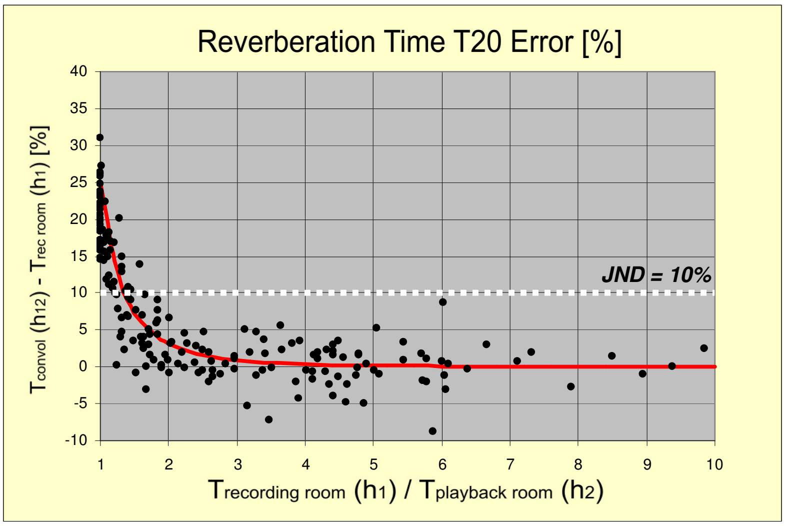

Figure 2: Difference between $T_{playback}$ en $T_{recording}$.

The scatter plot of the reverberation time T20 in Figure 2 shows that for the selected impulse responses (Table 1) the variation of the percentual difference between $T(h_{12})$ and $T(h_1)$ lies within a band of +/- 10% around the trend line. Starting from a JND (Just Noticeable Difference) of 10%, a diffuse field recording, and the reverberation time as the only judgement criterion, we can conclude that for an accurate demonstration a listening room should have a reverberation time below half that of the recording room.

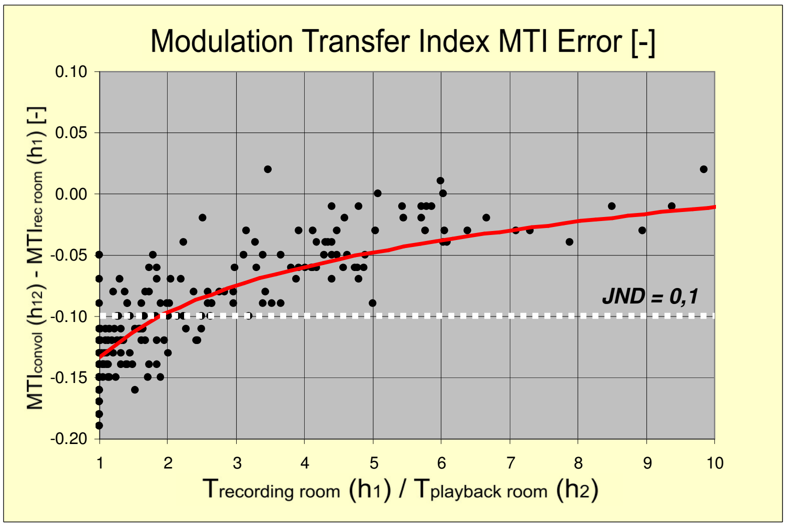

Figure 3: Difference between $MTI_{playback}$ en $MTI_{recording}$.

The scatter plot of the modulation transfer index MTI in Figure 3 shows that for the used impulse responses (Table 1) the variation of the differences between $MTI(h_{12})$ and $MTI(h_1)$ lies within a band of +/- 0.05 around the trend line. Speech intelligibility experiments and demonstrations require a playback or listening room with a reverberation time at least 4 times shorter than that of the recorded impulse response, using a JND for the MTI of 0,1.

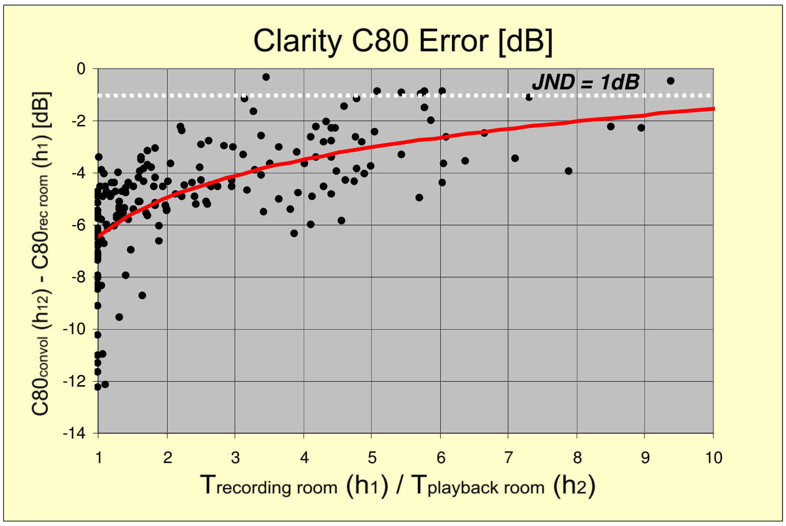

Figure 4: Difference between $C80_{playback}$ en $C80_{recording}$.

The scatter plot of the Clarity C80 in Figure 4 shows that for the used impulse responses (Table 1) the variation of the difference between $C80(h_{12})$ and $C80(h_1)$ lies within a band of +/- 3 dB around the trend line. When it is important to demonstrate or judge the details of sound definition or brightness, using a JND of 1 dB for the Clarity, you have to use a playback or listening room with a reverberation time of more than a factor of 10 lower than the reverberation time of the recorded impulse response. That means for most clarity or definition demonstrations you need a room with the properties of a sound recording studio.

6 CONCLUSIONS

- To accurately reproduce a steady sound energy decay rate (related to the reverberation time), the playback room is allowed to have as much as half this decay rate, under diffuse soundfield conditions.
- For energy modulations (related to speech intelligibility) the decay rate of the playback room, under diffuse soundfield conditions, should be at least four times smaller than that of the recording room.
- Initial energy ratios (related to definition and clarity) require auditive judgement in a direct soundfield.

7 ACKNOWLEDGEMENTS

The authors wish to thank all involved students of the Eindhoven University of Technology, Department Architecture, Building and Planning (Unit Building Physics and Systems). They provided good quality impulse responses, during their practical work for the course ‘Room Acoustics’. The authors also wish to thank Jan Hak (Acoustics Engineering, Boxmeer) and Renz van Luxemburg (Level Acoustics, Eindhoven) for reviewing this paper.

8 REFERENCES

[1] M.R. Schroeder, “Integrated-impulse method for measuring sound decay without using impulses”, Journal of the Acoustical Society of America, **66**, 497-500 (1979).

[2] M.R. Schroeder, “Modulation transfer functions: definition and measurement”, Acustica, **49**, 179-182 (1981).

[3] S. Müller, P. Massarani, “Transfer-function measurements with sweeps*”,* Journal of the Audio Engineering Society, **49**, 443-471, (2001).

[4] *Acoustics - Measurement of room acoustic parameters - Part 1: Performance rooms*, International Standard ISO/DIS 3382-1 Draft: 2006 (International Organization for Standardization, Geneva, Switzerland, 2006).

[5] T. Houtgast, H.J.M. Steeneken, “The modulation transfer function in room acoustics as a predictor of speech intelligibility”, Acustica, **28,** 66-73 (1973).

[6] *Sound System Equipment – Part 16: Objective rating of speech intelligibility by speech transmission index,* International Standard IEC 60268-16: 2003 (International Organization for Standardization, Geneva, Switzerland, 2003).

[7] *Acoustics - Application of new measurement methods in building and room acoustics*, International Standard ISO 18233: 2006 (International Organization for Standardization, Geneva, Switzerland, 2006).

[8] C.C.J.M. Hak, J.P.M. Hak, R.H.C. Wenmaekers, “INR as an Estimator for the Decay Range of Room Acoustic Impulse Responses”, *AES Convention Amsterdam* (2008).

[9] C.C.J.M. Hak, J.P.M. Hak, “Effect of Stimulus Speed Error on Measured Room Acoustic Parameters”, *International Congress on Acoustics Madrid* (2007).

[10] C.C.J.M. Hak, J.S. Vertegaal, “MP3 Stimuli in Room Acoustics*”, International Congress on Acoustics Madrid* (2007).

[11] R.H.C. Wenmaekers, C.C.J.M. Hak, H.J. Martin, L.C.J. van Luxemburg, “Air Absorption Error in Room Acoustical Modelling”, *Proceedings of Acoustics 08 Paris* (2008).