ACOUSTIC REVERBERATION: A BASIS FOR SOUND RECORDING IN MODERATELY ANECHOIC ROOMS

Dissertation

by

Philip Baron

submitted in fulfilment of the degree

Master in Engineering

to

Prof. M.Case

Department of Power and Control Engineering
School of Electrical and Electronic Engineering
University of Johannesburg
Doornfontein

30 August 2009
Johannesburg

ABSTRACT

Acoustic reverberation is put under the spotlight. A review of the theory was conducted followed by a look into digital artificial reverberation. Measurement methodology is presented including a review of the recently published ISO standard pertaining to reverberation. Experimental testing was conducted for four acoustically different environments with one of them almost completely anechoic. The reverberation characteristic of these four environments were measured and analysed according to the relevant ISO standards. The results were then used in a further study of digital artificial reverberation applied to impulse and vocal sounds. The anechoic sounds were artificially reverberated using *Cool Edit Pro* software to mimic the sound obtained that had natural reverberation present. The focus was on the RT as well as the EDT of the decay slope. The artificial method of applying reverberation was evaluated using two methodologies, firstly objective methods relying on mathematics; secondly, by subjective personal evaluations using a statistical analysis of a listening test questionnaire. Both the objective and subjective results confirmed that digital artificial reverberation methods could be applied successfully to impulse sounds and vocals. The results provide a basis for the motivation of computerised methods in the studio recording process especially for rooms that are moderately anechoic.

Declaration

I, Philip Baron, hereby declare that this research dissertation, which I submitted to the University of Johannesburg, is my own work and has not been previously submitted by me to any other institution.

P. Baron

Date: August 2009

Acknowledgements

A special thanks to my supportive family.

Thanks to:
University of Johannesburg, Lars Henrik Morset and WinMLS, and *Rooikraal Landgoed* Farm

TABLE OF CONTENTS

List of Figures

List of Tables

Notation

α	absorption coefficient
$\overline{\alpha}_E$	area-averaged random-incidence energy absorption coefficient
B	bandwidth (Hz)
c	velocity of sound is air of 344m/s
f	frequency in Hz
G	sound strength
I	intensity of a sound (W/m^2)
λ	wavelength in meters (m)
L_x, L_y, L_z	room dimensions, length, width and height (m)
n_x, n_y, n_z	room mode numbers
p	probability value (p-value)
p_m	measured sound pressure
p_o	threshold of hearing (20 µPa)
p_0	equilibrium hydrostatic pressure
ρ	population correlation
ρ_0	ambient density of the medium (kg/m^3)
Q	directivity factor
r	correlation value
S	is the total surface area of the room boundaries
t	time (s)
$\hat{T}$	estimation of the expected RT (seconds)
T_K	temperature in Kelvin
t_r	test statistic
RT_{20}	equivalent reverberation time from 20dB sound drop
RT_{30}	equivalent reverberation time from 30dB sound drop
RT_{60}	reverberation time for the full 60dB drop (Also T_r and T_{60})
μ_1, μ_2	the mean of a population
V	room volume (m^3)
γ	ratio of specific heats
z	distance from centre point of source

Acronyms

ANSI	American National Standards Institute
AR	artificial reverberation
atm	atmosphere (Standard)
BPF	band-pass filter
CD	compact disc
dB	decibel
DC	direct current
DF	degrees of freedom
DRM	digital rights management
DSP	digital signal processing
EDR	effective-decay-range
EDT	early decay time
EEG	electroencephalogram
FFT	fast Fourier transform
GB	gigabyte
HPF	high-pass filter
IACC	inter-aural cross correlation coefficients
IACF	inter-aural cross correlation function
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
ITD	initial time delay
ITDG	initial time delay gap
LEDT	lateral early decay time
LF	lateral energy fraction
LPF	low-pass filter
MLS	maximum length sequence
MP3	derived from MPEG-1 Audio Layer 3, moving picture experts group
MPEG	moving picture experts group
RH	relative humidity
RMS	root mean square
RT	reverberation time
SANS	South African National Standard
SNR	signal-to-noise ratio
SPL	sound pressure level
THD	total harmonic distortion
UJ	University of Johannesburg
USB	universal serial bus

Chapter
1

1 INTRODUCTION

1.1 Introduction

Hearing is fused to many human activities, for example; enjoying professional musicians performing in a live concert, hearing an ambulance's siren in traffic or even just having a verbal conversation; sound provides a large part of our human experience. The art of high definition sound reproduction is key to maximising the sense of hearing. The introduction and widespread availability of digital signal processing (DSP) has enabled even the most basic home theatre system to provide a new experience of audio during television viewing or music listening, that only select few were able to enjoy 15 years ago. The combination of the advances in digital electronics, acoustic software and studies on sound effects, have allowed the average consumer to obtain an unparalleled level of acoustic enjoyment. This has not been as widely available and readily possible in the past. Hearing damage, whether due to old age, accidental exposure or congenital factors can be a traumatic occurrence; thus, taking advantage of this wonderful gift of hearing is more of a duty than a choice.

Acoustics is a large subject that can be studied both empirically relying on objectivity as well as psychologically, relying on subjective perceptions. The focus of this study is on acoustic reverberation and the application of artificial reverberation.

1.2 Motivation for this Topic

Why has reverberation time (RT) been the focus of this dissertation when there are many acoustical properties available for study? As highlighted in the ISO (International Organization for Standardization) (2008:v) standard pertaining to reverberation measurement, there are several reasons to measure RT. Reverberation time and sound level measurements are core acoustic parameters, especially when solving environmental or occupational noise problems including; sound pressure level (SPL) from noise sources with regards to noise control, the intelligibility of speech, and the perception of privacy in a room. Music appreciation is directly related to RT including recording studio design and design of performing arts venues.

Measurement of reverberation applies to all sorts of room types, including classrooms [to which some countries apply certain building codes], offices, workshops and industrial plants, sports halls, churches and even motor vehicle interiors. Reverberation time is one of the measures required for the specification of absorptive materials for buildings including insulation measurements put forward by ISO 140 and ISO 3740. This subject area has a unifying ability in bringing together acoustics, civil engineering, architecture, musicology as well as unifying objective and subjective sides of the human experience. In terms of measurement repeatability and accuracy, RT has been found to be a better index than many other acoustic quality indexes (Bork, 2002).

There is consensus that reverberation is a central component for the facilitation of music appreciation (Zhang, 2005:19). This applies to the researcher, the architect as well as the recording engineer. In terms of auditoriums, RT is said to be the most important characteristic of the hall (Berg & Stork, 1982:213). Reverberation is one of the most important post-production tools available to the sound engineer or computer musician (Murphy, Howard & Tyrrell, 1998).

Computer aided technology has become an integral part of our daily lives. Software based solutions provide an alternative to the hardwired electronic methods and reduce the equipment footprint as well as the cost. With the advancements in acoustic holography in its ability to provide variable reference points for an acoustic control system, there is a further reliance on DSP techniques. The ability to compensate for physical room acoustical characteristics has become part of the acoustics field. The introduction of software based AR methods increases the range and flexibility of the acoustical control. A look into the efficacy and degrees of control of AR is integral in the field of ambiophony, electronic architecture and in particular, the recording studio.

Is it possible to isolate RT and study it in isolation? Ando and colleagues (2005) have researched the physiological process of the way RT is handled by the brain and believe that reverberation is an acoustical index that can be described independently from other indexes.

1.3 Contribution of this Study

A study on acoustic reverberation including a review of the ISO standard pertaining to the measurement of RT was conducted. A study of four acoustically different locations was carried out to determine their RT. A comparative analysis of sound recordings and sound playback was then undertaken with the use of computer aided AR. The objective of this study was to prove that adequate results could be obtained by use of software compensation for recordings and playback undertaken in moderately “dead” recording environments that could match recordings in livelier performing rooms that have natural reverberation. The application range is wide, including the recording studio, the concert hall and movie theatre. For example, in the music recording studio there is a considerable reliance on the room’s shape, size and finishes. Obtaining the correct natural reverberation characteristic in a recording room is often time consuming and costly owing to the complex layout and design requirement of the room. Performance halls too are designed specifically for certain RTs, which too are complex and costly in their design. The incorporation of AR has the ability to reduce the cost of the room finishes and can compensate for rooms that do not provide adequate natural reverberation. An explanation as to how this work fits into the music recording industry is highlighted in the final chapter under the heading “The Bigger Picture”.

Further uses for variable reverberation have become popular in performing arts venues whereby different RTs are required for different performance types. The idea of artificially adjusting the RT for performing arts venues has become a mainstay in this field. The degree of control and the efficacy of such methods are pertinent in their further development and acceptance.

Statistical analyses of the subjective appraisals were conducted as part of this study in terms of the auditory assessment of RT and early decay time (EDT). Interesting findings regarding the subjective evaluations of reverberation were uncovered, which raised new questions as to how competent are people in differentiating between reverberant and less reverberant sounds. This study therefore combined the objective (mathematical) and the subjective (personal) sides of acoustic reverberation. Thus, a holistic view is presented which is in keeping with the nature of acoustics.

As part of the academic staff of a local university, it was decided that a subject of acoustics should be added to the new degree that the university would soon be offering. While acoustics is a large field, the main area of study would be in sound recording and reinforcement and its measurements. The subject would be offered in conjunction with the general subjects of electronics and electrical engineering, and thus would make use of the principles learnt in those subjects as a foundation for the acoustics module. The experiments undertaken have been structured in such a way that they can be used as a teaching tool for students who want to study sound and reverberation in their engineering studies. A focus on the practical aspects and its application has been included.

1.4 Outline of Dissertation

Chapter 2 offers a summarised review of acoustic theory including practical issues that are relevant to reverberation. The chapter also introduces Digital AR and psychoacoustics.

Chapter 3 reviews the theory of measurement for RT using the recent ISO standard as a reference. There is a focus on the impulse method for the measurement of RT.

The practical experimentation part of this study begins in Chapter 4. This chapter consists of a study of four acoustic locations that are measured and analysed in terms of their RT and EDT.

Chapter 5 uses the results from Chapter 4 and applies AR methods to a relatively anechoic sound impulse to such a degree that it matches a much livelier impulse sound. This process was conducted for three acoustic locations. The results are evaluated using maximum length sequence software. This chapter has been termed the “Objective Test”.

A subjective testing method is introduced in Chapter 6 where two sample groups are used to determine if they could distinguish the artificially reverberated sound from the naturally[1] reverberated sound. The results were statistically analysed.

The final chapter brings this dissertation to a close. Conclusions are drawn and possible future studies are discussed.

Extensive use of images are presented throughout the experimental chapters (Chapters 4 and 5) to provide the reader with a graphical impression of the data. This includes numerous photos and computer screen views, which is common in the field of acoustics (particularly for studies on performance venues).

1.5 Research Scope

The RT measurement of a room is not the only acoustic property of interest when examining the sound characteristic of a room/space. Other areas of interest include the measure of relative sound pressure levels, early/late energy ratios, lateral energy fractions and intra-aural cross correlation functions, however the central theme is on RT, EDT and application of digital AR.

1.6 Research Method

A large part of the study was based on empirical testing and verification of theoretical knowledge and thus a classical approach was undertaken that could be described as positivistic. However, the subject matter is not a clear-cut science, particularly in the subjective perception of sound; thus, a different research design was used for the subjective evaluation of sound (Chapter 6). In this subjective section, a questionnaire design was incorporated that was statistically evaluated (borrowing the methodology from the social sciences). While this method could still be described as positivistic from the perspective of the social sciences, in terms of classical engineering methods, it was a different method.

Acoustics has both objective and subjective aspects to it. Thus, in order to provide a full picture of the study area, both objective and subjective methodologies were used. Objective methods such as empirical measurements and mathematical evaluations were conducted, while subjective methods such as personal appraisals and individual listening tests were relied on for the subjective section.

[1] Reverberation that can be accounted for by the room design (dimensions, materials and layout) and occurs passively. When the natural reverberation value is manipulated in a manner that changes this natural response, then it is termed artificial reverberation.

Chapter

2

2 ACOUSTIC REVERBERATION

2.1 Introduction

This chapter starts with a focus on the fundamentals of sound including the pertinent topics such as reflection and absorption of sound and how they are related to the sound field. Acoustic reverberation and its measurements are then discussed including a few practical issues regarding SPL, feedback and microphone placement. Performance venues are discussed including the recording studio. Artificial reverberation as well as some advantages are also discussed. The digital workstation is introduced as it was used extensively in the editing of the audio samples in this study (chapter 5 and 6). The chapter closes with a short discussion on psychoacoustics.

2.2 Physics of Sound: An Overview

2.2.1 Definition

Sounds are pressure variations in our environment that the human ear can detect. A sound exists if a change in pressure or displacement occurs in an elastic medium that results in a person perceiving this disturbance through his/her auditory system, or by the pickup system of a measuring instrument (Beranek, 1954:3). The number of pressure variations per second is called the frequency of the sound and is measured in Hertz (Hz). The frequency of a sound produces its distinctive tone. The lowest frequency of a sound determines the pitch[2], while a pure tone occurs when there are no overtones. Pure tones are a rare occurrence, as most sounds are combinations of different frequencies and/or harmonics superimposed together with the average young human ear only hearing up to ±20kHz. The human ear perceives sound logarithmically. The wavelength of sound represents the distance between adjacent pressure maxima [or minima] and ranges from just a few centimetres to a few meters. Octaves are used as conventional auditory sound intervals (Barron, 1993:10).

[2] It has been found that increasing the intensity of sound can change the pitch that is heard. For instance, if a 100Hz audible signal is increased from a loudness of 40 to 100, there will be a 10% decrease in the pitch heard (Tremaine, 1974:20). Pitch has been linked with the basilar membrane (delicate part of the cochlear) which exhibits a nonlinear response between pitch and frequency.

2.2.2 Why Measure Sound?

Measuring sound begs the question, what aspect of sound should be measured? For example, one area of interest could be noise measurements. An objective measure of sound annoyance is difficult to quantify as this would be a personal value and would differ among people. However, precise scientific measurement is important in building designs, musical/entertainment performances and the work/school environment. The measurement of human hearing sensitivity is called audiology. By obtaining environmental sound level measurements and having knowledge of the human ears' sensitivity, one can better manage noise in our environment such as at airports, highways and our homes, with the goal of improving our quality of life. Other aspects of attempting to objectify sound may pose problems to the researcher as perception of sound is often personal. Measuring favourable sounds has similar challenges to that of measuring unfavourable sounds.

2.2.3 Velocity of Sound

The speed of sound has been calculated to be 331,5 m/s at an air temperature of 0°C at 1atm (Kinser, Frey, Coppens & Sanders, 2000:528). They used the following equation to determine the speed:

$$c^2 = \gamma p_0 / \rho_0 \qquad [2.1]$$

Where:

γ is ratio of specific heats

ρ_0 is the ambient density of the medium (kg/m^3)

p_0 is the equilibrium hydrostatic pressure

$$c_0 = (1.402 \times 1.01325 \times 10^5 / 1.293)^{0.5} = 331{,}5\, m/s \qquad [2.2]$$

This theoretical value has been found to be very close to actual lab measurements of sound (Kinser et.al., 2000:121). To obtain the velocity of sound at a temperature other than 0°C (273K):

$$c = c_0 (T_K / 273)^{1/2} = c_0 (1 + T / 273)^{1/2} \qquad [2.3]$$

Where:

T_K is temperature in Kelvin

Thus at 23°C, the velocity of sound is calculated as follows:

$$c = 331{,}46(1 + 23/273)^{1/2} = 345{,}1\text{m/s} \qquad [2.4]$$

If the direct wave needs to travel 5m from the source to the receiver then it would take just over 14ms, while a reflected wave travelling a distance of 15m would take

43ms. If one were sitting in an auditorium where the distance from the performer is 20m, the time taken for the first direct wave to reach the listener would be 58ms, while a long reflected wave travelling say a total distance of 100m to a distant listener would be delayed by 232ms. This example begins to illustrate the cause of reverberation[3].

By knowing the speed of sound we can obtain the wavelength:

$$Wavelength\,(\lambda) = \frac{Speed\,of\,sound\,(c)}{frequency\,(f)} \qquad [2.5]$$

Where:

c is velocity of sound (m/s)

f is frequency (Hz)

λ is wavelength in meters (m)

Using the above equation we can calculate that at 20Hz, one wavelength is just over 17 meters, while at 20kHz, it is only 1,7cm. High frequency sounds have short wavelengths compared to low frequency sounds. The density of the medium has an effect on the speed of sound, but for this example, the above equation will suffice. According to Smith, Peters and Owen (1996:1), the above equation is accurate enough and acceptable for building science applications with respect to acoustics.

2.2.4 Loudness

Sound has a dynamic characteristic and is dependent on intensity. Loudness of a sound is the size of the auditory sensation produced by that sound (Olsen, 1967:251). The unit of loudness is the *phon*, which is linked to sound pressure level (SPL) at 1kHz. The *phon* unfortunately does not give a good indication of the human reaction to loudness. Firstly, as the frequency range of the human ear is non-linear and secondly, there is still argument about what a doubling of loudness equates to as some researchers say 10dB while others say less. A subjective unit of loudness has been found and is termed the *sone,* which is defined as the loudness experienced by a person listening to a pure tone of 40 *phon* loudness level (Everest, 1981:38). The *sone* and *phon* are still used but as most instruments make use of SPL in dBs, loudness is referred to as dB SPL from here forward. The dB SPL makes use of the 20µPa threshold of hearing as its reference level.

2.2.5 The Decibel

Sound is a pressure phenomenon. The human ear detects the pressure fluctuations based on the pressure or amplitude. The lowest fluctuation that the human ear can

[3] Simplified example assuming all other variables such as air temperature, wind flow etcetera are held constant.

detect is 20μPa. This is five billion times less than the normal atmospheric pressure[4]. At a value of 20μPa, the human eardrum would move a distance that is less than the diameter of a single hydrogen molecule (Brüel & Kjær, 1984:6). Owing to the fact that the human ear can accept pressure variations that are more than a million times its threshold, a dB scale is used to better manage the large numerical values. By using a logarithmic scale such as the dB scale, the response of the human ear can be practically represented. The value of 0dB is based on the threshold point of the human ear, which is 20μPa and thus dBs do not represent an absolute but rather a relative value.

To calculate the SPL in decibels:

$$SPL = 20\log\frac{p_m}{p_o} \qquad [2.6]$$

Where:

p_m is the measured sound pressure

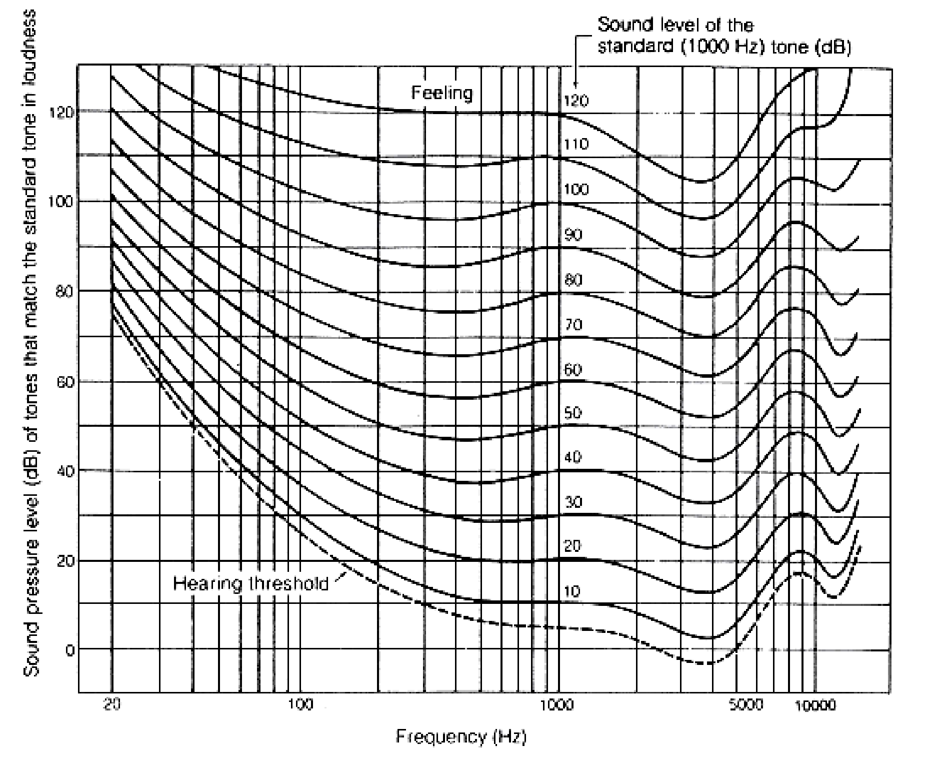
p_o is the threshold of hearing (20 μPa)

2.2.6 Hearing Range

The human ear can hear from 0dBs to over 130dBs, with the latter being the pain threshold. The subjective perception of sound level does not show complete linearity with that of power radiated from a sound source. One reason is that the human ear has differing sensitivities across the frequency range. Figure 2.1 shows the frequency response of the human ear.

[4] The standard atmosphere (symbol: atm) is defined as being precisely equal to 101.325 kPa. Thus 101.325kPa divided by 20μPa is equal to $5{,}066 \times 10^9$.

Figure 2.1: Equal loudness contours[5] (Robinson & Dadson, 1956).



A further complicating factor is that for each frequency, our ears do not always perceive consistent SPL increments that are applicable to another frequency's SPL increase. In particular, the low frequency range has an uneven distribution of perceived loudness when compared to the same SPL of higher frequencies. For example, generally, a young person's ears should be able to hear a 1kHz sound at an SPL of 25dB, but would only be able to hear a low frequency sound of say 60Hz if the SPL is increased by a further 30dBs. From Figure 2.1, the perceived loudness of each contour exhibits both a non-linear shape along the frequency axis as well as each contour tracing a slightly different shape. Another complicating factor relates to the time duration of the signal. There is a difference on the perceived loudness for steady state versus impulse sounds. Generally the shorter the sound impulse (less than 70ms), the lower its loudness is perceived to be (Brüel & Kjær, 1984:8). The SPL meter's sensitivity is based on the equal loudness curves and is discussed further in chapter 3.

2.2.7 Sound Intensity and Sound Power

Instantaneous sound pressure at a location is the incremental change from the static pressure due to the influence of a sound wave in a given instant (Kang, 2002:2). The effective sound pressure is the RMS value of the instantaneous sound pressure for a

[5] There are at least 17 equal loudness contour graphs to choose from. While the first popularised contour mapping was presented by Fletcher and Munson in 1933, it was later found that it was not completely accurate. In 1956, Robinson and Dadson presented their contour map, which has been used extensively. The ISO226:1987 standard was based on this loudness contour map of Robinson and Dadson. Further research has shown that this too was not entirely correct. A new standard has been set out namely, ISO226:2003 [or TC43] which seeks to portray a more accurate loudness contour map based on 12 studies starting in 1983 (Suzuki & Takeshima, 2004; ISO, 2009).

given time period (Kang, 2002:2). Sound intensity is the average sound power per unit area and is direction specific. Sound is transmitted in air as longitudinal waves as the gas particles are shifted in the path of the travelling wave (Kang, 2002:4). Sound energy is transferred from the vibrating sound source with the sound pressure being dependent on the power of the source as well as its air coupling ability.

2.2.8 Frequency of Sound and Frequency Weighting

Most sounds consist of more than one frequency component. Thus, to measure SPL a range of frequencies can be specified. These are commonly called frequency bands or octave bands.

2.2.8.1 Octaves and Octave Bands

Musical notation is given in octaves. One full series of eight notes on the musical scale is one octave, which is a pitch interval in music. Any two notes, which are an octave apart, have an approximate frequency ratio of 2:1 or 1:0,5. Each octave is divided into 12 notes (chromatic scale), called semitones, which are visible as black and white keys on the piano keyboard. An interval in octaves, between any two frequencies, is logarithmic to the base two of the frequency ratio (Tremaine, 1974:12).

Sounds that contain energy over a wide range of frequencies are divided into ranges called bands. An octave band is a common division of frequencies identified by their centre frequencies which are 63, 125, 250, 500, 1000, 2000, 4000, 8000 and 16000Hz. Octaves are commonly used to describe a filter's behaviour, for example, the roll-off of a low pass filter may be given as attenuating at 3/dB per octave. This can include a gain value as well, for example, a gain of 6dB/octave defines a change of 6dB for each doubling or halving of frequency. An octave filter with a centre frequency of 1kHz allows frequencies between 707 and 1414 Hz, but rejects all others. A third octave covers a range where the highest frequency is 1,26 times the lowest frequency. Acoustic measures are frequently undertaken in octaves as a matter of convention and can be seen throughout the applications chapters 5 and 6.

2.3 Reflection, Absorption, Diffusion, Diffraction and Transmission of Sound

2.3.1 Sound Propagation

A single source radiates sound energy in all directions as a concentric wavefront. The wavefront can be compared with the effect of a stone thrown into a pond where ripples propagate from the point of impact outward towards the banks. When the ripples hit the bank, a return wave is often seen, which is similar to the sound wave reflecting off a nearby wall as an echo.

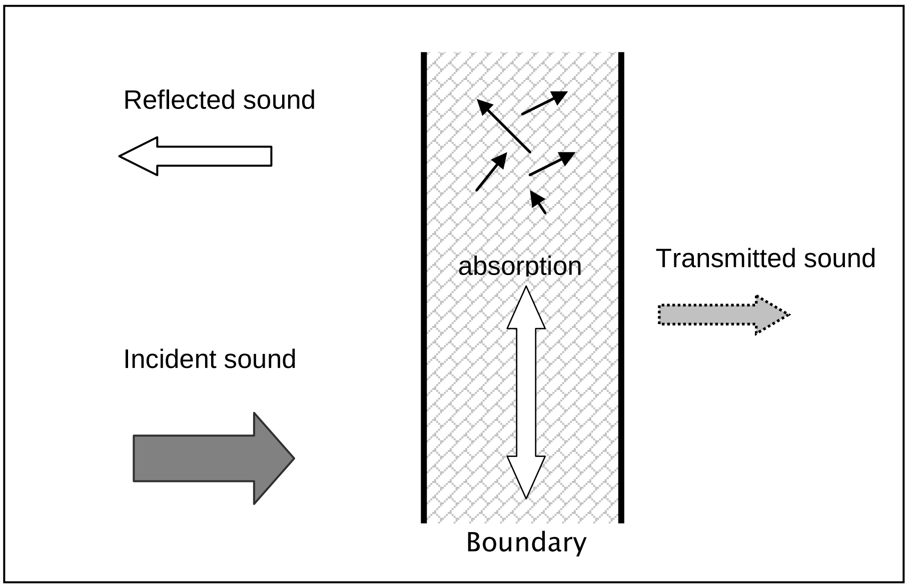
2.3.2 Sound Field in front of a Boundary

When sound is directed towards a boundary surface, there may be any combination of reflection, absorption and diffusion depending on both the sound parameters and the boundary type. Sound energy that is reflected from a room boundary contributes to the sound-field in the room, while absorbed sound dissipates as heat, and transmitted sound energy propagates through the boundary surface. Reflected waves that have wavelengths of incident smaller than the dimensions of the reflecting surface, have an angle of reflection equal to the angle of incidence. Sound waves that strike a wall and go in many directions are termed reflection with most of the sound reflected back into the room where it originated. The reflected sound may give rise to repeated semi-independent sounds or echoes. The geometrical pattern of the reflected sound can be used to identify echoes/problematic areas in a performance room by use of *ray tracing* (Brüel & Kjær, 1988). An echo is defined as a sound reflection that arrives more than 50ms after the direct sound [or has a path length of at least 17m longer than the direct sound] and can be heard almost distinctly from the original sound. Waves that are heard within ±55ms usually have the effect of changing the sound quality and are thus not normally termed echoes (Eargle & Foreman, 2002:171). Reverberation differs from echo, as reverberation is a grouping of many echoes from numerous surfaces that create an overlapping of echoic sounds.

If the majority of reflected sound is spatially and temporally dispersed, the reflection is termed diffuse and the reflecting surface (usually irregular) is called a diffuser (Cox & D'Antonio, 2004:1). The diffusely reflected sound's angle of incidence bears no resemblance to the angle of reflection (Kang, 2002).

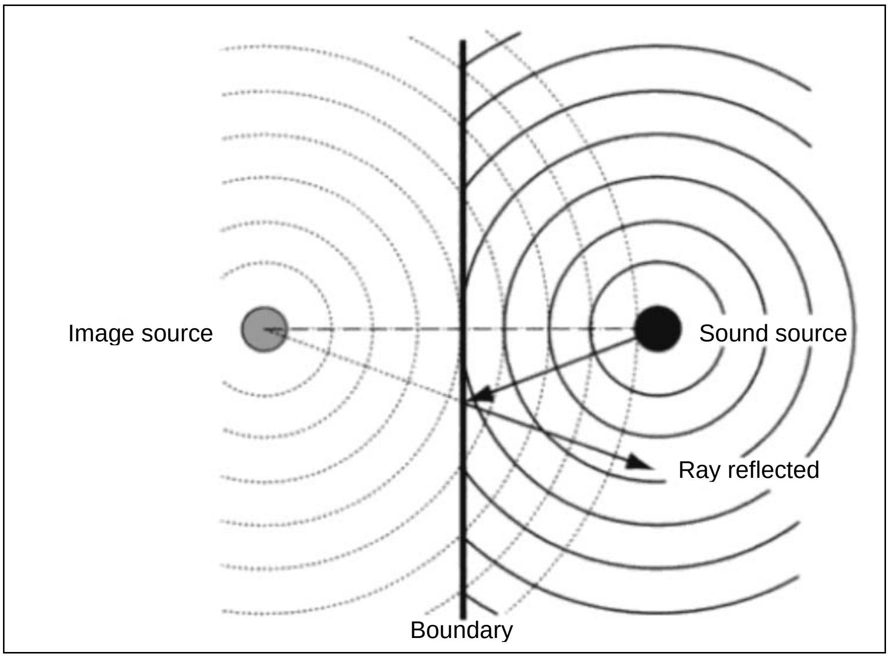
If the surface has an absorptive ability, the material is able to reduce most reflections and thus the sound would tend to decay without repeat. While absorption and reflection are at opposite ends to each other, there are many combinations between these two extremes. There are also excessively irregular surfaces where the sound tends to scatter and in some instances, for example a wall surface may have been designed for this very effect. The next figure shows a simplified example of sound reflection and absorption at a boundary. Some sound energy is transmitted through the wall. Sound that is not reflected back or transmitted through the wall is absorbed by the wall and is lost as heat dissipation. The reflection of sound from a wall normally involves some degree of scattering.

Figure 2.2: Sound reflection, absorption and transmission.



A sound wave that impinges on a smooth boundary surface encounters reflection from the plane surface. The pattern of reflection is shown in Figure 2.3 where the reflected sound seems to originate from a mirrored image of the source. If the surface is curved, the same effect applies and can be solved using geometry.

Figure 2.3: Sound source and its associated image during a plane surface reflection (Kang, 2002:8).



2.3.3 Interior Furnishings

Apart from the boundary surfaces, room furnishings affect the sound field within the room. Generally, cloth based furnishings such as couches, clothing and curtains are classified as absorbers. Harder surfaces such as tables, metallic/hard wooden objects such as chairs, act as diffusers, as they tend to reflect the sound impinging on it and thus disturb the wavefront. It is not always a clear classification of which furnishings act as absorbers and which act as diffusers as some do both (Nisbett, 1995:37). The sound frequency and wavelength also affect whether a device is generally absorptive or diffusive; for example, a drinking glass, while made from diffusive material, would only affect very high audio frequencies as the lower ones would move past it. A curtain can be considered an absorber but a thick heavy drape that hangs slightly off the wall would have a much higher absorptive characteristic and thus the thickness and position of the furnishings are also important in determining their abilities.

2.3.3.1 Diffraction and Refraction of Sound

Diffraction describes the bending or spreading out of sound around obstacles in its path. The obstacle size and shape as well as the frequency of the sound wave interact in a complex manner. Diffraction usually occurs after sound impinging on a recess or surface protrusion and is more significant for lower frequencies (Kang, 2002:8). Refraction describes the change in speed of sound as it moves from one medium to the next.

2.3.4 The Absorption Coefficient

The absorption coefficient α, describes the ratio of acoustical power absorbed to the total sound striking a surface and is a function of frequency. The equation is as follows:

$$\alpha = 1 - \frac{reflected\ energy}{incident\ energy} \qquad [2.7]$$

Absorption coefficients normally range between the ideals of zero (no absorption) and one (total absorption). There are many reference values for various materials and surface finishes, which are given over several octave bandwidths covering the range from 125Hz to 8kHz. For example, if a surface has an absorption coefficient of 0.4, it will absorb 40% of the power and reflect the remaining 60%.

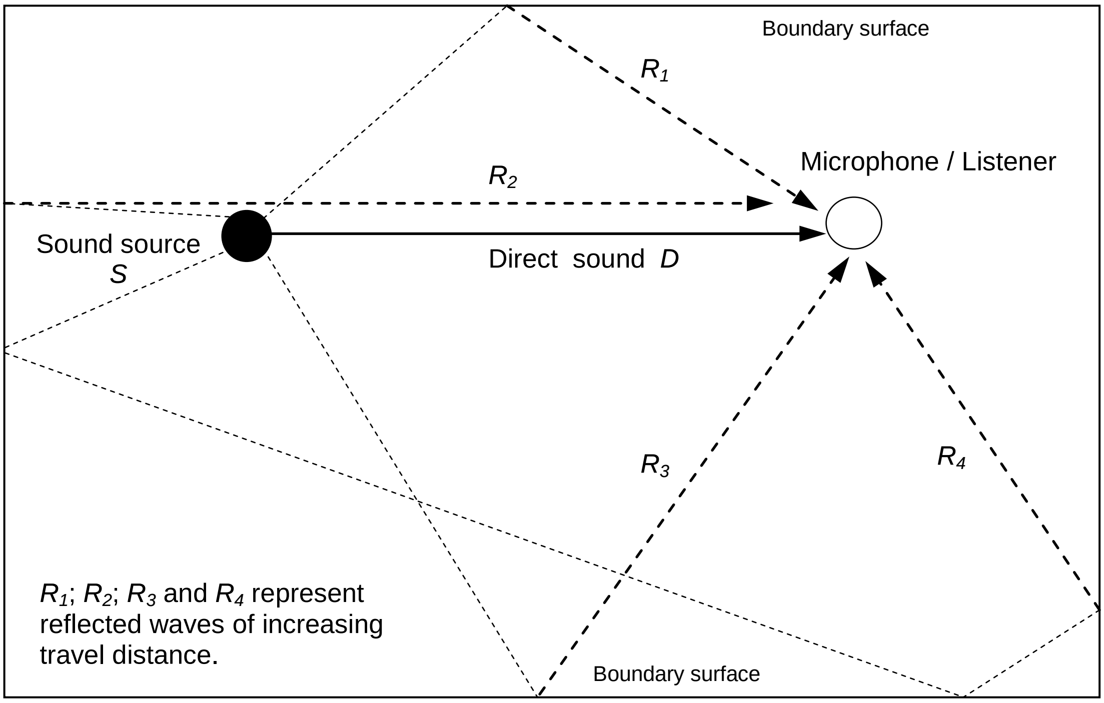
In terms of enclosed rooms, once the walls have been built, there is little that can be done to improve insulation, that is, to reduce transmission between adjacent rooms. The reason is that the characteristic response is already inherent in the design and construction of the walls (Eargle & Foreman, 2002:11). However, improvements can be made to the boundary surfaces of the walls by reducing reflections being reradiated into the room by use of damping materials. Manufactured treatments such as; foam wedges, pyramids and various other shapes, fibreglass panels, Helmholtz traps and many other devices can be used to improve the room’s sound (Gervais, 2006). Materials differ in their frequency response with respect to absorption, thus by manipulating the proportions of the room’s absorbers, it is possible to gain some control over the musical warmth of a room as well as improving the clarity for speech (Brüel & Kjær, 1988). The absorption characteristic of a boundary surface/material can be ascertained by use of an impedance tube or a reverberation room (Kang, 2002:6).

2.4 Development, Growth and Decay of Sound Fields in a Room

Using Figure 2.4, consider the source *S,* which is energised to a steady state and outputs an audible sound. The direct sound will travel in various directions and creates numerous paths between the source and the receiver. The shortest path is path *D* (direct sound) which takes the shortest time to present itself on the microphone/listener. At this instant when the first sound wave reaches the pick-up, other waves are still en route and depending on how much distance they need to travel, this determines the time delay in sound reaching the microphone. The direct wave will reach the listener first while the reflected wave R_4 will reach the microphone last. Owing to the varying distances that are present, many sound reflections reach the microphone over a certain period. If the first sound impinging on the microphone is taken as occurring at a time value of zero seconds, then anything later can be labelled as zero plus the time taken for each successive wave.

The first reflected waves heard at the microphone are termed early reflections and are usually in the range of between 25ms and 100ms. In the next time instant, that is, after 100ms, decaying sound will arrive from all directions in the room more or less equally (Eargle & Foreman, 2002:18).

Figure 2.4: Multiple sound waves from a single source (Adapted from Everest, 1981).

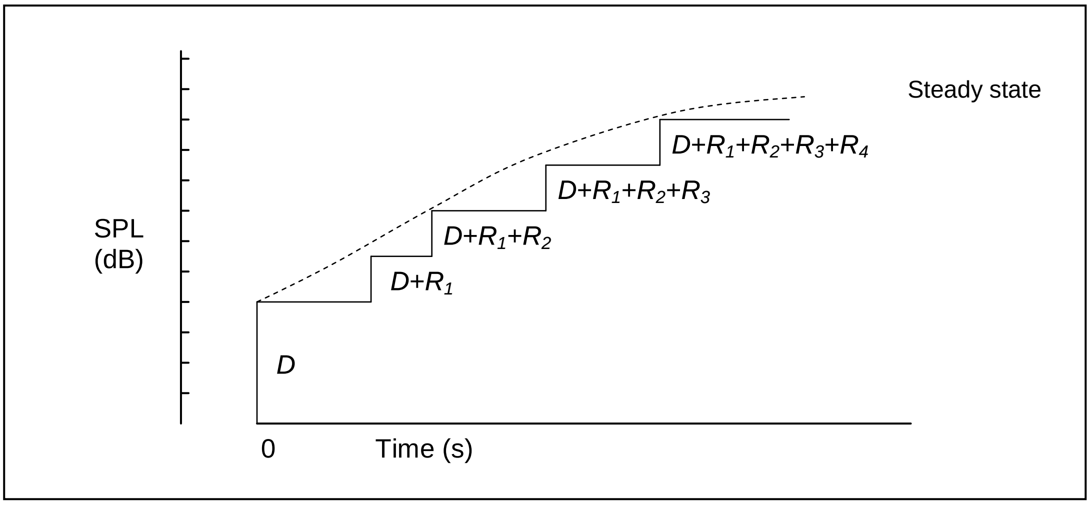


As time elapses, more sound waves will be present at the microphone. If we analyse the SPL at the microphone, we will see that the pressure level will develop until a steady state situation arises, assuming that the source energy has not been adjusted during its excitation. This phenomenon is shown graphically in Figure 2.5. It is assumed that at point zero on the time axis no reflected waves have reached the listener yet. At this point only the direct wave has reached the pick-up. The sound pressure measured at the pick-up would be slightly less than that obtained at the source owing to losses, such as spherical divergence and small losses in the air (Everest, 1981:132).

As the reflected waves R_1 to R_4 reach the microphone, the SPL grows to a higher value than if there had been no reflected waves. The exact growth value of the SPL would need to take into account the different wave phases and their respective wave shape at the combining point (Everest, 1981:132). Figure 2.5 assumes that each successive reflected wave has an additive ability when combined with the direct wave. If one looks deeper into this assumption, one may see that this is not always the case. It may occur that a reflected wave is out of phase [or even partially out of phase] with another reflected wave and has a cancelling effect. Another factor may be that part of the reflected wave cancels with part of another wave while still being in phase with a third wave, and thus adds to the SPL growth. A last option is that two waves may add to each other for one time interval but subtract in the next time interval owing to their different wave shape. Overall, a net additive value was assumed.

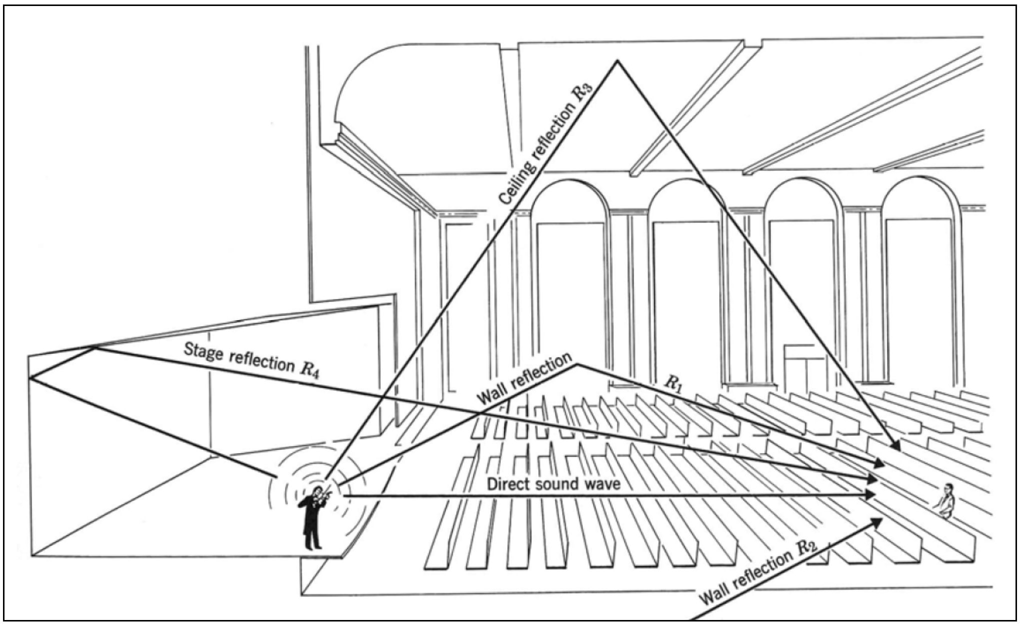
The reflected waves need to travel longer distances and thus their pressure values can only be noticed a short time after the original direct wave; thus, the SPL level at the microphone does not reach its peak instantaneously. The total obtainable SPL that can be generated in the room is dependent on the following: the input energy at the source, the efficiency of the radiating source device, the losses due to heat dissipation of reflecting surfaces and the small air losses. If the radiating source is energised by a constant input power value and its efficiency remains stable, the SPL in the room will reach steady state after a certain period. This is a function of the travel distance of the reflected waves. If the excitation energy at the source was increased, a new steady state SPL could be obtained. A graphical pattern of the growth and decay of sound in a room is usually plotted on a logarithmic scale, which has the effect of adjusting the slope to a linear shape. In practice, there are large amounts of reflected waves that add up to increase the sound level and thus the shape of the graph is curved (non-logarithmic scale) or linear for a logarithmic scale. Figure 2.5 shows a step-like growth and decay, which are for explanation purposes only. The dotted line represents a more true to life rendition of the actual response.

Figure 2.5: The positive development of sound pressure over time.



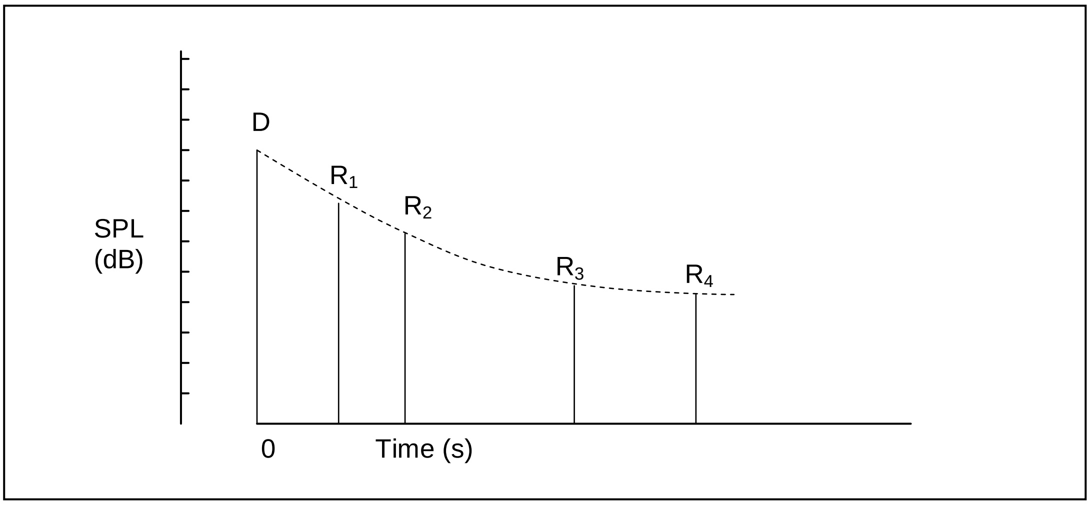
A better example that contextualises the reflected waves is shown in the next figure, which is of a performing auditorium.

Figure 2.6: Sound propagation in an enclosed space (Beranek, 2004:21).



Continuing with the explanation of sound growth and decay, if the source is suddenly stopped from radiating, a reverse situation to what has just been described occurs. At the instant when the source is de-energised, call this time t_{stop}, no more sound waves will leave the source irrespective of which direction they were heading. The SPL level at the microphone starts to reduce but does not drop to zero instantaneously. The sound waves that had just left the source prior to it turning off are still en route and are heard for a time after the source was quiet. Some of the waves are direct while others are reflected. The direct waves tend to be more prominent as they have encountered minimal loss between the source and the microphone. The direct waves thus present with a higher sound pressure at the microphone than the reflected waves that have encountered mainly heat losses through reflection and/or refraction. However, the direct path is shorter and thus the direct waves are lost first. The reflected waves while offering a lower sound pressure, persist for some time before the SPL drops completely (assuming no background noises). The waves that were still travelling lose their energy feed and are overcome by frictional losses. For example, a wave that was first reflected by a sidewall and then was reflected onto the ceiling before finding its way to the receiver has already encountered loss due to heat dissipation on first wall followed by a loss due to the ceiling and a frictional air loss. Each loss reduces the energy present in the travelling wave and may even reduce it to a point that is beyond human auditory range. The SPL reduction seen at the microphone has an exponential decay response. This decay time is definable as the speed of sound in air is stable[6]. Figure 2.7 illustrates this process of sound decay after the source is no longer energised.

Figure 2.7: The decay wave shape after source is no longer radiating.



2.4.5 The Attenuation of Sound in Air: Inverse Square Law

If a point source (small and spherical) excited a free field, the sound intensity would be calculated as:

$$I = \frac{W}{4\pi z^2} \qquad [2.8]$$

Where:

z is distance from centre point of source

I intensity of a sound (W/m^2)

From Equation (2.8), it can be deduced that sound intensity at any point is inversely proportional to the square of its distance from the source (ideal conditions). This is appropriately termed the inverse square law. Thus, for every doubling of distance, there is a 6dB drop in SPL (Berg & Stork, 1982:30). For example, if there are no nearby

[6] The above explanation assumes that the enclosed space has a reverberant effect. This can be contrasted with an anechoic environment where the sound waves have no reflecting materials en route and thus dissipate into the air or through absorptive materials at boundary surfaces.

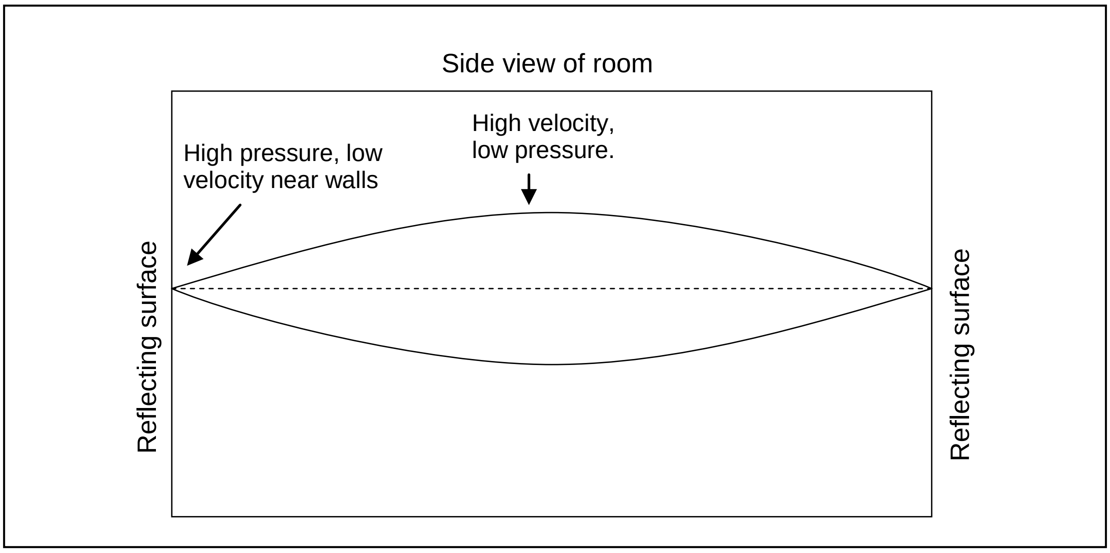
reflecting objects (free-field condition) and one measures an SPL of 80dB at 1m, then by moving to 2m away from the source, the SPL would be 74dB. The absorptive and reflective properties of objects in the way of a travelling sound wave are dependent on numerous factors. One such factor is the wavelength. For example, a 10kHz signal has a wavelength of 3,4cm and thus even small reflective objects en route would have a noticeable impact on this wave. Even the top of a microphone is generally bigger than 3,4cm and would have an effect on the sound field. Higher frequencies are easier to attenuate than lower frequencies, which is evident when one hears a nearby disco and can mainly hear the bass notes. Thus, at higher frequencies, attenuation increases beyond the inverse square law.

A point source is uncommon and most sources are firstly larger and secondly have some directionality. The directivity factor is known as *Q*, and is defined as the ratio of intensity at a given distance and angle from the source to the intensity at the same distance, assuming a uniformly radiating source (Kang, 2002:5).

2.4.6 Standing Waves in a Room

When a sound source is placed between two parallel walls, a standing wave may be set up with a lowest possible wave frequency of half a wavelength, which is equal to the distance between the walls as shown in the next figure.

Figure 2.8: Simplified representation of a standing wave between two surfaces.



The crest and troughs of the wave represent the highest air particle velocity while the zeroes have the highest air pressure. Thus, for Figure 2.8 there would be a maximum velocity at a distance between the walls at $\lambda/4$. To obtain a maximum sound absorption in a room, the sound absorbing materials should be placed as close to quarter wavelength as possible so as to interfere with the high air particle velocity (Eargle & Foreman, 2002:22). Rooms represent cubic dimensions and thus support an array of standing waves. Certain frequencies in a room are excited more than others are. The simplest solution to three-dimensional standing waves occurs when the room has a rectangular shape.

A mode is a standing wave that is frequency and room dimension dependent. In the acoustic field, the low frequency modes are of most interest. Each room has a unique modal signature. A mode/standing wave occurs in all room types to some degree, even professional recording studios (Gervais, 2006:16). These standing waves are most problematic at the low frequency range. There is argument as to what range of frequency; for example, Everest (1981) believed that 300Hz and below was problematic, while Gervais (2006:16) argues that *Everest's peak* should be dropped by 100Hz. Gervais (2006) motivates his point by citing that fibreglass compensates for the frequencies above 200Hz. To calculate the frequency of the room modes, the following equation is used:

$$f = \left(\frac{c}{2}\right) \times \sqrt{\left[\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2\right]} \qquad [2.9]$$

Where:

c is the speed of sound in air (345.1m/s)

n_x, n_y *and* n_z are the room mode numbers (integers)

L_x, L_y *and* L_z are the room dimensions, length, width and height (m)

There are software alternatives to manually calculating the room modes. They require the user to input the room dimensions and the software calculates the frequencies in which the axial, tangential and oblique modes would occur.

Modal waves are set up when the room dimensions coincide with a specific wavelength [or small wavelength range]. Under such conditions, the room dimensions do not allow for adequate natural decay of the wave and the wave either strengthens (amplitude increased) or weakens (amplitude decreased). Parts of waves that have their amplitude decreased are termed nodes, while the amplified parts are termed antinodes. Antinodes are located between nodes with the condition becoming worse the smaller and more reverberant the room is (Gervais, 2006:17).

Modes can be described as axial, tangential or oblique. Axial represent the strongest mode type and require two parallel surfaces, while tangential are about 3dB lower and require two sets of parallel surfaces. Oblique modes require six surfaces and are about a quarter as strong as axial modes. Above 300Hz, the modes are combined with other room acoustic effects and thus the main focus is below 300Hz as stated.

2.4.6.1 Explanation of how Room Modes can be Problematic

A rectangular room with dimensions of 5x4x3m (length, breadth, height) would have the following axial modal identity, as set out in Table 2.1.

Table 2.1: Axial Modal activity for rectangular room (100-200Hz range).

AXIAL	**Parallel Length**	**Parallel Width**	**Parallel Height**
Mode Number	Hz	Hz	Hz
2	*(68.8)*	*(86)*	**114.6**
3	**103.2**	**129**	**172**
4	**137.6**	**172**	*(229.3)*
5	**172**	*(215)*	*(286.7)*

When a mode is excited, a standing wave would be present in a room. Along the standing wave, there are pressure maxima and minima points. Table 2.1 of axial mode frequencies shows some of the frequencies between the 100-200Hz ranges. If a microphone were positioned at a point in the room where the standing wave of say 137,6Hz was at a pressure minima (at a node), then the recorded sound would have a reduced bass frequency sound. The bass frequency would sound "thin" (lack of bass sound). This condition and its severity are dependent on the room dimensions, room mode type – axial, tangential or oblique as well as the point where the sound is excited and heard. If a listener were sitting in an antinode location, the bass sound would be thought of as being "boomy" (excessive or overpowering compared to mid and treble sounds) owing to the increased amplification at the pressure maximum location. The point of excitation is important and can be practically demonstrated that some positions in a room have more bass than others. This is not only limited to bass, as many frequencies can be affected in the same way. This concept becomes more complicated as one looks at the tangential modes as well as the harmonics of each mode (Gervais, 2006:18). Compensation for the peaks and nulls introduces further complexities to this problem, especially when DSP methods are used such as adding or subtracting certain frequencies in the mixdown[7]. The point of departure is that room modes are important both in the recording phase as well as the listening phase of music appreciation, as the same phenomena may occur while mastering a recorded sound track. When conducting the RT analysis of the small room (chapter 4), the room modes had an effect on the SPL measured. The measurement of the small room's SPL is found in Appendix A.

With regards to recording studios, the ideal goal would be to obtain as close a reproduction of the actual performance as possible. After the tracks are merged into the mixdown, one would like to hear the final product as accurately as possible. The room that the musicians were recorded in as well as the listening/control room are subject to modal activity, which may change the recorded and/or playback sound. The main effect is increasing or decreasing certain frequencies owing to a resonant/cancelling effect of specific notes within the recorded/playback sound. Speaker boundary interference is a further factor in this regard, particularly in the listening room. This is an important point and the use of AR techniques can assist in this problem.

[7] Process where separate audio tracks of a multiple track recording are combined into a single track.

2.5 The Reverberant Sound Field

2.5.1 Reverberation Time Calculations

When a sound source has excited a room and is then turned off abruptly, one can hear a relatively smooth decay of sound in the room, especially if the surfaces are reflective. This persistence of sound is known as acoustic reverberation. This reverberant sound has a time component. Reverberation can be measured and is defined as the space-averaged sound energy density in an enclosed space that decreases by 60dB (ISO, 2008:2). Simply put as: the time required for the level of the sound to drop by 60dB, often expressed as T_{60}, T_r, RT60 or just as RT.

Walter Sabine conducted experiments in some of the lecture halls at Harvard University at the beginning of the last century, where he studied the impact of absorption on reverberation and the time constants associated with it. Sabin noticed that reverberation was a function of room volume and the amount of sound absorption within it. His revised formula for RT is as follows:

$$RT_{60} = \frac{0.16V}{S\overline{\alpha}} \qquad [2.10]$$

Where S is the total surface area of the room boundaries and $\overline{\alpha}$ is the effective absorption coefficient of the room. V is the volume of the room in m³.

To determine the absorption coefficient for the room, the individual absorption coefficients (α) of the boundary surfaces need to be known including their sizes (S). The absorptive ability of certain materials can be understood by measuring the SPL of a steady-state sound excitation. In a completely reflective environment, the SPL should increase indefinitely with the increasing reflections, however as there is always some absorption, the SPL reaches a steady state. The absorption coefficients are readily available and can be obtained from acoustic manuals, software programs, building material books, noise control textbooks or the internet. Absorption coefficients are also available for different gases and for liquids.

Sabin's equation is known to become inaccurate as the average absorption coefficient increases (Kang, 2002:14). Thus, many authors have argued for the use of Eyring's equation for recording studios rather than Sabin's one (Everest, 2001:160). Eyring and Norris' equation is as follows:

$$RT_{60} = \frac{0.16V}{-S\ln(1-\overline{\alpha}_E)} \qquad [2.11]$$

Where $\overline{\alpha}_E$ is the area-averaged random-incidence energy absorption coefficient.

For highly reverberant rooms (α<<1):

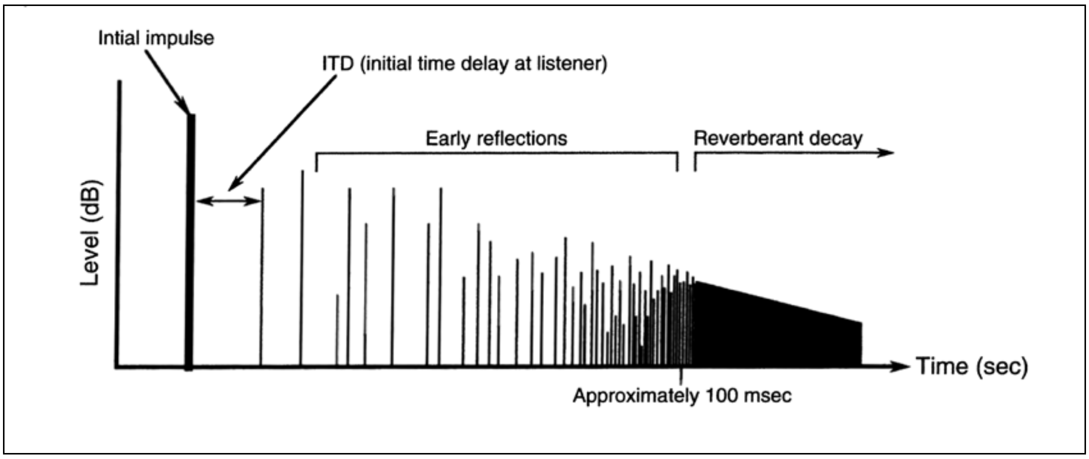
$$\alpha = \alpha_E \quad [2.12]$$

The classic Sabine and Eyring equations are based on the assumption that the sound field is diffuse. Long enclosures do not have a consistent single RT value and length dependent values have been found, which means that applying the classical RT equation to extended linear passages/enclosures would not produce accurate results (Kang, 2002:27).

2.5.2 The Reverberation Decay Slope

Analysing the decay slope, one can see that there are important parts in the sound decay. The decay roll-off is characterised by three terms: the initial time delay gap (ITDG), the early reflections and the reverberant decay as shown in the next figure.

Figure 2.9: What the listener hears (logarithmic scale) (Eargle & Foreman, 2002:22).



2.5.2.1 Early Decay Time or Early Reverberation Time

Early decay time (EDT) represents the initial sound decay response. If a long enough time is allowed between successive musical notes, then the full RT of the sound is heard; however, continuous notes played in succession may overrun the lower reverberant sound and thus it may not be heard. For this reason, EDT is of interest as the EDT may still occur before the next note is played and thus the audience will hear at least some of the decaying sound between successive notes.

The EDT and the RT can be compared with each other. To calculate the EDT, a measurement is required from the time the sound was stopped until it drops by 10dB. The time is then multiplied by 6 to normalise it to the same reference as the RT_{60}. The EDT has been found to be a more important predictor of the subjective perception of reverberation than the standard RT_{30} or RT_{20} measure (ISO, 1997:14; Yoshimasa, 2005:para 5). According to Beranek, EDT is a better predictor of acoustical quality than the full RT, as EDT is what is heard by the audience for most of the time during the performance (2004:24). The EDT has been shown to be dependent not only on the reverberant environment, but also the type of sound excitation used, and can differ substantially from the rest of the decay characteristic (Ivie, 2007). In a diffuse sound field, the EDT should have the same roll-off characteristic as the RT_{30} and would be linear in nature (Kang, 2002:15). As both RT_{60} and EDT are required for analysing the decay characteristic of the sound, both these parameters were studied in the experimentation chapters of this dissertation.

2.5.3 Ideal Reverberation Times

The scientist may ask what is the optimal RT? The problem with this question is that the answer is based on subjective likes and dislikes; thus, no precise answer can be given without a follow-up argument by acousticians who do not agree. There are recommended RTs available that have been researched to be aesthetically pleasing but other factors may confound this. Some authors list their recommended RTs with respect to sound type; classical, opera and so forth, and some authors highlight the size of hall as well. For example, a church represents a mixture between speech and music. While there are many types of structures used in church design (including interior fittings and room volumes), so too are there many types of church proceedings, some being mainly choir orientated, others being mainly speech focused with various degrees of each. It would be an oversight to propose a generic RT specification for the many types of speech and musical performances. The main argument against optimal RT is based on the following two points:

- People may prefer certain RTs only because it is what they have become used to when listening to their music, rather than there being an objective measure that is independent of experience.
- The influence of RT may interfere with other unknown factors.

A general guideline would be a better option and thus put rather crudely, environments that are geared for speech should have a lower RT than musical environments. However, there are obvious extremes especially when applied to speech. Excessively long RTs muffle sound and can make it difficult to understand spoken words as well as increasing the SPL. Thus, environments such as airports and classrooms and so forth would require a lower RT. Rooms for speech are easier to specify than music rooms and much work has been done in obtaining minimum standards for speech intelligibility. An example of which is school classrooms, while for music appreciation this is not as clear-cut. If the reverberation is too short, the sound is termed “dry” and is described as uncomfortable particularly for performing musicians.

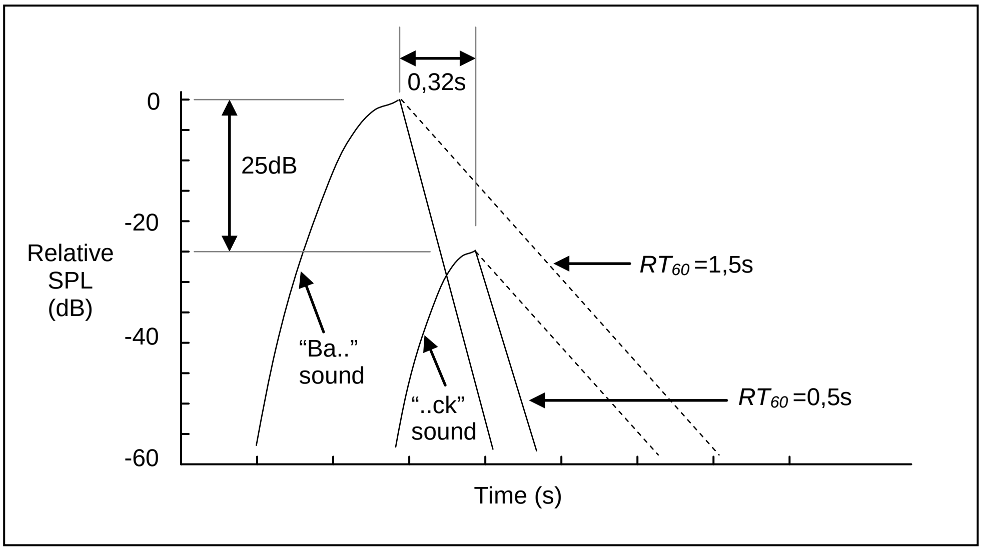
The ideal RT for the recording studio poses one of the most difficult tasks to determine. Various music genres play in the studio including different instruments that each requires a different RT. It is a common occurrence for a recording engineer to record different musicians in different rooms (or make changes to the room treatments if no other rooms are available). In such cases, the engineer may require each musician to be recorded on his/her own separate track and then after each one has been edited, the mixdown can take place (Everest, 1981). Another method would see the engineer manoeuvring different treatments around the different musicians within the same room, including adjusting microphone positions to enable the correct amount of reverberant sound. Artificial reverberation that is software generated assists in solving this problem as it has the ability to compensate for the room inadequacies. An experimental application of this feature is covered in chapter 5 of this dissertation.

2.5.3.1 Reverberation Time and Acoustic Intelligibility for Speech

Reverberation extends the excited sound for a period, which is room dependent. The key feature of room design for speech is the ability of the room to allow the speaker's content to be heard accurately. Using Alton Everest's (1981:147-148) simplified example, the influence of RT on speech can be explained as follows:

Words often have differing sound loudness syllables; for example, the word "back" consists of an abrupt starting sound of "ba" and ends with a softer consonant sound of "ck". Everest (1981) determined the average SPL and periods for the word and its parts. The "ck" sound is about 25dB below the first syllable and has its peak SPL at about 320ms after the "ba" peak. Figure 2.10 illustrates the decay layout for these two sound transients.

Figure 2.10: Decay layout for the word "back" with short and long RT conditions. (Extracted from Everest, 1981:148).



It is assumed that both "ba" and "ck" sounds decay at the same rate as they take place within the same room. Let us set the decay rate for the room to be 0,5s. From Figure 2.10, it can be seen that the peaks of both syllables can be heard individually and thus the word "back" would be heard clearly in this room. However, if the same test was taken in a room where the RT was 1,5s, then the dotted lines illustrate the new condition where the decay slope of the "ba" sound would mask the entire "ck" sound. Thus, the listener would hear "ba" clearly, but the "ck" part would be mixed-in with the reverberation of the "ba" and thus the intelligibility of the speech would be reduced. Without hearing the "ck" part of the word, it would be near impossible to identify the word without the sentence context to assist the listener, provided that the sentence too was not unintelligible. Thus, rooms designed for speech should have a lower RT than rooms designed for music.

A distinction needs to be drawn regarding speech used for conversation and speech heard in a musical context, commonly referred to as vocals. The reason for this distinction is that there seems to be different comfort levels for different types of voiced language. For instance, the acoustic properties of a school classroom should be tuned for optimal clarity rather than for artistic imagery. The art of deciding on how much reverberation to aim for is to some extent based on clear empirical science, which can be replicated and explained clearly, and to another extent based on psychological (cognitive) factors that are not as easily explained. This highlights the joint requirement of both scientific methods based on objectification as well as methods that are subjective and interpretive in nature. In this dissertation, both the subjective and objective evaluations were incorporated as part of the study.

2.5.3.2 Too Much Reverberation – is it Noise? Can it be reduced?

For reverberation to be enjoyable there seems to be an upper and lower limit. The context assists in defining whether it is noise or not, for instance, hearing reverberation on your cell phone can be annoying; however, if you were sitting at a live opera and there was no reverberation, one might feel that something was not quite right. Further, if you are sitting at the opera and there is so much reverberation that the audience cannot make out syllables then this may too become a problem. I think opera was not a good example as often one cannot make out the syllables anyway but I am sure the point was understood. If there is so much reverberation, it becomes difficult to decipher speech as discussed. In such a case, reverberation would be experienced as noise. Even music that has excessive reverberation could be experienced as noise; however, the tolerance with respect to music would be higher. We may now ask a different question: how much reverberation is good? This too was shown to be a difficult question to answer. It is like asking how much ice cream is too much. Different people have different comfort levels for their music and this is based on personal preference amongst music lovers and differs across music genres.

In a context that finds reverberation to be noise, how can this noise be attenuated? In situations where there is too much reverberation, many techniques have been applied to remove reverberation. These techniques are commonly applied to speech enhancement for mobile radio communication (Marro, Mahieux & Simmer 1998:241) rather than for music. Other examples include car telematics for GPS navigation control and Bluetooth voice car kits.

Marro et al. (1998) have shown in their study of de-reverberation techniques using microphone arrays with post filtering, that only a limited reduction in reverberation was possible. They were able to remove echo and localised noise quite well, however. In their study, they classified reverberation, echo and background sounds as noise, or put another way; anything outside of the original sound produced by the source was seen as noise. They believed that reverberation causes a degradation of the audio quality and that transmitting reverberation and the other noisy signals would reduce the sensation of naturalness and presence. They attempted to obtain the original “noise free” component from the recorded signal by use of microphone arrays and filter techniques. It has been proposed that by estimating the inverse of the filter corresponding to the acoustical path between the source and the pickup, one may obtain a de-reverberated outcome (Marro et. al.,1998: 240). This process is complicated and various technical factors have been highlighted. One such factor was found to be the need for accurate knowledge of all the filters and the requirement for them not to have common zeroes. It is proposed that if an exact inverse can be generated, the reverberation can be cancelled from the recorded signal.

A further method involves modelling the acoustical path as an all-pole filter and groupings based on vector mappings of the room transfer function. A study of the room response is thus required. However, these techniques still need to allow for the intelligence or the wanted signal to be relatively unaffected by the filtering mechanisms. This is difficult when the pass-band contains the wanted signal and the noise component. This is particularly true for acoustic reverberation as that component is considerably similar to the original signal in its frequency response.

2.5.4 Practical Considerations of Reverberation Time

2.5.4.1 Environmental Factors

Humidity, wind speed and temperature are some environmental factors that affect sound waves. High frequencies are highly dependent on humidity and attenuate at a rate higher than that of the inverse square law (Eargle & Foreman, 2002). This problem is more noticeable outdoors than indoors owing to the larger travel distances for the former, and thus poses a problem for outdoor setups that require long distance directional sound coverage. Temperature affects the velocity of sound with sound travelling faster in warmer air than in cooler air. Differing wind speeds affect the speed and directional perception of sound. For example, a sudden wind direction change can result in a distant sound not heard and then when the wind direction reverts, the distant sound can be heard again.

The relative humidity (RH) also affects the absorption property of air with dry air having a stronger influence. This absorption property is noticeable especially in large auditoriums. According to Nisbett (1995:42), an additional 3,5dB loss can be expected at 30m distance from source for frequencies at about 8kHz (RH at 50%). This effect may be further exacerbated if the sound balance is controlled from the back of a hall.

In general, the higher the frequency and RT, the larger the effect of air absorption; however, this relationship is nonlinear. Below 2kHz the absorption is negligible. In rooms where the RT is short, the RT reduction is minor and is usually taken as negligible in such measurements, especially if the room size is small (Eargle & Foreman, 2002:414). In rooms that have large RTs, the losses can be significant and steps need to be taken to accommodate this. The RH, wind and temperature values were noted for the reverberation study conducted in this dissertation owing to the stated effects.

2.5.4.2 Echoes, Feedback, SPL and Microphone Issues

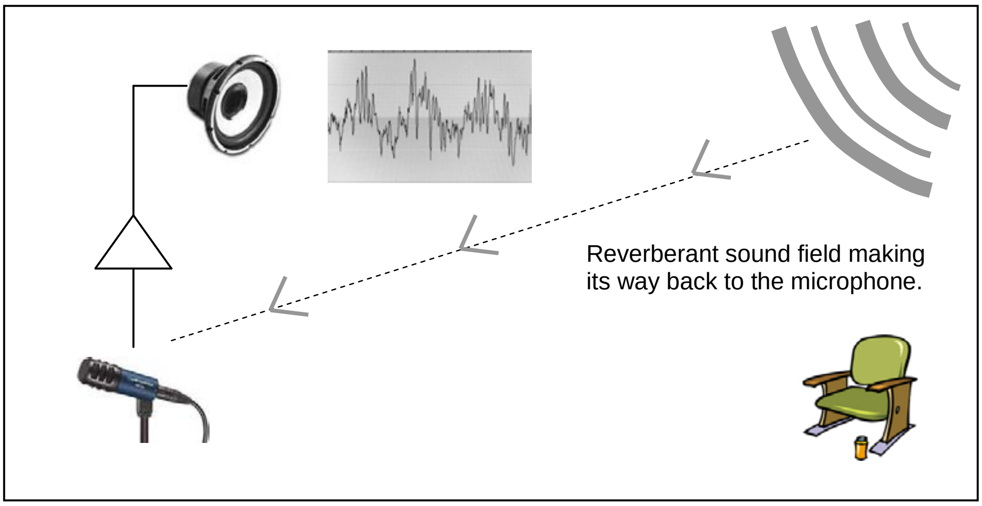
2.5.4.2.1 Echoes and Feedback

To obtain a smooth textured sound, it is important to have the reflected waves follow the direct waves with close succession and uniform attenuation. However, an acoustic environment that produces a noticeable echo would have a loud reflected sound following the direct sound, but with a magnitude high enough to stand out above the other attenuated reflected waves. The echo wave may have a constant frequency and have wavefronts that have constant amplitude normal to the phase velocity (plane wave). This is an undesirable effect.

With respect to an outdoor setup, when two loudspeakers are placed 30m apart and the listener is standing close to the one speaker, he/she will hear the other speaker's sound at a delay of at least 85ms. The sound from the far speaker will be perceived as echo. This scenario is one of the problems that face sound reinforcement engineers when setting up outdoor rigs (Eargle & Foreman, 2002:171). Reflections and echoes indoors are similar to reflections outdoors excepting that indoor setups have many more echoes.

In terms of sound reinforcement, the microphone's purpose is to pick up sound directed onto it and send it to the amplifier for amplification, which is heard through a loudspeaker. If the sound from the loudspeaker is picked up at the microphone owing to the system volume set too loud [and/or microphone/speaker placement incorrect], then this too gets sent back through the amplifier and so on until an annoying oscillation occurs [and the audience cover their ears]. Owing to reverberation's ability to increase SPL in a room, the potential for feedback in an indoor system increases (Eargle & Foreman, 2002:174). The microphone picks up the sound that acts upon it which in the case of a reverberant room would be higher; thus, reverberation effects need to be factored into the arrangement of an acoustic setup as shown in the next figure.

Figure 2.11: The increased possibility of feedback due to reverberation.



In a lively room, the reverberation may be high enough to uniformly increase the SPL throughout the room. A microphone anywhere in the room will pick up reverberant sound, which may result in feedback even in areas where there is not much direct sound (Eargle & Foreman, 2002:220).

2.5.4.2.2 Microphone, Balance, Scale of Width and Reverberation

In live orchestral performances, the amount of reverberation present is important in the perception of the stereo and the scale of width. If the reverberation is too dry or the audience is too close to the musicians, the perception may be that the musicians are cramped or the orchestra is undersized (Nisbett, 1995:88). The balance between direct and indirect sound is critical as well as the perceived *sound size* of the musicians, for example, when the whole orchestra plays together, obviously, it should be experienced as completely full sounding. However, when just the string quartet play, the sound image should be smaller or they will seem oversized (Nisbett, 1995:88).

With co-incident microphone reinforcement, the sound image and reverberation can be manipulated in several ways; for example, simply increasing the distance between source and microphone increases the reverberation. This technique was used to obtain extra reverberation for the lecture hall location that was used for one section of the listening tests (chapter 5 and 6). Adjusting the polar response pattern of the microphone changes its ability to pick-up the reflected sound from the hall and thus one can reduce/increase the reverberation, which in turn affects the sound image of the source. Changing the polar pattern from say cardioid to rear bidirectional will result in more reflected sound being reinforced.

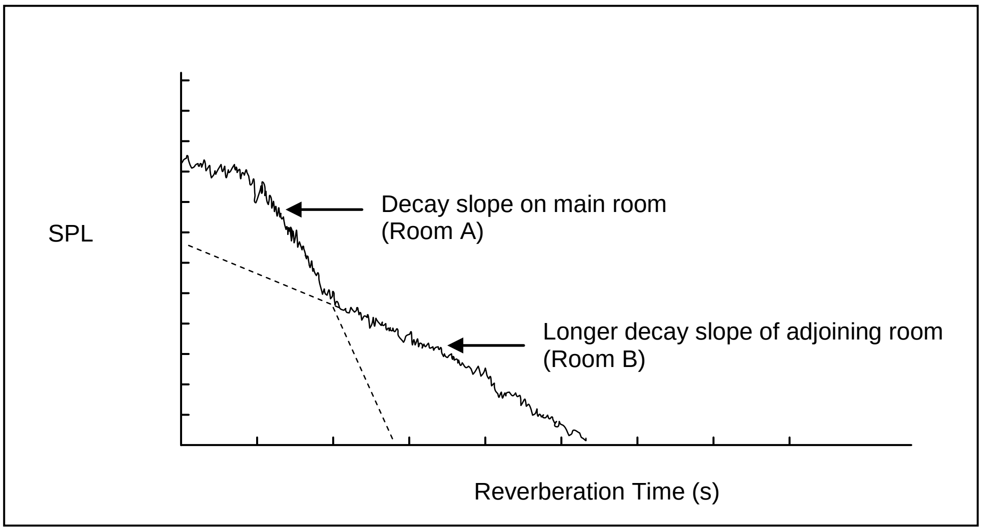
2.5.4.2.3 Pitch Change and Reverberation Decay

Reverberation has the ability to change the pitch of sounds during the decay slope (Everest, 1981). This effect has been attributed to two main factors, firstly, to the perceptual dependence of pitch on sound intensity and secondly to the possibility of energy shifts between normal modes (Everest, 2001:150). An older study performed by Balachandran (1976) who used fast Fourier transform, shows that it is the former rather than the latter (psychoacoustic) factor that causes this condition. However, this has been disputed and another school of thought believe that the pitch is in fact unchanged; harmonic blending of the sounds together are perceived as a pitch change.

2.5.4.2.4 Reverberation Time and Coupled Spaces

Using Figure 2.12 it can be seen that there are two distinct RT slopes. This often occurs when there is an adjoining room to the room under test. If the room under test (room A) has a faster decay, but the coupled room (room B) has a longer decay, the result looks like there are two distinct RT decays. This condition often occurs in public gathering spaces, homes and offices. If this condition occurs in an office room where employees are gathered around in a meeting, those employees who are sitting in the room (room A) and are near a coupled room (room B) may find it more difficult to understand the speech of those who are not near the coupled space. The reason for this effect is that background noise (reverberation) would be higher and speech intelligibility would be lower.

Figure 2.12: The RT slopes of two coupled spaces.



2.5.5 Performing Venues and Reverberation Time

Obtaining the correct acoustical properties in a hall, auditorium or recording studio requires an architectural study. This includes a look into the shape, enclosed area size, wall filling, boundary and interior materials as well as the placement of the source and audience. There are numerous variables; for example, the type of music, type of performance and style of music, which all affect the acoustic design requirement. The two main factors that influence the RT decay characteristic are shape and distribution of absorptive materials whether outside or indoors. For example, trees, buildings and the shape of the land all play a role in the reverberation response even though the sound is outside (Kang, 2002). A room may have a focusing effect if it is cylindrical in shape as the curved surfaces can focus sound to a certain area of the enclosure with the result of a variance in RT (Cox & D'Antonio, 2004). Each musical performance requires its own special brand of acoustic character. The RT and subcategories associated with it can be adjusted by various methods. Microphone placement techniques is one such crucial topic in this field. Sound treatment[8] is another key ingredient to obtaining the right sound. These two methods rely on physical changes in the environment as opposed to AR, which can be applied to the recorded sound directly.

A brief look into some of the features of performing environments with the focus being centred on reverberation follows.

[8] The use of sound-absorbing materials to give a room a desired degree of freedom from echo and reverberation.

2.5.1 The Recording Studio

2.5.1.1 General Purpose Sound Studios

Sound balance can be natural or artificial. Natural balance relies on the acoustics of the instrument and the recording/performing room for the reverberation effect. A single microphone can be used and on rare occasions can be obtained for more than one musician playing together (Nisbett, 1995:118). To obtain the natural balance of the instrument, the recording studio is also treated as an instrument. For example, a guitar has a sound, which is primarily composed of the reflections from the nearby wooden board. The shape, size and density of the board affect the radiated sound. Similarly, the natural sound studio has character and the reverberation from the natural studio needs to be harnessed from all these three dimensions (Nisbett, 1995:119). The reverberation obtained from the room can be manipulated with microphone choice and placement, but falls short of the artificial methods in obtaining reverberation.

Generally, speech studios require less reverberation than a music recording studio; however, there are exceptions, such as pop music has been found to be recorded well in a relatively dead room (Bartlett & Bartlett, 2009:3). In addition, some musical genres are composed knowing that they are to be played in the open air where there is little/no reverberation; such is the case with certain Eastern music types (Nisbett, 1995:41).

The majority of voice recordings are recorded in mono. The mono recording contains all the reverberation and colorations as well as the original vocals. If voice is recorded in stereo, the placement of the microphones become increasingly critical as some playbacks are then converted back to mono. This needs to be accounted for in the recording process. However, not all recordings would be heard in stereo even though they were recorded in stereo. A stereo recording played in mono has a different reverberation response with the mono response tending to have a dry sound product. A possible solution to this problem would be to reduce the reverberation balance in the stereo recording and then add more reverberation to the mono channel. Reverberation cancellation occurs frequently in crossed figure-of-eight balance (out of phase parts) which is one reason for motivating the use of hypercardioids or crossed cardioids, which exhibit a reduced effect (Nisbett, 1995:95). The loudness, aural selectivity and phase distortion are also affected in stereo to mono conversions.

2.5.1.2 Radio Drama Studio Setups

Many South African radio stations have radio drama programmes. These programmes consist of dialogue, music and sound effects that help the listener imagine the story. The studio is often divided/partitioned into “scenery” areas that have different absorption and reflection treatments. The aural perception of the recorded sound is a factor of the timing of the early reflections as stated earlier. Manipulating the timing of the first reflections can assist in changing the perceived size of the imagined room, but there are limits to this practice (Nisbett, 1995:109). It is difficult to obtain the exact acoustic properties of the imagined radio drama location; for example, obtaining the response of a large lively auditorium would require at least a large room that has a lively response. The RT value alone is not enough in obtaining the recording room’s aural perception to match that of the imagined room. While an empty lively room may give an RT of 1,5s, which may be similar to that of a large auditorium, the listener would probably be able to distinguish between the two. The room size exhibits itself in the type of frequencies that are reflected back and thus the RT may be similar, but the quality and reflected frequencies would differ highlighting the room dimensions (Nisbett, 1995:109-110). This phenomenon was encountered in the reverberation analysis (chapter 5) whereby the small room had a similar RT to that of the arts theatre yet they did not have the same sound character.

While it is accepted that the early reflections are important in the aural perception of the room, the quality of the reflected waves are also important. A room that is supposed to acoustically mimic that of a bathroom would need to have reflective surfaces and would still need to be small. Variety may be obtained by reducing the amount of reflective boundaries and so the "bathroom" may now sound like a "foyer"; however, there are limits and the story line also helps improve the imagery. Another technique is to manipulate the microphone source distance to obtain the aural effect of say a large reverberant room, but then shortly after the effect was noticed, revert back to a clearer sound as listening to long RTs for voice may become annoying. Artificial methods of reverberation can be used to reduce the physical changes that are needed. For example, instead of making changes to the actual studio by moving screens and placing reflective/absorptive materials in strategic places, the AR software can be set to adjust for these parameters directly. This saves time and reduces the physical equipment; however, a moderately anechoic studio is required for this method to be most effective.

2.5.1.3 The Dead Studio

There is an increasing requirement for "dead" studios, especially with regards to pop music (Nisbett, 1995:41). Nisbett (1995:41) feels that the studio acoustics play little or no part in the production of modern pop music. While there are acceptable norms in classical music genres, such as certain instruments playing at certain volumes and with specific tonal balance, pop music does not obey the same rules and may sample a piano or even a flute, which is expected to play at the same sound level to that of classical brass. The manipulation of the sound level required also means that the room should not allow the instruments to spill over into the microphones and are thus set for lower volume instruments. In setups such as these, reflected waves are seen as a nuisance (Nisbett, 1995:42). The use of AR provides the recording engineer with the sound he/she requires.

There are some disadvantages to dead rooms. One such problem is the effect of a dead room on a musician's experience of his/her own music. A dead room has the effect of attenuating the direct sound that impinges the boundary surfaces to such a degree that the amount of reflected waves are reduced somewhat. Musicians such as those who play string instruments often rely on the reflected waves as a marker for loudness and tone quality. In dead rooms, these musicians may play louder to compensate for the lack of reverberance, which may offset the internal balance of the total orchestra sound. Many musicians find dead rooms musically unpleasant (Nisbett, 1995:43). In terms of this study, the motivation is for moderately anechoic rooms, not completely anechoic, as the latter is expensive to construct. Further, almost all music (and speech) requires some reverberation. Thus, it is not necessary to create an anechoic space before AR can be accomplished.

2.5.1.4 Adjusting Acoustic Properties of a Studio

Sometimes different acoustic responses are required in the recording room. Some absorbing materials are fitted in such a way that they are movable and allow for varying the room's response. A useful tool is the screen, which is a panel that consists of wood with a layer of padding on one side. Screens can be used to separate instruments, adjust the high-end studio response and reduce ambient sound amongst other features. Screens can be used to reduce the RT of the studio room as a whole by reducing the path lengths between reflections within the room. They are also effective in reducing parasitic reverberation from nearby unwanted open recording rooms (Nisbett, 1995:46). Other treatments include porous materials that are applied to boundary surfaces and include, foam, rock-wool, carpeting and drapes. Soft absorbers work mainly in attenuating mid to high frequencies. They are not efficient in reducing low frequency sound (below 500Hz) (Nisbett, 1995:43). A common belief is that using carpets to treat all boundary surfaces will result in a room being sound treated. While carpets are not only a fire hazard, they do not provide adequate absorption of low frequencies. Another popular material is egg crates. In a test conducted by *Riverbank Acoustical Laboratories* on the sound absorption ability of egg crates, they found that below 350Hz the absorption coefficient of the crates amounted to less than 0,08 Sabins. Only at frequencies above 500Hz was there significant absorption, which unfortunately was not a linear result either; there were some higher octave bands that exhibited lower α values than that of lower octave bands (Gervais, 2006).

Room treatments require a careful evaluation followed by acoustic testing to determine if the theoretical predicted results were practically obtained. A room that has not met its design specification would then be assessed to determine what frequencies are problematic within the room. To reduce low frequency sounds "bass traps" are used. For example, to reduce 40-80Hz sound, one could use Helmholtz resonators, corner foam devices or the placement of panels across the corners of the room (Gervais, 2006). Another method relies on the movement of a membrane (often termed membrane absorbers) in a piston-like motion that is free to move with the low frequency sound wave thus absorbing some of the sound energy. These velocity absorbers can be suspended in space and hung a certain distance away from the wall for wavelengths calculated where the velocity of the wave would be a maximum.

The control room or mastering room of the recording studio provides the reference environment for the final sound. The reverberance of this room is a key feature and the monitoring speaker's response is of concern. While it is possible to remove the room's effect by using headphones, this is not always a comfortable option. Treatments for the control room are equally important as for the recording room. Treatments are not only confined to studio use and are also used in performing theatres and the like.

2.5.2 Performance Halls/Auditoriums

Great sounding halls do not necessarily have the same acoustic sound. There are variations between different halls even if some of their RT specifications are the same. There is also variation within the hall and while the sound differs in different locations, it may still be acceptable (Griesinger, 2007:131).

2.5.2.1 Does the Reverberation Time Predict the Sound Character of a Hall?

If two halls have the same RTs, do they sound the same? It is found that they do not. One of the simplest explanations as to why this occurs can be obtained simply from looking at Sabin's equation for the calculation of RT. There are two variables, V and α (or S). Different combinations of room volume and absorption may lead to the same RT. We know that sound emitted from a source travels in all directions and is reflected/absorbed at the boundary surfaces. While each hall has a different interior footprint, each hall would then have a unique absorption and reflection character, which means that the listener would hear a difference in a different hall owing to the changes in direct, reflected and absorbed sound even though the RT may be the same. The Sabin equation averages the absorption to obtain an average RT[9]. The intermingling of the direct and indirect sound waves allows the listener to hear the unique sound that the hall generates. Thus, with each hall having different layouts, the mix would differ from hall to hall.

2.5.2.1.1 <u>EDT and RT Variations within the Same Hall</u>

It is commonly found that average EDTs are lower than the RT_{30} values and the measured averages are lower than proposed theoretical design values. It has also been found that the RT and EDT may vary for different locations within a room. This is particularly true for the EDT. For example, Bradley (2005) conservatively found that changes of only 30cm could translate to 150ms differences in EDT. Other acoustical indexes such as sound strength and lateral fraction[10] too exhibit changes.

Conductors work with the response of the concert hall in order to obtain the sound they like by adapting to the hall's strengths and weaknesses. For example, conductors may ask the musicians to stretch out the endings of notes to enable a reverberant effect, or in one noted case, ask the violinists to play out of unison when the hall's response was dry (Beranek, 2004:3). Tactics like these allow the music to sound full, which is a key component related to acoustic reverberation.

The EDT is sensitive to room geometry and can vary considerably with changes in location of the listener. Early decay time is particularly sensitive to strong early reflections that reinforce sound in the first 100ms (Templeton, 1997:63). In the

[9] In the application/experimentation chapters, octave specific filters were used to determine the RT for each frequency band. This provides a RT for each octave, which is more revealing than an averaged RT.

[10] This is the fraction of lateral energy arriving at the listener and is compared to the total sound energy that arrives at the listener. It can be specified as early or late depending on the time interval used to measure the lateral energy (Templeton, 1997:64).

experimentation section of this dissertation, numerous physical positions were used to measure the EDT for the above stated reason.

2.5.3 The Anechoic Environment

An area without audible echo and/or reverberation can be both indoors and outdoors. Large open spaces that have no nearby trees and buildings can be seen as semi-anechoic. The higher the sound source is placed off the ground, the less the reverberation and echo. Thus, by placing the sound source on a high pole in an open wind free area, one can obtain sound recordings that have little reverberation present. Indoor anechoic rooms or chambers have a more complicated and expensive design requirement to achieve the same goal. Ultra absorbent materials are used on all the surfaces to absorb the radiated sound waves. The absorbent walls are often made out of thick wedges of fibreglass [or in some cases polyurethane] foam. The wedge shape traps sound waves by reflecting them onto another part of the wedge rather than back into the room. The foam itself then absorbs the vibrations (turning them into heat). The deeper the wedges are, the lower the frequencies of sound that can be absorbed. Suspended panels may be used underneath the walls and floor to further absorb the sound waves. Measurements taken in anechoic areas are called free field measurements, as there are little or no additional sound waves except the direct waves travelling from the source. It is rare to find a pure anechoic room. Semi-anechoic rooms are easier to find including open-air locations.

2.5.3.1 Open Air Acoustics

The reverberation is practically zero and although completely dead sounding environments are not favoured, it is sometimes unavoidable. For example, interviews with sportsmen/women on a cricket pitch or a soccer field would be one such instance. The microphone signal would be rather “dead” sounding and could be digitally edited to improve the reverberation to counteract the lack of environmental reflections. There are advantages with an outdoor recording with one advantage being the clarity with which the speaker can be recorded as there are few interfering reflections (assuming little background noise). The speaker also tends to speak louder in a “dead” environment.

As this study required a semi-anechoic environment, an open-air location was used for the “anechoic” part of the experiments. In such locations, the only significant reflections are from the ground plane.

2.6 Artificial Reverberation

A sound is considered artificially reverberated when the natural decay signature of sound is simulated using artificial techniques. The original natural sound setup differs from that of the AR setup. Earlier forms of AR include echo chambers, plates and springs. The older techniques have largely been surpassed by the advent of advanced analogue electronic devices and then later further surpassed by the development of

DSP in audio processing. To put this in perspective, the mechanical plate effect that consists of a thin steel sheet and tubular housing, which is tuned to give a *single* specific reverberation response can be mimicked by a DSP device. This DSP device can provide the “mechanical plate” sound as one of potentially thousands of individual reverberation sound characters. Nisbett (2003) makes a joke about this topic in his sixth edition of *The Sound Studio* where he says:

> You want reverberation? Just buy an artificial reverberation (AR) device put it in a channel and power it up. Do you want really long reverb? Turn the knob to, say 5,5 seconds. Maybe scan through the seemingly endless list of acoustic environments … (p:203).

Artificial reverberation is frequently used in recording studios that lack natural balance or where there are dead spots in the multi-microphone balance. Reverberation devices are also routinely used in reinforced music mixing. The term is sometimes incorrectly referred to as adding ‘echo’ even though one would not like to be left with added echo as the finished product (Nisbett, 1995:195). Mixing decks, DSP processing units, acoustic control systems and other devices can be used to manipulate the RT. The same applies to the RT of vocals where the natural passive reverberation from the performer would be called the natural reverberation, while AR would be the manipulated output that differs from the original signal. This is usually obtained by use of microphone level changes, polar pattern response changes and/or adjustments on the mixing deck. The AR method used in this study was software based and has its main application in the recording and mastering of audio, with particular focus on the recording studio use.

2.6.1 Further Benefits of Artificial Reverberation

2.6.1.1 Variable Reverberation

The RT is usually assumed to represent a *single measure* of an enclosed space’s sound decay. Many books highlight good RT for specific applications, for example romantic auditoriums should have an RT of at least 1,9s. This can be misleading as music is composed of many frequency ranges and the character of the sound may be better if certain frequency ranges have a longer RT than that of the other ranges and so forth. With artificial methods, one can obtain a variable RT for the complete frequency range, for example, classical music is usually played in venues with RTs in the range of 1,8s to 2,5s. However, it is also favourable to have longer RTs for lower frequency components such as RTs of 2,5s to 3s. By incorporating AR, a variable RT can be obtained across the frequency range (Nisbett, 1995:201).

2.6.1.2 Instantaneous Changeability

It is also possible to adjust the RT on the go, for example if there is an upcoming gap in the performance and only a single instrument or vocalist uses the gap. Their part can be enhanced by adjusting the RT during this softer sequence. When the full complement of the performers continue, then the RT can be reduced again. This is a tricky adjustment and requires a trained ear and knowledge of the on-going performance.

The perceived depth and volume of the room can also be adjusted by changing the time before the first reflected sounds are heard. In AR setups, this can be adjusted without making large-scale changes to the room (assuming a relatively dry room - one with low RT). Increasing the pre-delay (the delay before the reverberation is heard) also allows the listener to hear the full direct sound, including any transient parts that may have been masked by the early reflections (Nisbett, 1995:201). The reverse can take place whereby the reverberation can be brought closer to the direct sound. This would normally also change the distance and depth of the perceived sound field. However, manipulating the early part of the reverberation sound can still obtain the same perceived hall size; although the listener would feel that they were now closer to the sound source (Nisbett, 1995:201). Artificial methods can be incorporated to add differential reverberation. For example, a singer in an orchestra can have a shorter RT added than that of the instruments. The singer can also have different RTs in different parts of the song (Nisbett, 1995:145).

While the above is not impossible in terms of analogue methods, digital programs provide more control and functionality especially in terms of the equipment footprint.

2.6.1.3 Repeatability of Sound and Fine Control

Atmospheric conditions affected the reverberation output of some of the older physical devices such as the echo chambers (non-isolated – open to the atmospheric conditions) to such an extent that there was a different response when the weather changed (Nisbett, 1995:197). Electronic and digital methods are relatively unaffected by such environmental changes and provide repeatable acoustic responses. Digital methods also provide precise settings that allow for programmability of function.

2.6.2 Reverberation, Delays and Special Effects

2.6.2.1 Delaying the Sound

This is the process of playing a signal while at the same time storing the same audio signal in some form of storage medium, and then replaying all or part of the original signal after a pre-set time, either once or repeatedly. Delays can be long or short with many varieties such as *slap-back*, *multi-tap* and *ping-pong*. Presently, delay effects can be found in a stand-alone device or as a feature in software based recording and editing audio suites. Delay effects can be used to time-shift an audio 'block' forward or back to improve the total balance of the track and can also be used for changing the pitch of the audio sample (Nisbett, 2003:4).

The natural reflected sound has or should have a similar sound to that of the direct sound; for example, if the reverberation of a female vocal is heard and compared to the direct sound, one would be able to say that the reverberation was from the women's voice. However, if an effect (pitch change for example) was added to the direct sound and then applied to the reverberation, the reverberation although derived from the original voice is now at a completely different pitch. Other effects can be chosen with endless possibilities with the use of software based effects to such a degree that the very definition of reverberation would be difficult to apply. Special effects are often used in movies where strange voices are needed for the characters, such as in science fiction films. Nisbett (2003:213) describes how *Star Wars' Jabba the Hutt's* voice could be generated by use delays, pitch changers and the addition of sub harmonics to create his hollow alien voice. While the special effects team in a motion picture would welcome this effect, the recording engineer and musician may cringe at the sound of such extreme effects. This highlights a problem in terms of artificial methods of reverberation: that of overuse. As easy as it is to create and manipulate the reverberation of music and speech, it is just as easy to destroy the integrity of the sound whereby it becomes unusable. In practice the effects are often overused (Nisbett, 2003:213).

2.6.2.2 Reverberation and Colorations

Natural reverberation relies on reflections from the room's boundary surfaces. With each room being different from the next, some frequencies are absorbed more than others and the frequency mix and individual decay characteristic differs. A situation occurs where the reflections of certain frequencies persist while others decay; this phenomenon is called coloration. Natural and artificial reverberation are susceptible to colorations with the latter sometimes being used to supplement the former. The procedure is not simply used to fill in the gaps. Rather, some coloration is important especially with regard to musical instruments, as they often rely on the very fact that certain frequencies will stand out above others. Thus, a careful study of the natural reverberation response is required before supplementing it with AR (Nisbett, 1995:196). In terms of this study (particularly for chapter 6), applying AR across the entire audio sample will not simply solve the problem of inadequate reverberation. The audio sample needs to be assessed to determine parameters such as pitch, graphic equalisation, loudness and timing of the additional reverberation. This further highlights the point stated earlier that two audio samples with a matched RT and EDT do not necessarily sound the same. The colouration of the sound needs to be addressed as well. If one is attempting to match a predefined sound as was accomplished in this dissertation, the task can be complicated if one does not have a good ear for these changes.

2.6.2.3 DSP Solutions to RT Adjustments (Electro-Acoustic Solutions)

2.6.2.3.1 Signal Processing and DSP

Signal processing components are used to modify sound and usually takes place between the mixer [or in the mixer] and power amplifier. Signal processing components include graphic equalisation, filters – high/low/band, limiters and delays. Analogue signals are converted into a digital format. The processing that takes place, which is based on digital, ultimately binary representations is termed DSP. Microprocessors have the ability to process multiple signal components with several processed signals converted to a single digital component.

The most popular DSP solutions make use of microphones, speakers and amplifiers to counteract the total average absorption within a room. Digital signal processing methods have been used successfully in churches, performing theatres and concert halls. Electro-acoustics have been used successfully in the equalisation of sound as well as for delay effects. The main advantage of the electro-acoustic method is that it allows for variable RT, and frequency dependent ranges can be applied. For example, increasing the RT in the lower frequencies by adding late reverberation is favourable in concert halls. However, it may not be that simple, as some critics have said that there is an unfortunate side effect of the sound becoming muddy. New techniques have been introduced to address this side effect (Griesinger, 2007:130). A possible solution would be to mic the musicians at a close range and allow the DSP to generate the controlled reverberation with frequency dependent RTs. Griesinger (2007:130) talks about one such case with respect to a concert hall. In this case a DSP solution was applied without official announcement to the public who were subsequently very impressed with the conductor’s ability to get the sound right. The hall was the *Berlin Staatsoper* and the main effect was an increase in the RT from 1,1 to 1,8 seconds at 150Hz, which resulted in improved blend and envelopment without effecting the sonic distance of the vocalists and soloists (Griesinger, 2007:130).

2.6.2.3.2 The Transversal Filter

The Fourier mathematical method allows for a variety of features that can be applied to audio signals owing to the unique relationship between the complex frequency response (amplitude and phase) of a sound impulse and the corresponding time domain response. The analogue sound can be digitised and equalisation treatments can be accomplished within a wide range of values. For example, delays can be introduced merely through software tools rather than hard wired electronics. Digital methods allow for accurate and easy equalisation, which can be accomplished by measuring the impulse response of the desired filter in the analogue domain. This is mapped onto a set of coefficients in a transversal filter in order to obtain the equivalent response in the digital domain. There are numerous advantages when working in the digital domain over the analogue domain. The incorporation of software to adjust the RT provides further range of control, which is one of the motivations for this study.

2.6.3 The Digital Audio Workstation

The digital Audio Workstation (DAW) uses software and a PC that is connected through an audio interface such as a sound card. There is a reduced infrastructure footprint as it is possible to have a single DAW for recording, editing and mixing of audio, which all take place in the digital domain. The main components would be a PC that has large memory capacity both random access and magnetic storage. One may think that the PC would need to have a high specification like a server computer, but even a laptop can be used as a DAW, although this would not be recommended. The reason that laptops are not recommended is that large monitoring screens make it easier to use the software, which was the case in this study where on-site tests were recorded using a laptop but transferred to a larger desktop for further processing.

The main advantage of the DAW is that one system can replace numerous complex hardware components. As software has great expandability, “plug-ins” can be imported into the software, which now takes the place of the hardwired electronic devices; thus, one could have an abundance of acoustic related tools that are all software based. There is an increase in this trend in recent times with many recording engineers enjoying the benefits of the digital methods available in the DAW.

A distinction can be drawn by the way the software provides AR. There are software packages that adjust the parameters in real-time, which are used mainly for live performances, and then there are those that are used in the editing of already recorded sound. For example, one recording software package that can be used in the editing of audio is *Cool Edit Pro ver. 2*. It has 27 built in reverbs and the ability to load new plug-ins for additional ones. These reverberations are adjustable and have numerous settings but are only applicable to sounds that have already been recorded. This software package has been studied in more detail in Chapter 5 and 6.

It is noted that a common problem exists in the field of music recording whereby the same settings may be input into different manufacturer’s programs or DSP units yet they produce different results (Miller, 2007). The main reasoning behind this is that the different manufacturers define their parameters in a different way with both being acceptable. Added to this, many recording engineers and musicians do not have the knowledge of the inner workings of DSP methods or a comprehensive understanding of electronics. It is common to find a recording engineer and/or musician stay with a certain device or manufacturer and avoid change purely because they feel they cannot get the same sound from another device or manufacturer (Miller, 2007). In the final chapter, this problem of lack of universal standards across the manufacturers and software companies has been recommended for further study.

2.7 Psychoacoustics

It is commonly thought that sound is mainly associated with the ear and that sound can be explained in terms of physics, and the process of hearing can be narrowed down to mechanics. Research has repeatedly challenged any objectification of the listening experience. A landmark study challenged the idea of there being a fixed upper frequency limit of 20kHz to human hearing. Tsutomu and colleagues (1991) set up a listening experiment where they played back a recording that had active frequencies up to 60kHz. They set up a speaker system and included an independently powered tweeter that was able to excite frequencies above 26kHz. The tweeter was switchable to be on or off during the test. An electroencephalogram (EEG) was incorporated as part of the listener’s response data. The finding was that the subjective evaluation of the music played was altered by whether the high-frequency tweeter was turned on or off as well as changes to the EEG were noticed (Tsutomu, Emi, Norie, Yoshitaka, & Hiroshi, 1991).

While most music is designed for our hearing range of 20Hz-20kHz, many musical instruments have a large portion of their vibration energy well above our hearing range; such as the cymbal instrument which has ±40% of its energy between the range of 20kHz and 100kHz (Boyk, 2000). In terms of everyday appliances, certain noisy electronics are designed to have their oscillation noise above 20kHz so that they do not annoy us. The switched mode power supply is one example. If the frequency of the oscillating driver circuit is reduced to 8kHz it can be most unfavourable for the nearby person's auditory system. This raises interesting questions as to the effects of unconscious sounds and the experience of listening and attitude changes. The study of *subjective* human sensation and perception of sounds is termed psychoacoustics and is described as the study of the psychological correlates of the physical parameters of acoustics.

As the ear is only a part of the hearing chain, audio processing requires a study of the hearing system, the integration of this system and the brain as well. For example, the inner ear does a large amount of signal processing in its ability to convert sound waves into neural stimulus. Audio compression methods such as MP3 (derived from MPEG-1 Audio Layer 3, moving picture experts group) take advantage of this signal processing and remove the wave data that was imperceptible to our hearing. This is to achieve a reduction in file storage space (Ahlzen, & Song, 2003:510). The MPEG group have developed several successful methods for dealing with audio data, which are psychoacoustic based (Pohlmann, 2005).

In terms of music appreciation, Zhang (2005:20) believes that two main categories are important with respect to psychoacoustics: these are temporal and spatial factors. The time at which a given set of musical sound is acquired and related to other sounds and their respective time of hearing relates to the temporal category. The structure and reflection pattern of sound in an enclosed space is a central component in temporal factors. In terms of noise, different sounds mean different things to different people. The sound of a loud engine exhaust may be music to the ears of the drag racer but noise to his/her neighbours. Often its one's attitude to the sound that can determine if one would call it noise or not. Unwanted sounds do not have to be loud before they annoy us. The loud crash of thunder can be as annoying as a creaking floor that is only a fraction of the sound level. Thus, psychoacoustics forms a critical part of the study of sound and hence has been incorporated into the experimentation section of this dissertation.

2.8 Conclusion

Reverberation is strongly related to its environment and thus environmental factors were discussed in this chapter. The impact of computer technology has provided the acoustician/performing artist with a variety of options. While the recording engineer may still prefer the physical control interface of real knobs and buttons, he/she is not confined to that option. The software based audio suites offer almost complete control from a single computer station, namely the DAW.

Reverberation has objective factors that can be studied in the same way as physical engineering phenomena such as mechanics, whereby empirical and objective testing can take place. On the other hand, there are parts to this subject that rely on subjective appraisal and are not so easily differentiated. Psychoacoustics is an important part of acoustics and should be studied to gain a holistic view of the field.

This chapter reviewed the foundations of RT. The next chapter tackles the topic of RT measurement.

Chapter

3

3 THEORY OF REVERBERATION TIME MEASUREMENT

3.1 Introduction

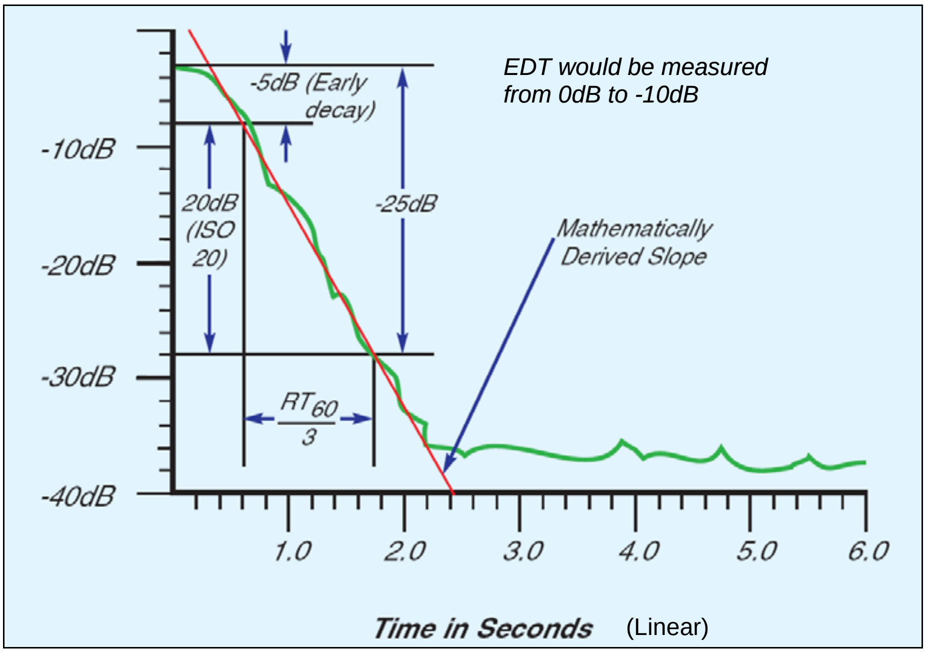
This dissertation had two goals: the first goal was to ascertain and compare the RTs of four different locations by use of impulse tests; the second goal was to manipulate an almost anechoic sound (impulse sound and vocal passage) to mimic each of the three other livelier recordings that took place in their respective naturally reverberant locations. The goal was to show that by using AR, one could match the almost anechoic sound, which was subjected to AR, to the naturally reverberated sounds. This was evaluated by objective methods (Chapter 5) and by subjective perceptions (Chapter 6).

In order to achieve these goals, precise measuring techniques were required. This chapter is devoted to the theory of RT measurement as put forward by the ISO standards. Recently a new standard was published (ISO 3382:2008) which supersedes the 1997 standard. A review of this new standard including the changes from the old standard was also conducted in this chapter. The SPL meter too was studied as it provides the sound level measurement that is required to measure RT. This chapter stands in as a theory of measurement for the next two chapters.

3.2 Measuring RT in Terms of the ISO Standard

To meet the definition of RT_{60} in its entirety, a full 60dB sound decrease would be required (actually a 65 or 70dB decrease in order to apply a noise floor). This means that an increase of say 65dB above the background noise level is needed before the sound is stopped. Depending on the environment, background noise could be anything from 35dB to 55dB. Therefore, for an increase of 65dB from 55dB one would be near the threshold of pain, and thus adequate hearing protection would be needed. However, even when one can obtain a loud source, the impulse level may overload the acoustic analyser and/or its software ranges. Some analysers have an attenuator attachment that can be inserted between the microphone and the preamplifier to reduce the SPL by 20dB (Brüel & Kjær, 2008a). This does not affect the reverberation measurement as relative values are used to obtain the reverberation graph. Obtaining specific impulse levels may also pose a problem, for example, if one would like a value of 110dBs and currently can only get 107dBs, then one would need to double the input power into the source. The reason for this is that in obtaining an increase of 3dBs for the same distance between source and recorder, a 100% increase in power is required. It is usually not possible to obtain a full 60dB decay; thus, ISO makes provision for a 20dB and 30dB decay measurement. The decay slopes are then mathematically extended to provide an equivalent RT_{60}.

Figure 3.1: Reverberation measurement for RT_{60} derived from RT_{20} (Ivie, 2007:3).



The RT_{20} or “ISO20” measurement starts at -5dB and ends at -25dB. This represents 1/3 of the full RT_{60} and thus needs to be multiplied by three. The ISO standard requires the reverberation measurement to be taken over specific sections of the decay slope. If there is sufficient decay available, then the “ISO30” or RT_{30} can be used which would require a starting point of -5dB and an end point of -35dB. To obtain the early decay characteristic, a short measurement is taken from the instant of the cessation of the sound excitation until a drop of 10dB has taken place. The EDT can be obtained using the best-fit linear regression as with RT_{30}, except it would be taken over a shorter period. Figure 3.1 shows a straight red line, which is the line that best represents the reverberant decay slope of the excited sound decay. The decay response is plotted on a logarithmic scale. This helps in obtaining a linear roll-off, however obtaining a straight decay slope for the full roll-off does not always occur. An approximation is often used by choosing the linear part of the decay response. Measurements are made in narrow bands (octave or 1/3 octave), rather than broadband (20Hz to 20kHz).

3.2.1 The Current ISO Standards

The three main standards pertaining to reverberation are as follows:

- ISO 3382:1997[11], Acoustics - Measurement of the reverberation time of rooms with reference to other acoustical parameters.
- ISO 3382-2:2008, Acoustics - Measurement of room acoustic parameters - Part 2: Reverberation time in ordinary rooms.
- ISO 3741:1999, Acoustics - Determination of sound power levels of noise sources using sound pressure - Precision methods for reverberation rooms.

At the time of writing, ISO were busy releasing two new standards (ISO 3382-1 and ISO 3382-3) that will cover the measurement of acoustic parameters. These two standards are scheduled to be released in 2009. The ISO 3382-1:2009 will provide methods for the measurement of RT and other room acoustical parameters in performance spaces incorporating digital measuring techniques. It will cover similar topics as set out in the ISO 3382:1997 document. The ISO 3382-2 standard sets out to provide a methodology for the reverberation measurement in ordinary rooms. The ISO 3382-1, ISO 3382-2 and 3382-3:2008 replace the ISO 3382:1997 standard.

3.2.2 The New ISO Standard and its Focus

The previous ISO document for the measurement of reverberation is ISO 3382:1997. This standard has now been surpassed by ISO 3382-2:2008. There are some changes introduced by the new standard 3382-2:2008. One such change is the focus on the RT measurement range of 20dB rather than 20dB and 30dB. In laboratory setups, it may be easier to obtain 30dB measurement ranges [which thus require a 45dB signal level over background noise[12]], however in field measurements with higher ambient noise levels, a RT_{20} is preferred. Added to this, there are well known findings that the subjective perceptions of reverberation are highly influenced by the early roll-off characteristic of the decay slope (ISO, 2008:v).

Standard 3382-2:2008 recommends that RT should be undertaken in rooms that have two or less people and thus the newer standard differs from the 1997 standard, which allowed for three different occupancy states. The newer standard has noted the effect of higher frequency sound absorption in air and set out two parameters regarding the effect thereof (ISO, 2008:2).

3.2.3 Measurement Methodology

The ISO 3382-2:2008 standard focuses on two methods, namely; the interrupted noise method and the integrated impulse response method, which remains unchanged from the previous standard. The amount of measurements and the extent of the frequency range required depends on one of the three methods used, that is, *survey*, *engineering*

[11] The ISO 3382-2:2008, together with ISO 3382-1 and ISO 3382-3, replace ISO 3382:1997. However, at the time of writing, neither ISO 3382-1 nor ISO 3382-3 were published and thus the ISO standards used in this study were ISO 3382-2:2008 and ISO 3382:1997.

[12] The ISO standard specifies a noise buffer of 10dB (ISO, 2008:5).

or *precision*. For example, the engineering method requires a frequency range of at least 125Hz to 4kHz, while the survey method requires a top frequency of only 2kHz (ISO, 2008:5). Measurements can be taken in 1-octave bands or 1/3-octave bands while the signal bandwidth needs to be greater than the octave band chosen. A flat signal output needs to be obtained for each octave band, or a pink noise signal can be used (ISO, 1997:6). The common frequency ranges are between 100Hz to 5kHz, which applies to both interrupted and impulse method.

3.2.3.1 Measuring Positions

There should be enough measuring positions used to cover the area of the room with the microphone position at least half wavelength apart (minimum 2m). The nearest reflecting surface should be at least quarter wavelength away. The sound source and microphone should be set up in such a way that direct sound should not be too dominant (ISO, 2008:4). The following equation for calculating the minimum distance between source and microphone can be used:

$$d_{\min} = 2\sqrt{\frac{V}{c\hat{T}}} \qquad [3.1]$$

Where

$\hat{T}$ is an estimate of the expected reverberation time, in seconds.

The newer standard differs from the older standard in that it specifies three methods of assessment for the room under test. Each of the three methods has been allocated the number of decays required, the amount of microphone positions, source positions and source-microphone combinations (ISO, 2008:4). The previous standard did not delineate a precise method but rather a visual interpretive analysis that relied mainly on the experience of the measuring engineer to locate "good" positions. For completeness the older standard's method has been briefly summarised next as it is still widely used.

When choosing microphone positions to measure RT, one should try to obtain positions that highlight the room's inconsistencies, such as areas near a wall, under a balcony or where there is a change in material in the area (furnishings or boundary). Locating suitable positions can be done by visual inspection taking note of where listeners would be present in the room. The ISO 3382:1997 standard recommends that the acoustician take into account the boundary surfaces and note the absorptive and diffusion properties of the surfaces including suspended surfaces. At least three measurements in the seating area should be taken for a room that has relatively consistent room volumetric dimensions and these measurements should allow for repeatability of results (ISO 1997:5, 6).

In this study, when conducting the RT of the performing arts theatre as shown in the next chapter, the above techniques were used, as they were beneficial in quickly locating the best measurement positions.

3.2.3.2 Recording the Reverberation Decay Slope

To obtain the RT of a reverberant area, one needs the following equipment:

- controllable sound source that can excite the reverberant area;
- audio pick-up device such as an electret microphone (omnidirectional and conform to IEC[13] 61260;
- real-time sound level meter/analyser that can record the decay waveform;
- timing device (usually part of the analyser's software); and
- filter network to apply the octave bands, either software based or hardwired electronic.

3.2.3.3 Steady State Sources

Sabine used a windchest and organ pipes to obtain his measurements. The reason for this may be that these instruments do not give consistent frequency sound waves. These instruments vary their intensity in different frequency ranges. It has been found that consistent sine wave sources give highly irregular reverberation decay shapes, which are difficult to compute into a single RT (Everest, 1981:138). Thus, by warbling a tone on a windchest [as Sabine did] results in rapid modulations of frequency within fixed limits around the basic pure tone frequency; thus, one can distribute the source energy into slightly different frequencies. Random noise sources provide a good way to obtain equal energy owing to its ability to provide uniform power spectral density at every frequency in the range of interest. A white noise generator is such a device and exhibits an aperiodic signal with equal loudness levels at each frequency [analogous to white light containing all wavelengths] (Watkinson, 1998:21). Pink noise on the other hand is obtained by filtering white noise. The choice of whether one uses a white noise or a pink noise source depends on the type of analyser one has. If the analyser records at fixed frequency gaps, say pass-bands of 100Hz, then white noise can be used as a flat response can be obtained at each measurement range with the frequency bandwidth remaining constant. If on the other hand one uses octave filters, pink noise is preferable as the frequency gaps increase for each additional octave filter; said another way, the bandwidth increases as one climbs the frequency range. For example, a 1/3-octave filter with a centre frequency of 100Hz would have a passband of 23Hz (89-112Hz) but if one moves up the frequency range, a centre frequency of 1kHz has a passband of 229Hz (891-1120Hz). Thus, pink noise would provide a flat response when comparing one passband to that of another even though the percentage bandwidths are the same, but the actual Hertz bandwidths are not the same.

3.2.3.4 Measurement Repeatability and Room Mode Decay Variations

One of the reasons for using pink noise to excite a room is that it exhibits a random response. For example, if pink noise is viewed on an oscilloscope, one would find a large amount of "noisy" waveforms moving unpredictably across the screen. However, if the pink noise is filtered by say a 1/3 octave filter, then one would see a sine wave that

[13] International Electrotechnical Commission

is continuously changing amplitude and phase. This shifting amplitude and phase has an effect on the modes that are set up in a room and thus can affect the reverberation decay slope characteristic (Everest, 1981). A room consists of axial, tangential and oblique modes, which tend to be quite close in frequency within one octave band. The modes have different amplitudes and are dependent on the room design. When using pink noise to obtain the RT of a room, the sound source is energised. At the point in time when the pink noise starts, random waves are coupled to the air and standing modes may be set up relative to the room dimensions. When the pink noise starts, many modes would be set up. An explanation of the possible effect of this is explained next: Two modes will be discussed, call them mode A and mode B. These two modes were energised simultaneously but because pink noise is random, mode A may be re-energised in the next time instant while mode B starts decaying as it was not re-energised. The process is entirely random. Another possibility is that both modes would be successively energised over more than one period. When the pink noise is suddenly stopped, with two modes and two periods, 16 different sound combination decays may be recorded. For example, if both modes were energised during both periods, then the decay slope may be a bit longer as the modes were both in re-energised phase just prior to pink noise being terminated. Thus, the decay may be longer than if neither mode were energised during both periods. The question arises as to what total modal layout would be present in the room at the moment of cessation of pink noise (Everest, 2001:141). Owing to pink noise's random nature, so too is the modal layout of the room after the sound is interrupted; thus, the decay slopes often have some differences in their layout. The frequency bands that are used when taking the RT measurement are also a factor. The lower octave bands are more susceptible to modal activity and thus there would be more variation in the RT decay spread of a few successive measures. Reverberation decays of the higher octaves tend to be duplicated without much difference though (Everest, 2001). Consequently, there is a requirement to take a set of decaying waveforms to reduce any uncertainties particularly for the lower octaves. When taking successive RT measurements, it is important to keep the setup the same or else new variables would be introduced.

3.2.4 Using the Interrupted Noise Method

Excitation of the room under test is accomplished by an omnidirectional loudspeaker source that is fed with a broadband random signal. The measurements should only be undertaken after the signal has reached a steady state before sound interruption. The excitation signal should be radiated for a minimum period of $T/2$ seconds and a few seconds for larger rooms (ISO, 2008:6). This method was not used in the RT study found in the next chapters and thus the detailed explanation is omitted. The Impulse response was used and is explained next.

3.2.5 Using the Integrated Impulse Response Method

3.2.5.1 Exciting the Room

There are many devices that can be used to provide an impulse for the RT measurement, for example a blank pistol shot[14], balloon or a spark-gap device amongst others. For best results, the sound source should be able to excite the room in a wide range of frequencies. To obtain the RT of the low frequency components in a large room, a high-energy discharge device may be required. Everest (1981:136) talks about the use of a canon to remedy this problem. The impulse device should not provide a reverberant sound itself and the frequency range should be wide enough to satisfy the octave measurements (at least 4 octaves wide).

3.2.5.2 Integration of the Impulse Response

The new standard provides less information regarding the backward integration method and refers the reader to the ISO 3382-1 standard for more details. At the time of writing, the latter standard was not yet available and thus the following method is from the ISO 3382:1997 standard. Backward integration (Schroeder method) is applied to the squared impulse response that is obtained for each octave band.

In a treated room where the background noise is very low, ISO 3382:1997 recommend that the integration start at the end of the impulse response. Following that, proceed to the beginning of the square of the impulse response shown below in its common format:

$$\int_t^{\infty} P^2(\tau)d\tau = \int_0^{\infty} P^2(\tau)d\tau - \int_0^{t} P^2(\tau)d\tau \qquad [3.2]$$

Where P is the impulse response.

3.2.5.3 Evaluation of Decay Curves for the Integrated Impulse Response Method

The integrated curve is analysed over a minimum of a 20dB range starting from 5dB below the total level. Larger ranges are ideal and an attempt should be made to obtain a 30dB analysis area. A least-squares fit line is used to obtain the slope for the RT.

3.2.5.3.1 <u>Measuring a Short Reverberation Time</u>

The measurements of very short RTs are limited by the filter and detector response. Noise in the tail end of the decay slope can be problematic for standard BPFs. For example, an impulse response can give rise to the distortion of the decay curve if there is ringing in the tail end. Lee and Ho (2002), recommend that the product of the 3dB

[14] A blank pistol can be used but has a limitation in that it does not provide the exact sound impulse repeatedly as well as suffering from non-linear effects close to the gun. In such a case where the impulse response is not exactly repeatable, more measurements should be taken and averaged accordingly. Using a loudspeaker as an impulse device is also not recommended unless further signal manipulation takes place such as if many pulse responses are synchronously averaged (ISO, 1997:16).

bandwidth of the filter and the RT of the system be at least 16. However, if the RT is short, it is not always possible to use standard BPFs and they proposed the modified wavelet transform, which can interpret the decay slope for lower bandwidth RT products. If the minimum product of reverberation and bandwidth time are less than 16 an increased distortion measurement tends to occur, as the product value is lowered owing to the influence of the detector and filter (ISO, 1997:9). The ISO recommendation is as follows:

$$BT > 16 \quad and \quad T > 2T_{\mathrm{det}} \tag{3.3}$$

Where:

T_{det} is the RT of the averaging detector

T is the measured RT

BT product of filter bandwidth and RT

If a test room has a very absorptive layout, the exact RT is difficult to measure. In low-coverage measurements, the limits can be reduced to (ISO1997):

$$BT > 8 \quad and \quad T > T_{\mathrm{det}} \tag{3.4}$$

According to Lee, if the wavelet filter[15] bank is used, then it can be utilised from a product of bandwidth and RT of four (Lee, 2003).

Listening rooms often have short RTs as they usually contain carpets, curtains and couches. Obtaining the RT roll-off in the low frequency band using a third-octave filter has been found to be problematic when the RT is low (Lee, 2002). In the next chapter, the RT was conducted for an outdoor area that was semi-anechoic. Areas that have a very low RT require an assessment of the recording equipment to confirm if they have the ability to sample and filter rapid sound decays. The recording equipment used for the semi-anechoic location was assessed in order to meet the ISO standard and is discussed further in the next chapter.

3.2.5.4 Spatial Averaging

The results of the RT measurements can be obtained for separately identifiable areas within an enclosed space or for the room as a whole. In the case of a large auditorium, this is an important test. In the next chapter, the RT and EDT were undertaken for three separate areas within the performing arts hall. The reason for this is that people sitting in different areas of the hall would be subjected to different RTs.

Two procedures for spatial averaging are presented by the standard:

[15] The wavelet filter performs a wavelet decomposition on a signal.

a) Arithmetic averaging of all the RTs: all the RTs used to obtain a mean RT for the room for all source-microphone positions. The standard deviation and variance are then calculated.
b) Ensemble averaging of the decay curves: the individual decays superimposed onto each other with their starting points synchronised.

In terms of the recording location chosen in this study, only the small room and the lecture hall used the arithmetic averaging method to obtain the spatially averaged RT and EDT. The arts theatre had three separate spatially averaged RTs and EDTs to provide a location specific RT and EDT. This is a common practice when measuring the RT and EDT in halls as the acoustician would like to obtain the specific RTs for the different areas in the hall. Using this information, he/she can then determine if there are major differences and these differences can be attended to.

The ISO 3382-2:2008 standard provides additional information on the uncertainty of measurement. They have also included a least-squares fit method as well as two proposed methods for non-linear decay curves.

3.2.6 Presentation of the Results

The RTs at each frequency should be presented in tabular format and can include graphs. The graphical method should conform to IEC 61260 in terms of the implementations of BPFs or spectrum analysing. The results should specify the RT range, that is, RT_{20} or RT_{30}. A test report should accompany the data measurements. The report should have a batch of information including room volume, condition of furnishings, room layout, temperature and RH, degree of precision, type of sound source and other technical information. The next chapter provides an exhaustive example of this data and was included for each of the RT measures undertaken.

3.2.7 Measurement Devices

Early approaches relied on recording devices that would trace the received SPL that was recorded, including its decay slope. Some of these devices proved problematic owing to the limitation of tracing speeds of the printing devices and now better methods have taken over. Currently, there are hand held acoustic analysers that have the ability to determine the RT and interface with a PC where software tools can be applied to the results. To obtain the RT, an SPL measurement is firstly required. Thus, sound loudness is a key feature in the measurement and analysis of RT. Owing to this reason, the SPL meter will be discussed in the next section.

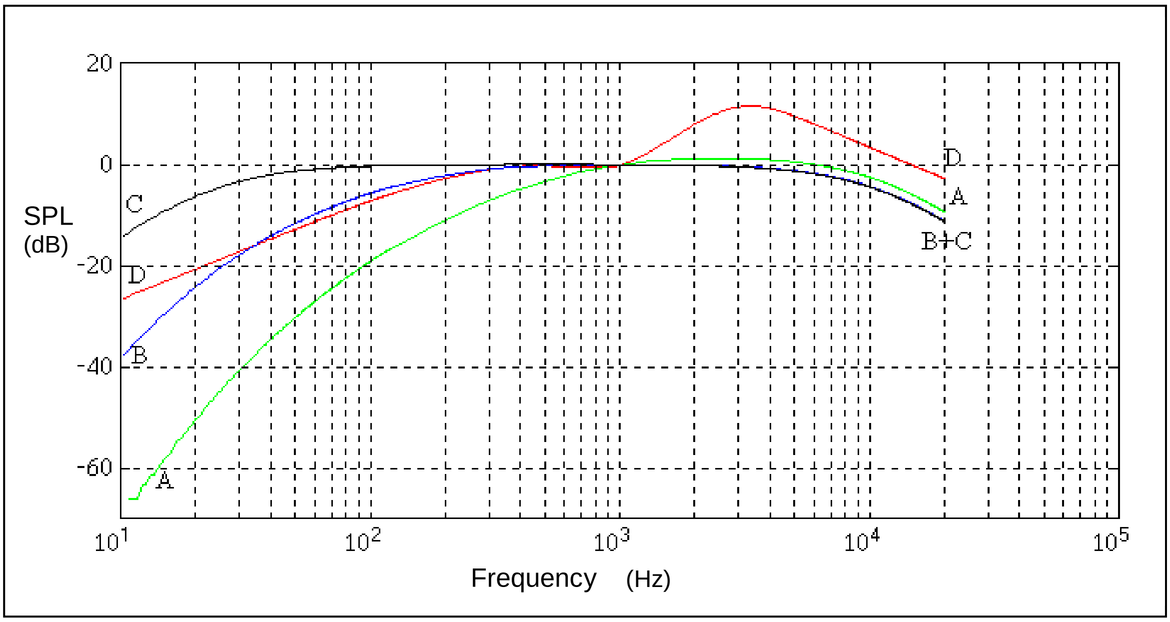
3.3 Measuring Sound Pressure Level

3.3.1 The SPL Meter

The device usually consists of a condenser microphone that converts the sound energy to an equivalent electrical signal. The condenser microphone is very sensitive and has a large audio frequency response. A pressure sensitive diaphragm forms one plate while the back plate remains fixed. A polarizing voltage of 50V to 100V is required. Vibration of the diaphragm causes a change in the capacitance with the result of a current change noticed in the load (Aldred, 1972:95). A condenser type microphone called an electret microphone is now frequently used. This type of microphone uses a material that produces a constant electric field without a dedicated power source and can be seen as an electrostatic equivalent of a permanent magnet (Watkinson, 1998:152). The name is derived from both *electr*ostatic and magn*et* respectively. The electrical signal generated from the microphone is amplified and then subjected to a weighting circuit. The weighting circuit allows for one of two (four if available) adjusted frequency-SPL sensitivities. The specifications for the weightings are set out in IEC 651. The “A” weightings are designed to reflect the response of the human ear and in particular, the response at low SPLs based on the equal loudness contour studies (Figure 2.1). The “B” weighting network is for the medium SPL range, while the “C” network is for an equal loudness contour at high SPLs (Brüel & Kjær, 1984:12). The “D” weighting is for use in the aviation field. The SPL meter displays dBA for “A” weighting and so on. Some SPL meters are able to measure the sound with an un-weighted option. This is termed a Linear or “Lin” network.

The “A” network is most popular and is specified by SANS ISO 3741:1999 (ISO, 1999). Networks “B” and “C” were designed based on pure tones, which can be generated in a lab, but are not found as commonly in the everyday environment. Figure 3.2 shows the different weightings at their respective frequencies.

Figure 3.2: A, B, C and D Weightings[16].



The A, B and C weightings mainly differ in the degree of sensitivity at lower frequencies, relative to 1kHz. The A-scale exhibits the least sensitive response. Two sounds judged to be of similar loudness would produce similar dB(A) values, although their un-weighted dB values would vary considerably.

[16] All the weighting curves have a normalised gain of 0dB at 1kHz.

The environment has different effects on different frequencies, and thus analysis of specific frequency bands is often required. There are various methods in filtering the signal from analogue op-amp type active filters to DSP types. The bandwidth of the passband of the filters are commonly one or 1/3 octave wide. The selected frequency range is then amplified and the RMS amount is shown on the display.

3.3.1.1 Detector Response

Sound pressure level meters that conform to the ISO 651 standard have two measuring speed options. The “F” or fast scale has a time constant of 125ms and the “S” or slow range’s time constant is one second (Brüel & Kjær, 1984:14). The output displayed is the RMS value for the respective period of 125ms or 1s respectively.

3.3.1.1.1 The Impulse Sound Level Meter

To obtain the SPL of a sound impulse, a faster sampling rate for the SPL meter is required. A sampling rate of 35ms is recommended. A *peak hold* function may be present on the meter that can save the maximum SPL recorded. This was used extensively in the experimentation section of this study.

3.3.1.2 The Microphone/SPL Meter in the Sound Field

The frequency response of a measuring microphone should exhibit a near flat plot across the active bandwidth. The positioning of the microphone during the measurement is important and is different for different standards organisations. For example, IEC require the microphone to be at a zero angle of incidence while the American National Standards Institute (ANSI) requires an angle of between 70 and 80 degrees. The microphone itself is designed with this in mind and thus a random incident microphone should be used as such, as if pointed directly at the source the reading may be too high. The microphone introduces a disturbance in the sound field by its obstructive presence and should be shaped to be as unobtrusive as possible. When taking measurements in a reverberant area, an omnidirectional microphone should be used.

3.3.2 The Practical Measuring Room

Most SPL measurements are conducted in neither completely anechoic nor reverberant areas but somewhere in between. The position of measurement becomes a critical issue. If measurements are taken within the near field of the source, large SPL changes may be noticed for small movements of the SPL meter. This is particularly true if the measurements are taken at a distance less than the wavelength of the lowest frequency emitted from the source, or at less than twice the greatest dimension of the noise source/machine, whichever distance is the greater. According to Brüel and Kjær’s (1984:21) manual, measurements in the near field should be avoided. However, taking measurements too far away also possess an error. When in the measuring room, the further one moves from the noise source, the closer one gets to a nearby wall and thus the chance of reflected waves being as strong as the direct waves is possible. This area is called the reverberant-field. Best results can be obtained by taking a measurement in the free field, which is between the near and reverberant field. The free field is localised by confirming that a 6dB drop occurs when the distance between recorder and source are doubled. It may occur that no free field exists, as the room may be too small or reverberant, which would require the use of a correction factor as stated in ISO 3746.

3.4 Conclusion

The measurement of RT is not limited to the performing arts. Many other applications require a RT measurement and thus this is an on-going field of adaptation and improvement. The measurement of RT is important in obtaining the absorption of material and room designs. There are various standards that have arisen to facilitate this field. For example, ISO have over eight dedicated documents (some with further revisions) ranging from basic sound absorption in a reverberation room (ISO 3382 part 1 and 2) to the measurement of random-incidence scattering coefficient in a reverberation room by use of sound scattering properties of surfaces, namely, ISO 17497 (ISO, 2008b). The measurement of sound pressure levels are closely related to RT measurements. ISO have over 15 dedicated standards for sound level measurements. Sound absorption measurements are important for the design of buildings both residential and commercial, the transportation industry, noise control as well as for the entertainment industry.

This section focussed on the theory of measurement according to the ISO requirements. These requirements were adhered to for the next experimental chapters.

Chapter

4

4 EXPERIMENTATION: REVERBERATION TIME MEASUREMENT AND ANALYSIS

4.1 Introduction

The practical experimentation of this dissertation is set out over the next three chapters. This chapter consists of the measurement and comparative analysis of the reverberation characteristics of four different environments. It forms the basis for Chapter 5 and 6, which is an experimental study of AR control specifically applied to sound impulses and vocals. There are numerous graphical illustrations in this section, which was required for the analysis of the decay curves and should be viewed in colour. Although this topic is presented apart from the next two chapters, they are tied together in that the next two chapters' data are based on this chapter's measurements.

In order to compare the RTs of the four chosen locations, measurement of the RT and EDTs for each location was undertaken. There was repetition as the method was largely the same. An attempt was made to reduce the repetitive nature of this section by only presenting new issues for each location. This chapter begins with an explanation of the procedure used followed by measurements of RT and EDT for the different environments. The environments differ in type and size and were purposively chosen for their diversity.

4.1.1 General Procedure Overview

Four different environments were selected for this section. The first three represent echoic areas while the fourth was the semi-anechoic or "dead" environmental reference response. The four chosen areas are as follows:

1. Small room
2. Lecture hall
3. Arts theatre
4. Anechoic type environment - Reference

An impulse test was used to derive the RT and EDT characteristic for the four test environments according to ISO 3382. The impulse tests were conducted using the same method for all four areas and were analysed similarly. The sound impulses were recorded with a microphone, which was connected via a high performance sound card to a computer. The recording software used for all impulses was *Cool Edit Pro*[17] ver. 2. The raw wave files were prepared and imported into *WinMLS*[18] acoustic software package for post-processing. Hak and Vertegaal (2007) have found that MP3 coding could be used for impulse tests without noticeable differences in reverberation time as well as other acoustic parameters; however, it was decided that for all recordings and analysis in this study, uncompressed pulse code modulation wave files were to be used. The SPL of the ambient and background noise were also measured to obtain the SNR. The minimum distance between source and SPL meter was 2m for all measurements unless otherwise stated.

The first echoic room test has more detailed graphical explanation while the remaining ones were reduced in content as a lot of the information was of a repetitive nature. The last test for the anechoic environment was covered in detail, as this test setup had the job of being the reference environment and differed from the previous in its almost anechoic response.

The octave filtered RT response is shown for all four locations. While the improvements in DSP and software tools have allowed for a higher accuracy in determining the impulse response waveforms, Oguchi (1998: para 5) feels that it is important to check the decay curve personally when making a critical judgement of the decay slope. Owing to this reason, graphical results have been included as well as tabular results. A table summarising the measuring parameters for all tests follows.

Table 4.1: Testing requirements.

	Reverberation Measurements
Equipment required:	SPL meter, wideband microphone (omnidirectional), oscilloscope and leads, laptop with acoustic recording software, two vertical stands, 2m ripcord, tape measure, spirit level, compass, thermometer, barometer and battery power source/inverter, sound card of least 16 bit depth and 24kHz[19], balloons and extended balloon popping arm Microphone cables used were XLR/F to XLR/M and all cables were balanced where applicable.
Applicable standards followed: (Method reference)	ISO 3382-2:2008; ISO 3382:1997 and SANS 3741 BR0047-13 (Brüel & Kjær, 1984) Measuring Sound. (Rev. ed).
Degree of precision	Engineering method (ISO, 2008)
Sound source and description	Impulse test using gas compliant balloons. Same manufacturer and type of balloon used for all tests. Inflation amount closely checked for all balloons via perimeter measuring with sewing (flexible) measuring tape.

[17] Adobe has bought the rights to Cool Edit Pro, which is now called Adobe Audition (Adobe, 2008).
[18] Acoustic analysis software designed by Morset Sound Development company.
[19] The sound card used was Creative Lab's *Audigy 2NS*, which has a peak sampling rate of 96kHz at a bit depth of 24. The sound card has been tested and evaluated by numerous persons/companies. One such review was conducted by Jean-Pierre Roche from Tom's Hardware website (2003). The results obtained were mixed. There was noticeable crosstalk between channels as well as a 0,5dB ripple occurring above frequencies of 8kHz. Apart from that, the sound card performed well and in particular, it was noted that its distortion levels were very low. In terms of RT measurement, a 0,5dB inaccuracy at frequencies above 8kHz was not found to be problematic; however, most measurements were conducted to a maximum of 8kHz.

	At least two impulse tests performed for each source-microphone configuration. Balloons popped using 1,5m wood dial stick with pin attached to the end. Reflections from researcher minimised. Sound source directivity meets the specification of providing dynamic range without interference by ambient noise.
Sound recording device and description	Half-inch omnidirectional electret condenser microphone (SPL meter). Shure SM58 microphone[20].
Decay curve evaluation method	At least RT_{20} (-5dB to -25dB) and where possible RT_{30} (-5dB to -35dB). Various methods used and some compared: Direct analysis from level-recorder; Least-squares fit; Schroeder method; and Maximum length sequence (MLS)
Octave-filtering	1-octave

4.1.1.1 Background Noise and Overloading

The different measuring locations had different values of background noise. The ISO standard required a 15dB gap between the background noise and the RT's lower limit of evaluation with no overloading to be present at the microphone. This requirement was followed; however, there were sporadic occasions when the background noise increased. No measurements were used during such conditions.

The background noise was evaluated using the relevant standard, that is, the ISO 3741:1999 document stipulates how SPL measurements need to be obtained while there is a presence of background noise [under heading 8.1.4 (p.11)]. In 1984, Brüel and Kjær (1984:29) presented a method that tackles the issue of background noise. If the background noise increased while conducting the RT tests, both the Brüel and Kjær's method and the ISO method were followed in analysing the noise component. In all but one case, the background noise was not significant enough to affect the recordings; nevertheless, the recordings that had an elevated noise component were discarded and a new recording was undertaken.

4.1.1.2 Frequency Range of Measure

The ISO standard specifies a measuring frequency range of 125Hz to 4kHz. The bands above and below are also important and Bradley (2005:171) recommends that the range should be from 63Hz to 8kHz. His motivation for the lower frequency band rests on the premise that the perceived strength of bass sound in halls depends on strong low frequency levels, rather than long low-frequency RT_{30} value influence. Thus, an attempt was made to measure all the octave bands from 63Hz up to and including 16kHz. Bradley does concede that there are difficulties incurred for his recommendation of the larger frequency spectrum use. The main difficulty is in obtaining a unidirectional impulse source that can excite this larger frequency range as well as the increased requirements in the amount of measurements.

Although most of the literature states that musical instruments have a frequency response of less than 5kHz, the researcher has chosen to disagree with that statement

[20] Although an omnidirectional microphone is required for RT measurements, a cardioid microphone (SM58) was also used mainly as a comparative reference as it was the vocal microphone used in the subjective tests in Chapter 6 of this experimentation.

as that generally represents classical instrument types. Synthesisers and software-generated effects have a larger frequency response and are used widely in non-classical music productions. Music genres like electronica, acid jazz, ambient and the like make use of sounds that are well over the 5kHz boundary. Added to this, even percussion instruments have their frequencies extending above 5kHz, excluding the higher harmonics.

This experimentation chapter was able to meet and exceed the ISO standard's frequency range requirement, and where possible, Bradley's requirement.

4.1.1.3 Audience and Occupancy

An empty hall/room obviously has a different acoustic characteristic to the same hall when it is fully occupied. Not only the audience but also the performing musicians may cause differences. The ideal scenario for studying the response of the test area would be to do it while the area is in a state that one would be generalising their results to. Some halls have seats that are covered in material and thus offer some approximation of what it would be like if the seats were occupied. However, correction factors may still need to be applied.

In terms of this experimentation chapter, the results were used as a basis for the application of AR (conducted in the next chapter) and do not require correction factors, as the rooms was still unoccupied during those tests.

4.2 Reverberation Time Analysis for Small Room

The desired goal was to locate a room that would be similar to that of a shower but with enough space for practical testing. A small rectangular room was identified for this test, which on initial subjective analysis exhibited a dominant low frequency response with an uneven frequency range response. It was hoped that the RT across the octave bands would be uneven and exhibit considerable variation. The test specific data is summarised in the following table[21].

Table 4.2: Summary of test setup data for first echoic location - small room.

	Measurement 1: Small Room
Date:	22 June 2009
Time:	19:20-21:45
Site:	Doornfontein, JHB (UJ) Lab 4319 (small enclosed lab within larger lab) John Orr Building
Room layout:	See upcoming photos and top-view line drawing
Volume of room:	Length 4,54m; width 3m; height 3,3m - large windowsill of 4,54x1x1,2m (additional) - Two reinforcing cement cross members along roof: 4x0,2x0,4m (encroaches room volume) Overall volume = 50,07m^3
Condition of room:	Empty of furniture except a single classroom type desk and plastic chair. Room was fully enclosed. Single entrance/exit door closed. Windows also closed.
Environmental conditions: Temperature, relative humidity	Partly cloudy cool night 16°C RH:39%
Degree of precision (coverage)	Engineering method (ISO, 2008)
Measuring height above ground/floor plane:	1,3m
Distance between source and microphone:	>2m (unless otherwise stated)
States of occupancy	Unoccupied (maximum of two persons in the test room)
Spatial averaging	Arithmetic mean RT of all source-microphone positions.
Background noise level	41-42,5dB
Peak SPL of impulse	108,7dB
Theoretical predicted result[22] using equation 2.10	$RT_{60} = \frac{0.16V}{S\bar{\alpha}} = \frac{(0.16)(50.7)}{(81)(0.04)} = 2.5\ seconds$

Reverberation time measurements can be undertaken with all states of occupancy; however, people, clothing and temperature are some of the factors that are related to absorption and thus for the experiments undertaken, occupancy has been stated as well as temperature and RH. Added to this, it is important to note that when doing comparative analysis, measurements taken under different environmental conditions need to be assessed with caution and where possible similarity in conditions should be

[21] In order to meet the ISO standard for RT measurements, the test specific data needs to be presented for each test, as shown in the table.
[22] For interest I wanted to compare Sabin's calculation to the actual measured result. This would be an average value across the octave bands.

ascertained, or at least stated as a limitation. This point has been complied with in this study.

4.2.1 Room Description and Layout

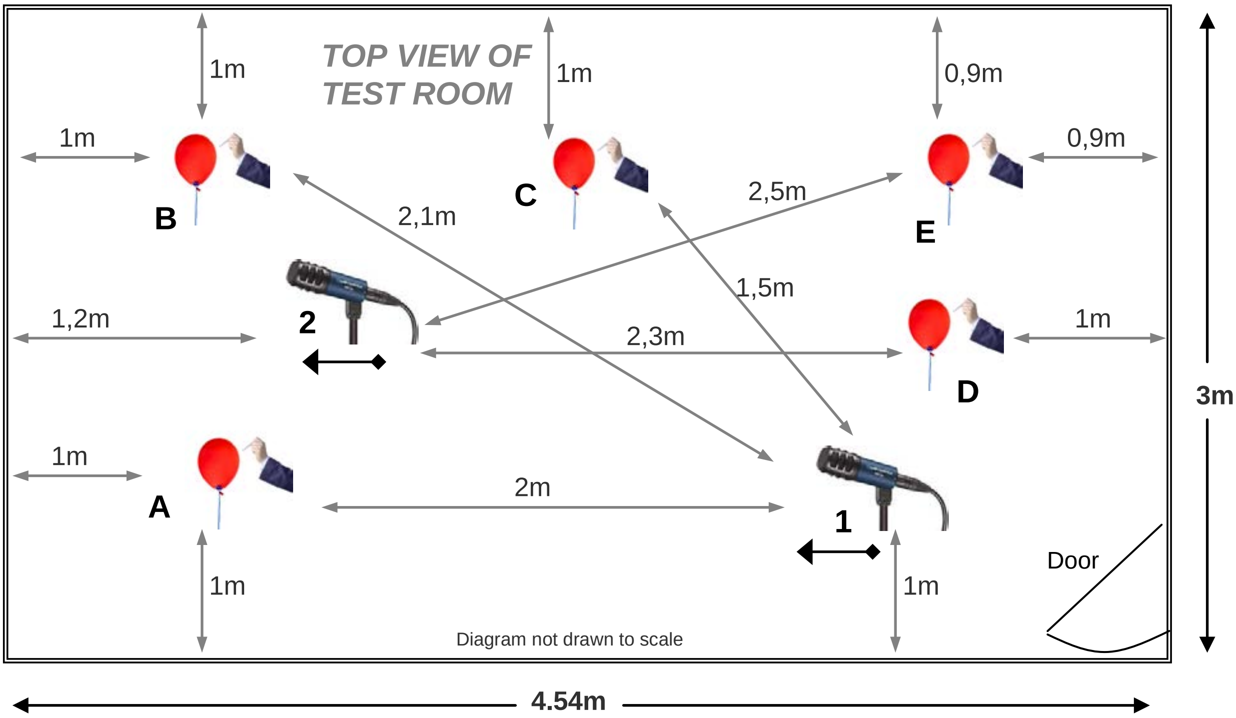
The room was rectangular and small. The vertical surfaces were all of cement cladding while the floor had a thin layer of melamine “tile” over the cement floor. Altogether, the surrounding surfaces were found to be reflective in nature. The roof was also cement and had two cross member cement protrusions. There was a row of windows running along the width of the room. The room was mostly empty. Figure 4.1 and 4.2 provide a view of the room while Figure 4.3 shows the top view diagram of the room’s measurement layout.

In terms of ISO, this room was considered a small room as the volume was less than 300m^3. Small rooms should have their source placements in a corner position, however there needs to be quarter wavelength gap between any reflecting surface and the source (ISO, 2008:4).

Figure 4.1 a-d: Room layout – Small room.

Figure 4.2: Example of test setup.

Figure 4.3: Top view of test room showing room layout and source-microphone positions for small room.



Key to diagram:

A, B, C, D and E: Source positions

1 and 2: Microphone locations

Microphone direction

Figure 4.3 shows the source-microphone positions and illustrates the attempt to follow the ISO standard for small rooms. Although the standard accepts a single source location for small rooms, the researcher undertook several different source-microphone locations.

4.2.2 Measurements

Five measurement configurations were undertaken with each source-microphone position having at least two impulse tests. In total 16 impulse tests were recorded. Two out of the 16 were discarded as they were too loud and the decay waveform exhibited a large amount of distortion[23]. While conducting the tests it was seen that there was little difference in the decay slope for the different source-microphone combinations within the room. This was in keeping with the known theory of the reverberant field as well as the theory of rectangular rooms and their modal activity. Owing to the small volume of the room, it was decided that five measurements were sufficient for this room. The minimum distance for this room was calculated from the following equation:

$$d_{\min} = 2\sqrt{\frac{V}{c\hat{T}}} \qquad [4.1]$$

($\hat{T}$ = 1,5s)
d_{min} = 0,626m

The ISO standard recommends a minimum of 2m between source and microphone but acknowledge that small rooms may require less than 2m. The equation allows for a distance of at least 0,63m as the room volume was very small. The following table summarises the measurement positions and their respective proximities to nearby boundary surfaces.

Table 4.3: Measurement position combinations for small room.

Measurement Combinations (with reference to the top-view diagram)	Distance between source and microphone (m)	Nearest reflecting surface from source (m)	Nearest reflecting surface from microphone (m)	Comments
1-A	2	1	1	Angle of incidence 0°
1-B	2,1	1	1	Angle of incidence not 0°
1-C	1,5	1	1	Angle of incidence not 0°
2-D	2,3	1	1,2	Microphone facing near wall (180°)
2-E	2,5	0,9	1,2	Microphone facing near wall

[23] A brief analysis on the effects of distortion on RT measurement has been addressed and is under the section dealing with the reverberation time analysis for the lecture hall. For all impulse tests, peak level meters were used and overloading on recordings were identified and discarded.

4.2.2.1 Influence of Instrument and Operator

It is generally thought that humans are sound absorbers, but for certain frequency ranges people can act as reflectors. According to Brüel and Kjær's (1984:25) manual, frequencies around 400Hz have been shown to reflect off people to such an extent that a 6dB change was noticed. Thus, the operator of the equipment can have the following impact on the measurement:

- reflection of direct and reflected sound waves,
- absorption of direct and indirect sound waves;
- blocking of reflected and direct waves; and
- an overall influence on the SPL meter by way of the above.

The instrument itself also obstructs the sound field. Ideally, the least amount of interference of the sound field should occur. Some solutions include:

- Meters with canonically shaped fronts.
- Extension rods or cables to allow for remote measurements and smaller meter body interference near the measuring point.
- The user positioned as far away as possible from the measuring device or at least arms length.
- The use of a thin stand (an adapted microphone stand can be used).

To determine the effect of the user's presence on the measurement, one can take the measurements while standing in various positions while noting each reading for each position. A final measurement then is taken where the user is totally out of the sound field. This experiment was conducted and can be found in Appendix A[24]. Necessary steps were taken to reduce the effect of the researcher during the taking of measurements. As this experiment was conducted in a small room, the location of the researcher during each measurement was kept the same for each successive measurement. The balloon was popped using an extension arm and the microphone had a tapered body with a small aperture.

4.2.2.1.1 <u>The Recording Microphone</u>

In this first experiment, two recording microphones were used. The second microphone was the *Shure* SM58. Although the SM58 is not an acoustic surveying microphone, it was quite close in response to the SPL IEC651 type 2 microphone used. A test was conducted to compare the two microphone responses and is shown in Appendix C. The main reason for introducing the SM58 into this part of the experiment is that the SM58 was used extensively as the reference microphone in Chapter 6 (subjective testing).

4.2.3 Results

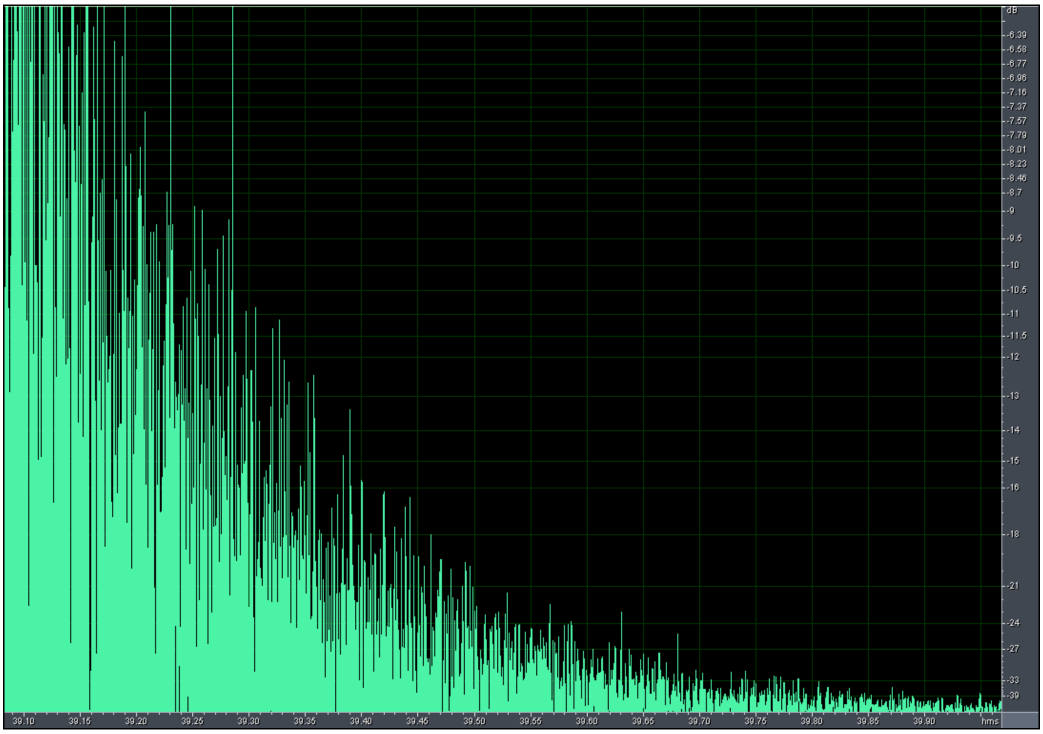
An example of the decay slopes as captured by Cool Edit Pro are shown in the next figure. The right hand vertical scale is in dBs with time on the horizontal axis in seconds. A zoomed view of a decay slope is shown in Figure 4.5.

---

[24] The results for this experiment were surprising. There was a significant variation in SPL while the measuring positions were held constant and only the researcher moved in the room.

Figure 4.4: Two decay slopes for small room.

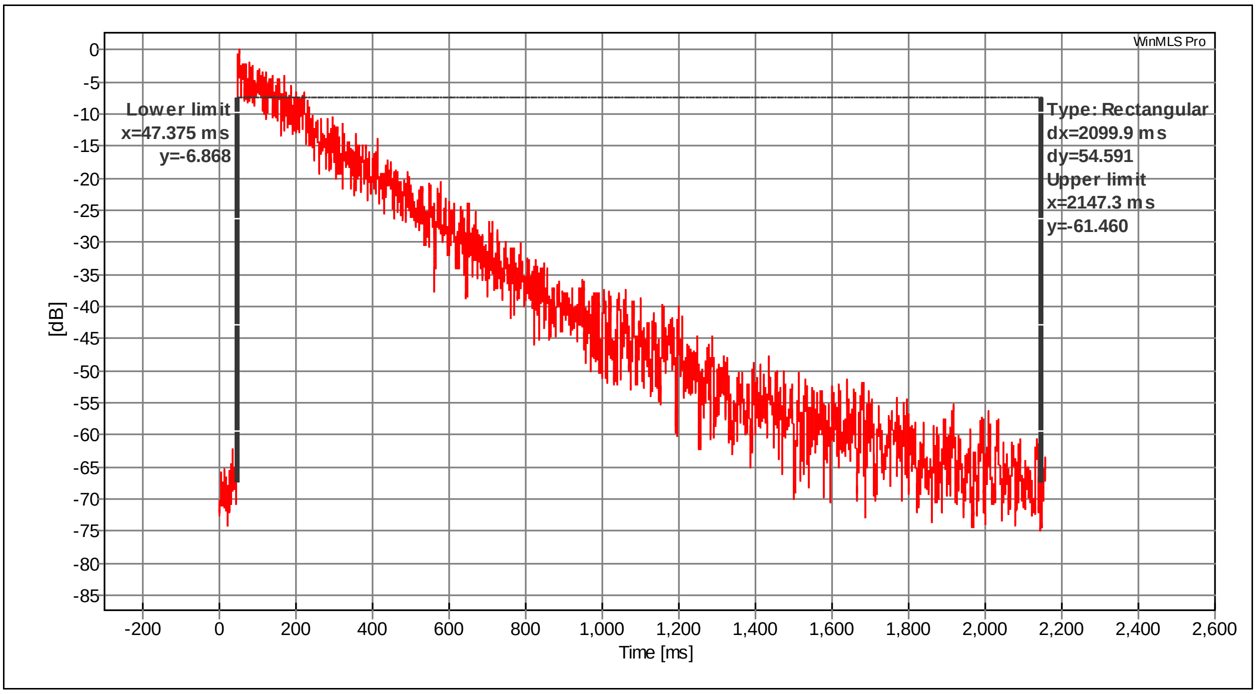
Figure 4.5: Zoomed view of decay slope – small room.



The sound impulse was recorded at a minimum sample rate of 48kHz with a bit depth of 16. The recorded sound impulse sample was then imported into WinMLS software for post-processing. It was possible to use the WinMLS for recording the impulse, however as Cool Edit Pro provided good samples of the recorded impulse and allows for comprehensive adjustments in terms of the positioning of the sound impulses on the computer screen (for illustrative purposes), it was decided to continue using it as the recording software. The impulses were filed and labelled according to the floor plan of the test room. These wave files were then prepared for WinMLS for analysis.

An example of the imported impulse plot is shown in the next figure. The plot has been adjusted to display the "energy in bins", as well as having the DC component removed. The truncation is shown in grey (dotted line) with the noise floor set to 5dB.

Figure 4.6: Normalised time-impulse response for source-microphone position A-1.



4.2.3.1 Schroeder Curves

The new standard, ISO 3382-2:2008 provides less information regarding the backward integration method and refers the reader to the ISO 3382-1 standard for more details. At the time of writing, the latter standard was not yet available and thus the following method was taken from the ISO 3382:1997 standard. Backward integration (Schroeder method) was applied to the squared impulse response that was obtained for each octave band, as well as for the total reverberant sound.

In a treated room where the background noise is very low, ISO 3382:1997 recommend that the integration be started at the end of the impulse response, then proceed to the beginning of the square of the impulse response. This is given below in its common format using two integrations:

$$\int_t^{\infty} P^2(\tau)d\tau = \int_0^{\infty} P^2(\tau)d\tau - \int_0^{t} P^2(\tau)d\tau \qquad [4.2]$$

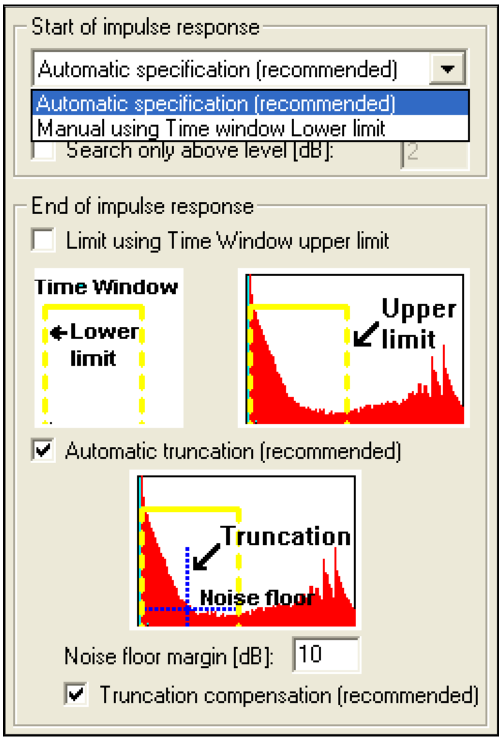
When applying Schroeder's method, the background noise is a factor that introduces inaccuracy (Morgan, 1997; Dragonetti, Ianniello, & Romano, 2009). The Schroeder method is sensitive to the choice of truncation time and it has been recommended by

Morgan (1997) that the truncation point be set to 5dB above the noise floor, while ISO 3382:1997 recommends at least 10dB above the noise floor.

The WinMLS software used in this study can determine the truncation point automatically or by manual settings. The point is chosen as the cross-point between the decaying response and the stationary noise floor. This is detected by viewing the full response or by using an iterative algorithm. A safety margin is introduced to improve the measurement quality, which is called the noise floor margin. However, high noise floor margins require high SNR ratios. The works of Lundeby, Vigran, Bietz and Vorländer (1995) provide a wider context to this issue as well as Bradley's (1996) step-by-step approach to optimizing the decay ranges for room measures. While computer programs can provide advanced analysis of acoustic properties, these programs need verification. Of high importance is the skill level of the user, which also affects the quality of results obtained. Choi and Cabrera (2005) discuss the efficacy of computer aided acoustic tools in their article titled "*Some Current Issues in Computer Modelling for Room Acoustic Design*". Through a round robin approach, they evaluate the efficacy of computer techniques for a range of acoustical parameters including RT and EDT amongst others. This topic has been visited in the last chapter under the heading "Further Study".

WinMLS performs the late truncation by a line fit and a compensation for the energy lost by truncation is estimated assuming exponential decay to infinity (Morset, 2004:416). An example of the truncation setting is shown in the next figure as a "print screen" view of the software's graphical user interface.

Figure 4.7: Print screen view of the truncation settings in WinMLS.



Using the same source-microphone location A-1, the following Schroeder curves were obtained for the first impulse test (please refer to floor plan in Figure 4.3, and associated Table 4.3 for location of all measurements). The first figure (Fig. 4.8) shows the curve with a noise floor set to 5dB while the second (Fig. 4.9) was set to 10dB as per ISO standard. Both plots are similar excepting that the x-axis (dB range) is 5dB shorter than the other is. Owing to the high SNR obtained during the measurement, a 10dB noise floor is easily obtained for this impulse test, while still allowing for adequate decay range for analysis of the curve.

Figure 4.8: Normalised Schroeder curve with noise floor set to 5dB.

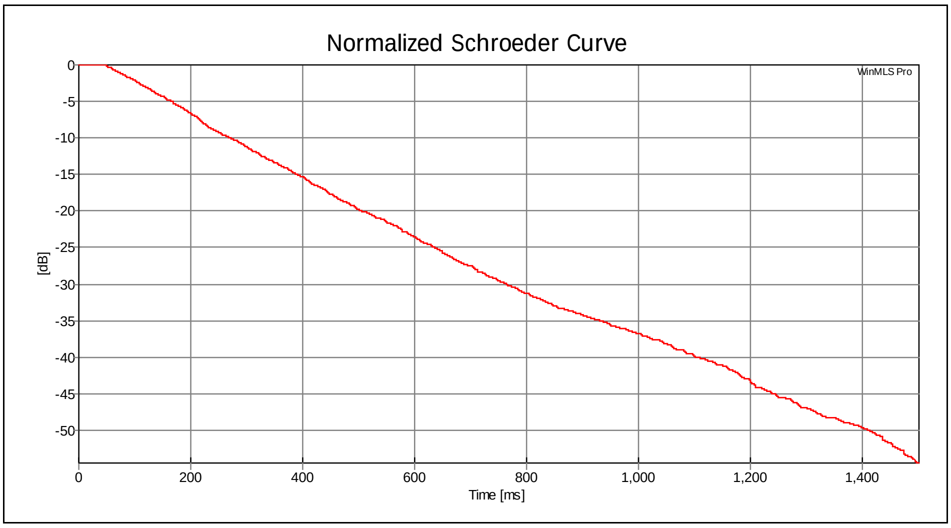
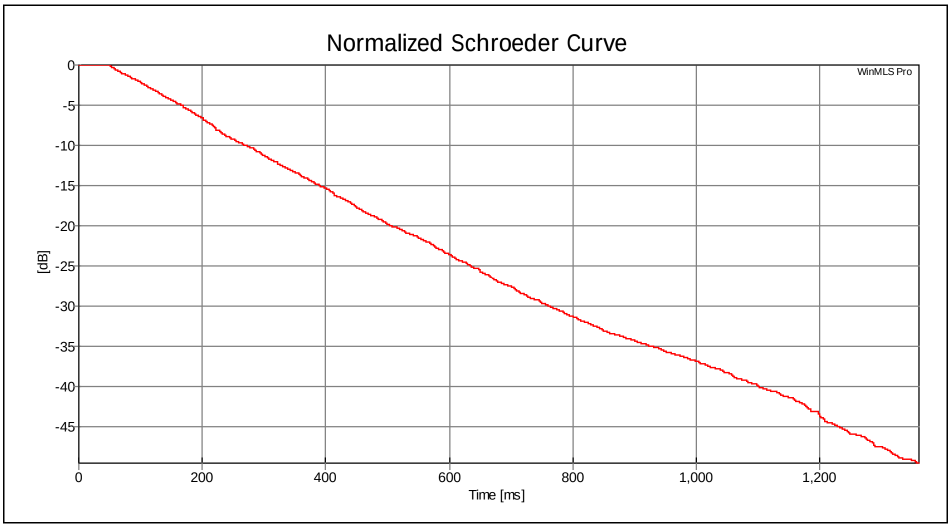
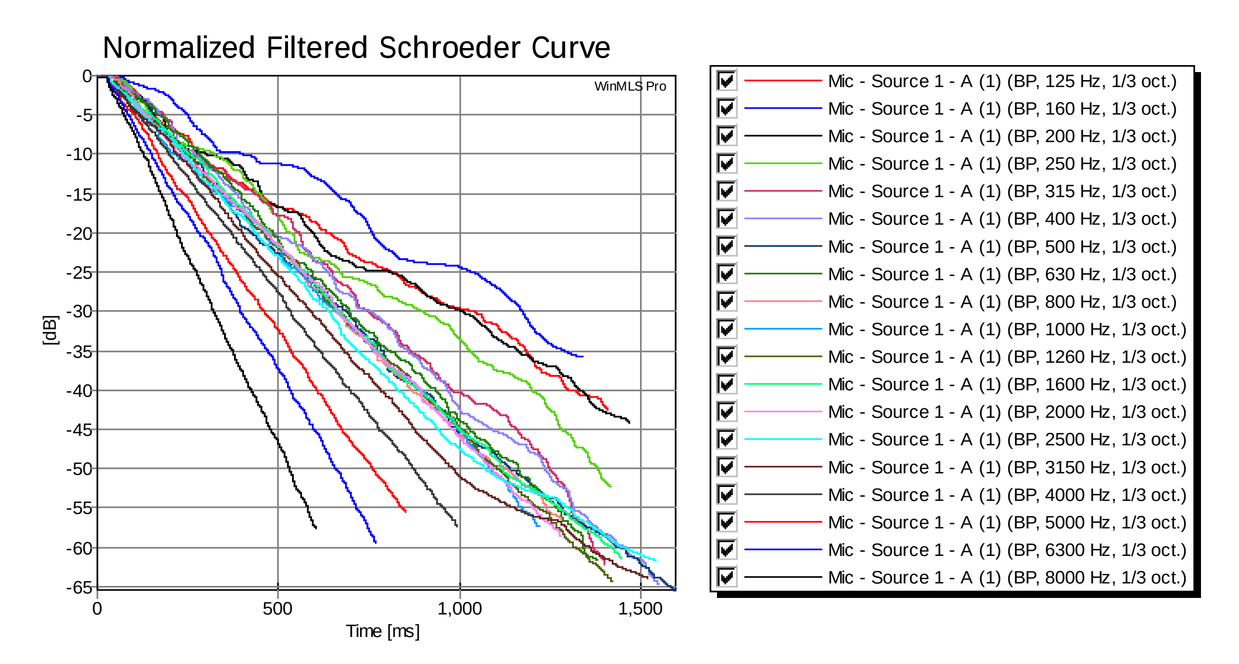


Figure 4.9: Normalised Schroeder curve with noise floor set to 10dB.



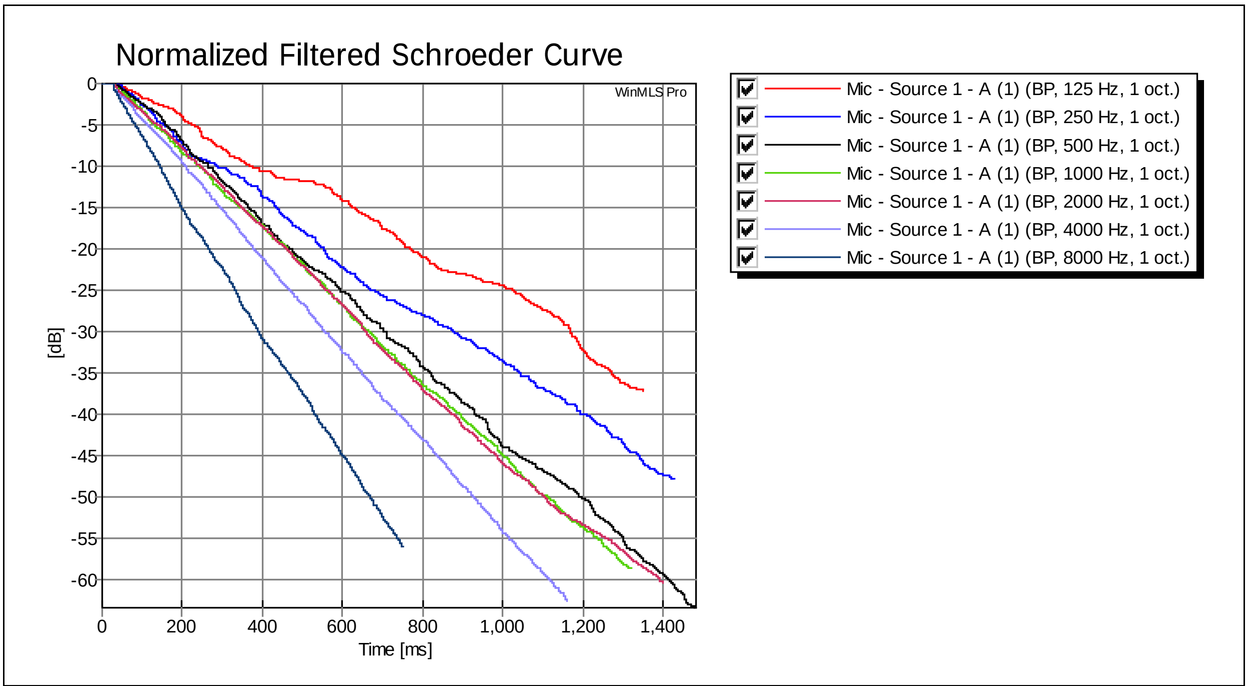
If the Schroeder plot has filtering applied, analysis of the individual octave frequency roll-offs can be performed as shown in the next two plots. The first plot shows 1/3-octave filters applied. The legend shows the different band-pass frequencies and their respective centre-frequencies. Sixth-order Butterworth filters were used for all Schroeder octave-filtered plots in this study.

Figure 4.10: Schroeder plot of a single impulse test with one-third octave filters applied.



For this study, it was decided to standardise on 1-octave instead of 1/3 as there were too many plots for clear analysis, especially when more than one impulse test was conducted. The next plot shows band-pass filtering of 1-octave bands for the same impulse test.

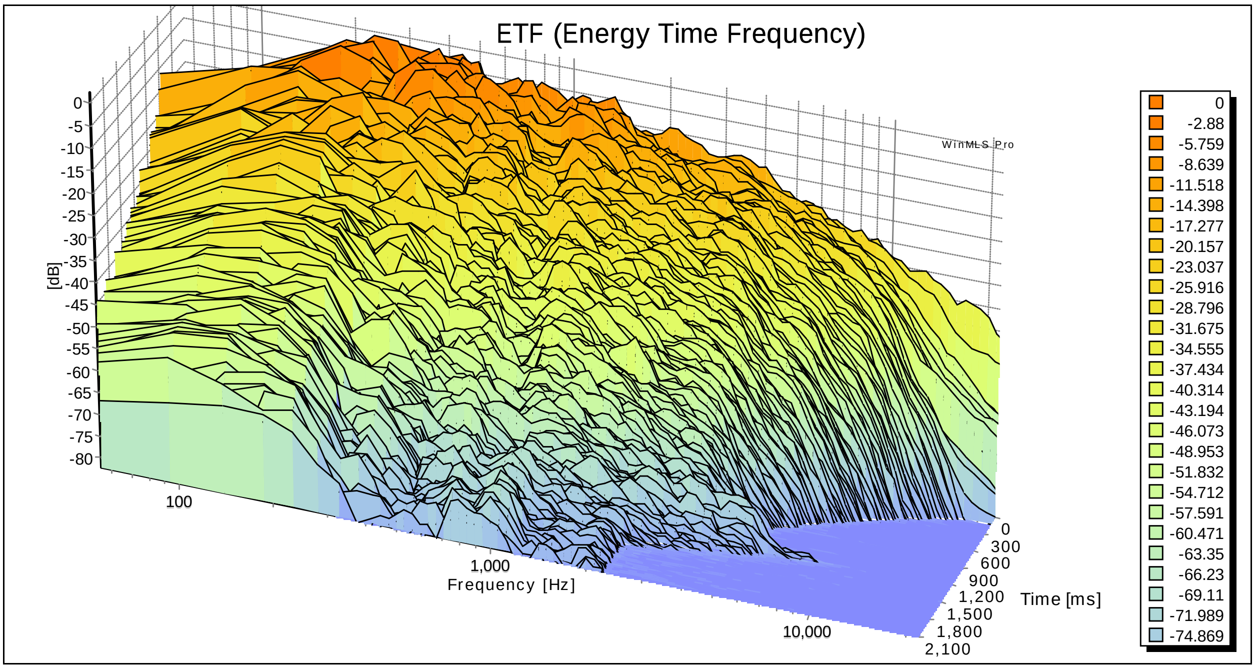
Figure 4.11: Schroeder plot of a single impulse test with 1-octave filters applied.



It can be seen from the previous two figures how the increasing frequencies tend to have the lower RTs.

A helpful plot is the waterfall type, which shows the relationship between sound level, frequency and time in a three dimensional view. From the plot, it can be seen that the majority of the higher frequency reflected waves (above 8kHz) have attenuated considerably after 400ms. This is in keeping with the theory of sound attenuation and high frequency components. In contrast to this, the lower frequencies were still active one second after the impulse, which was as predicted for this small rectangular test room.

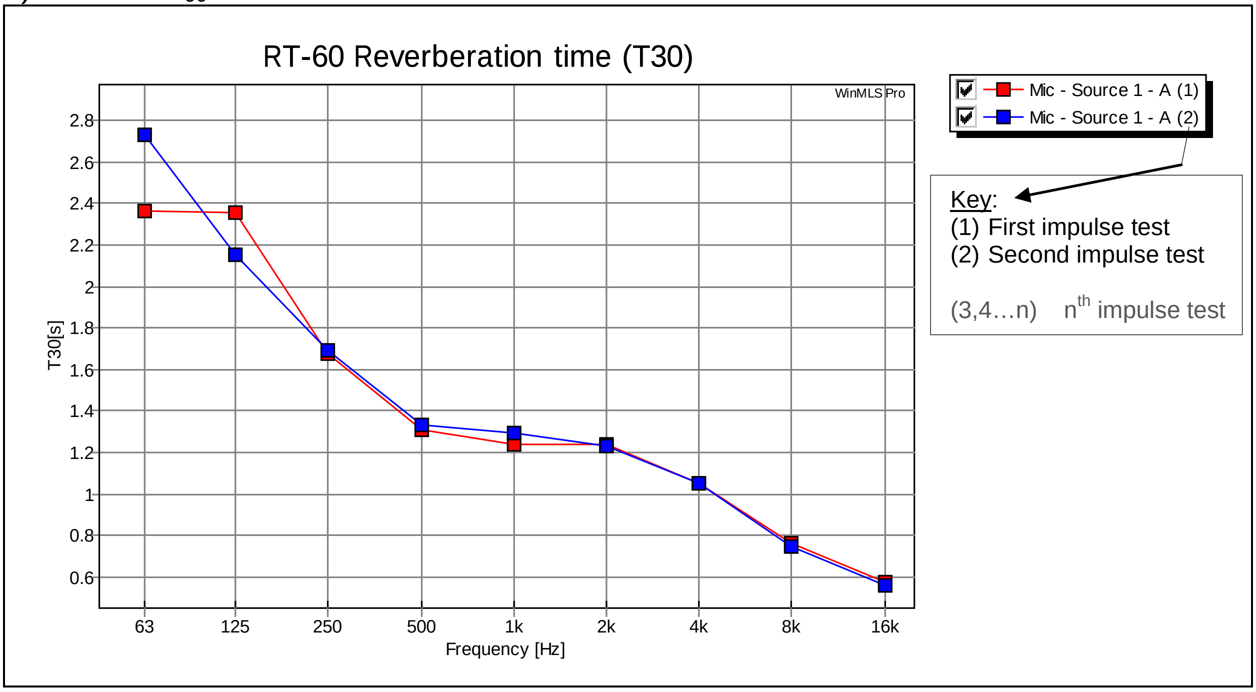
Figure 4.12: Waterfall plot for small room (test position A-1).



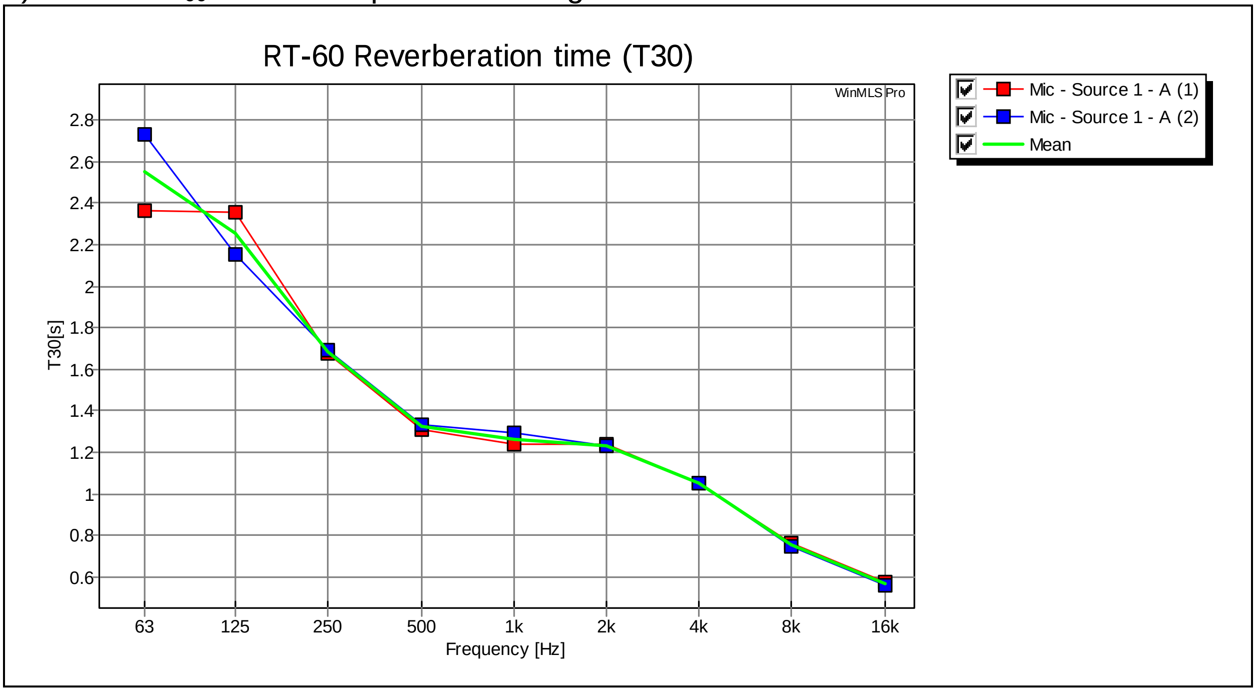
4.2.3.2 Octave Band Reverberation Decays

At least two impulses are required for each source-microphone location test as per the ISO standard. An example of two impulse tests conducted for one source-microphone location and its mean is shown in the next figure. The vertical axis represents the RT, while the horizontal axis shows the octave frequency bands. The first test impulse is shown as the red plot while the second test impulse is plotted in blue. Where possible an RT based on RT_{30} was conducted even though the new ISO standard allows for RT_{20} as a basis for RT. Larger roll-offs were permitted owing to the relatively high SNR and low ambient noise during tests. The next short analysis was conducted as an example of how data can be compared.

Figure 4.13 a & b: Deviation between two consecutive impulse tests for a single source-microphone position.
a) Plot of RT_{60}.

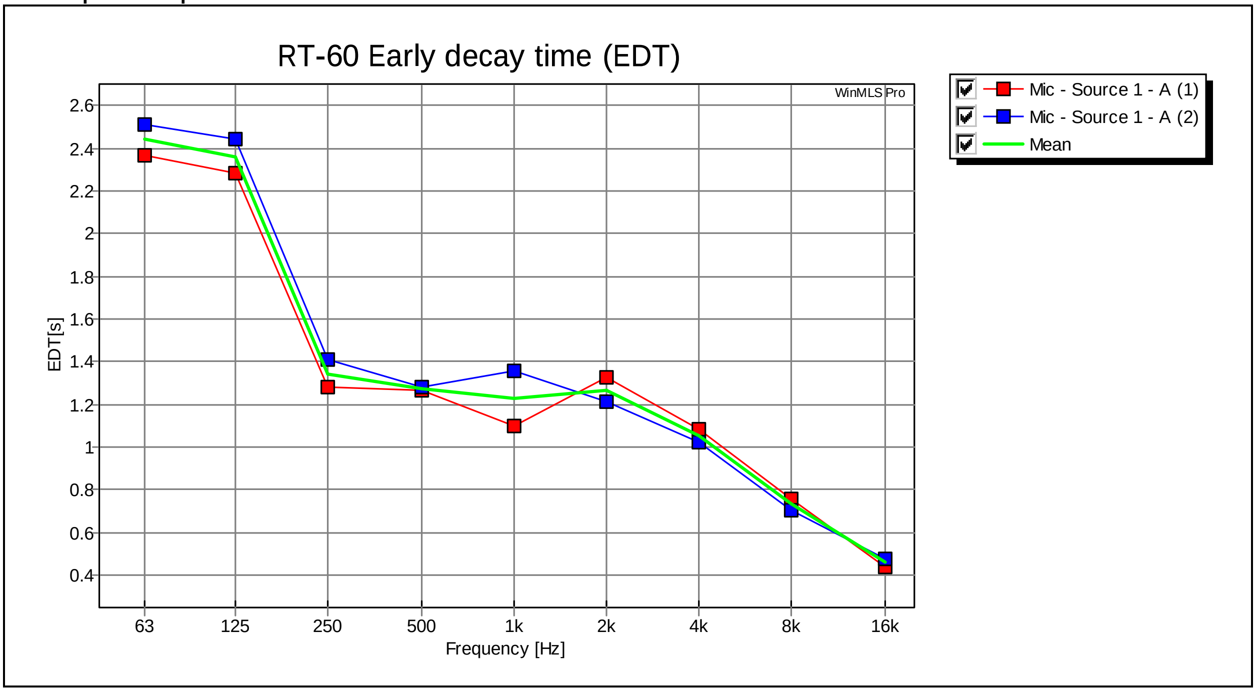


b) Plot of RT_{60} with mean plot shown in green.



The two impulses are closely matched excepting the lowest octave bands, particularly 63Hz. The 63Hz band deviation is discussed shortly as to why there was such a large deviation. The other source-microphone positions exhibited a similar characteristic in that there was not much difference between the first and second impulse tests for each source-microphone position excluding 63Hz. An average was used as the final measurement and is shown as the green plot in the previous figure. The impulse sound was greater than 45dB above the background noise level and thus an RT_{30} test was successfully conducted. The EDT had a slightly different response and is displayed in the next figure.

Figure 4.14: Plot of EDT for two consecutive impulse tests taken in one source-microphone position.



The tabular results are shown in the next table. The frequency bands and the associated indexes are also shown. The effective-decay-range or EDR [dB] parameter provides a quality index of the RT measurements and is generated automatically by WinMLS software. Any EDR values below 25dB are considered questionable and thus the RT parameters should be considered carefully (WinMLS, 2004). This was the case with some of the low frequency RT measurements at 63Hz (italicised in table), while the EDRs of all higher octave bands were well over 25dB and thus the measurements were assumed to be reliable. The 63Hz band had EDRs below 25dB and thus provided unreliable results. They have been included for explanation purposes only. The tables also include the correlation value for the RT measurements. The software uses a least-squares fit of a straight line to the chosen range of the decay curve for the computation of decay time parameters. The decay curves should approximate a straight line for the decay times to provide a quality indication of the RT. The correlation coefficient describes how well the decay curve matches that of a straight line. The coefficient should be close to -1 as that represents a linear slope. From Table 4.4 it can be concluded that the results were approximated well, as their correlation coefficients were better than -0,995 and some being perfectly correlated at negative one.

Table 4.4: Summary of test parameters and quality indexes for source-microphone position A-1.

Mic – source 1 – A 1st Impulse									
F [Hz]	*63*	125	250	500	1000	2000	4000	8000	16000
SNR [dB]	*20.6*	36.3	49.4	57	55.2	55.3	54.5	47.2	30.4
EDR [dB]	*22.6*	36.2	49.6	59.5	58.3	58.4	57.4	51.8	37.3
EDT [s]	*2.37*	2.29	1.28	1.26	1.1	1.33	1.08	0.76	0.44
RT30 [s]	*2.37*	2.35	1.68	1.31	1.24	1.24	1.05	0.76	0.58
correlation	*-0.998*	-0.998	-0.998	-0.998	-1	-0.999	-1	-1	-0.999
RT20 [s]	*2.37*	2.3	1.55	1.22	1.22	1.2	1.04	0.76	0.55
correlation	*-0.998*	-0.996	-0.998	-0.998	-0.999	-0.999	-1	-1	-1

Mic – source 1 – A 2nd Impulse									
F [Hz]	*63*	125	250	500	1000	2000	4000	8000	16000
SNR [dB]	*19.2*	29	40.8	53.6	53.5	55.1	55.8	48.2	30.1
EDR [dB]	*19.2*	28.8	42.6	57.1	56.2	58	58.3	52.8	36.5
EDT [s]	*2.51*	2.44	1.41	1.28	1.35	1.21	1.02	0.71	0.48
RT30 [s]	*2.73*	2.15	1.69	1.33	1.29	1.23	1.05	0.75	0.56
correlation	*-0.998*	-0.994	-0.999	-0.999	-0.999	-1	-1	-1	-0.999
RT20 [s]	*2.73*	2.27	1.6	1.3	1.25	1.21	1.07	0.74	0.54
correlation	*-0.998*	-0.996	-0.999	-0.997	-0.999	-1	-1	-1	-0.999

The table also shows both RT values derived from RT_{20} and RT_{30} for each impulse measurement. Two comparisons can be conducted, firstly the RT_{20} compared to RT_{30} for each independent test (intra-test) and then secondly the comparison between tests (inter-test). The intra-test comparison is shown in the next table.

Table 4.5: Intra-test comparison of RT_{30} and RT_{20}.

Mic – source 1 – A 1st Intra-test comparison									
F [Hz]	*63*	125	250	500	1000	2000	4000	8000	16000
RT30 [s]	*2.37*	2.35	1.68	1.31	1.24	1.24	1.05	0.76	0.58
RT20 [s]	*2.37*	2.3	1.55	1.22	1.22	1.2	1.04	0.76	0.55
Difference [s]	*0.00*	0.05	0.13	0.09	0.02	0.04	0.01	0.00	0.03
Percentage difference %	*0.00*	*2.15*	*8.05*	*7.11*	*1.63*	*3.28*	*0.96*	*0.00*	*5.31*
Mean [s]	*2.37*	2.33	1.62	1.27	1.23	1.22	1.05	0.76	0.57

The table shows that there was some notable difference between the RT derived from RT_{30} and RT_{20} with the 250Hz and 500Hz octave bands leading the way in this deviation. Apart from those two bands, the deviation was insignificant. Overall, the two RT measures correlate well with each other. The main reason was the high SNR obtained during the measuring process, as well as the consistency in the decay slope. Comparing the two impulses in an inter-test is tabulated next.

Table 4.6: Comparison between first and second impulse measurement for a single source-microphone position.

Mic – source 1 – A	**1st and 2nd**	**Inter-test comparison**							
F [Hz]	63	125	250	500	1000	2000	4000	8000	16000
T30 [s] (1st)	2.37	2.35	1.68	1.31	1.24	1.24	1.05	0.76	0.58
T30 [s] (2nd)	2.73	2.15	1.69	1.33	1.29	1.23	1.05	0.75	0.56
Percentage difference	*14.12*	*8.89*	*0.59*	*1.52*	*3.95*	*0.81*	*0.00*	*1.32*	*3.51*
T20 [s] (1st)	2.37	2.3	1.55	1.22	1.22	1.2	1.04	0.76	0.55
T20 [s] (2nd)	2.73	2.27	1.6	1.3	1.25	1.21	1.07	0.74	0.54
Percentage difference	*14.12*	*1.31*	*3.17*	*6.35*	*2.43*	*0.83*	*2.84*	*2.67*	*1.83*
EDT [s] (1st)	2.37	2.29	1.28	1.26	1.1	1.33	1.08	0.76	0.44
EDT [s] (2nd)	2.51	2.44	1.41	1.28	1.35	1.21	1.02	0.71	0.48
Percentage difference	*5.74*	*6.34*	*9.67*	*1.57*	*20.41*	*9.45*	*5.71*	*6.80*	*8.70*

With respect to the RT_{20} and RT_{30}, the percentage difference was better than expected. Excluding the 63Hz band, the octave bands match each other well with just two exceptions, namely the 125Hz band for T30 and 500Hz band for T20. There was a larger percentage difference for the EDT measurements than for the RT. In particular, the 1kHz band had a more than a 20% difference which was substantial. The other EDT bands also had notable differences.

While ISO recommend two impulse tests per source-microphone position, it may be advisable to take a few more for the EDT. In addition, the type of sound impulse has an influence on the repeatability of the EDT decay slope. Some sound impulse devices have been found to provide better repeatability than other devices. There are other reasons why one should undertake at least two tests for each source-microphone position and this is discussed in later tests for a different location. It also confirmed that it is an important feature of the standard[25].

The EDT may be more important in the subjective perception of the room reverberation and has been compared to the full RT in the next table (Beranek, 2004:24). The EDT exhibited a slightly different curve shape as well as having two significant deviation areas, namely 250Hz and 16kHz.

4.2.3.3 Comparing the Different Source-Microphone Positions

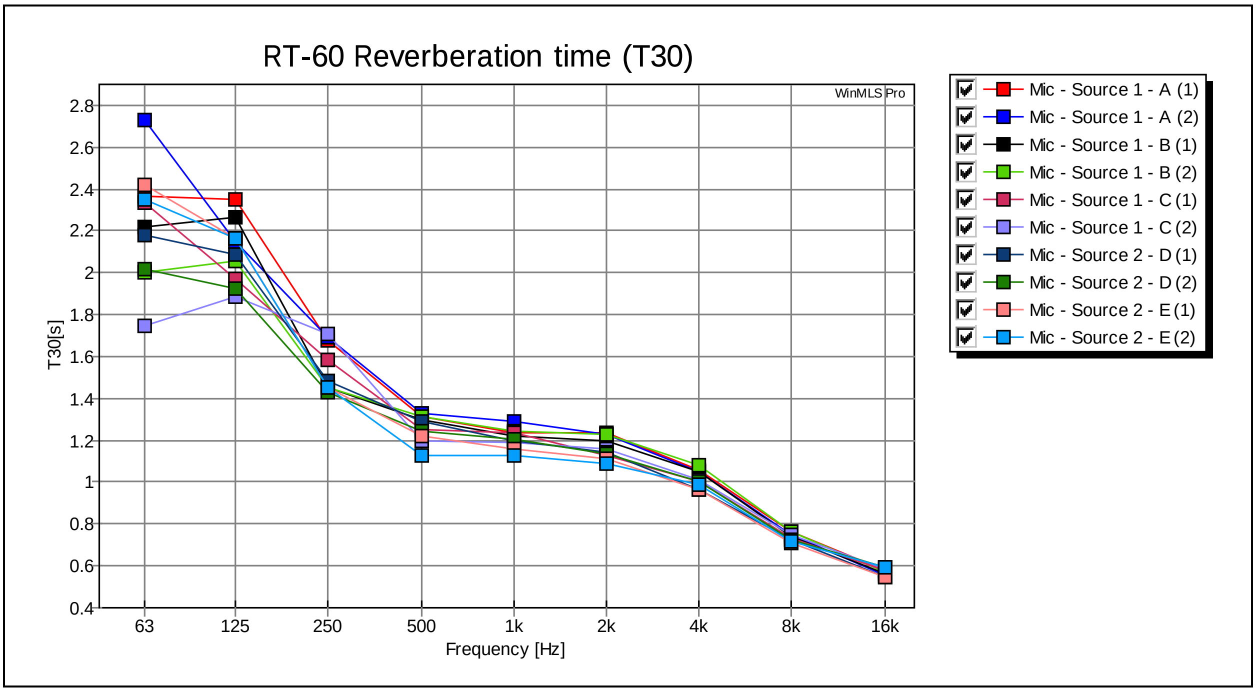
There were five microphone-source combinations and each had two impulse tests[26]. Thus, there were a total of 10 impulse tests performed for the RT analysis for the first venue. The plot is shown in the next figure. The top view floor plan (Fig. 4.3) shown

[25] While conducting RT measurements using the impulse method, an occasion arose where the two impulse measures did not correlate well with each other. Owing to there being a minimum of two impulses for each position, it was quickly noted that a mistake was made in filing the measurement, and the test was conducted again and the correct results refilled. On another occasion, the first impulse did not match that of the second. In this case, the ambient noise had increased, which was not noticed as the researcher was wearing ear protection. It was identified when the raw waveform was re-examined. Thus, two impulse tests assisted in providing a check system.

[26] More impulses were recorded, but for simplicity only two were used for each source-microphone position.

earlier maps out the chosen source-microphone positions. The bracketed numbers after the source-microphone combination in the legend refer to the first and second impulse test respectively.

Figure 4.15: Plot of RT_{60} for all source-microphone positions for the small room.



The RT_{60} shows consistent values for frequencies above 250Hz for all source-microphone positions. The variance increases as the frequency approaches the centre value of 63Hz. The 125Hz band also exhibits a fair amount of RT difference with the lowest value of 1,89 seconds for position *mic-source* 1-C (2). The highest RT measured was 2,35 seconds, which took place in position 1-A (1). Thus, there was a 0,46s difference in RT for this octave band. The EDR values for this band were all well above 25dB and the correlation coefficients were all close to negative one, thus the measurements were assumed to be acceptable. Table 4.7 gives a more accurate numerical result for the equivalent RTs of the plot.

The 63Hz octave band had the highest RT deviation and it also had the lowest EDRs for each plot. If the recommended EDR cut-off point of 25dB was applied, the next figure shows the result. The 63Hz band has no plots, which means that they were not acceptable in terms of the minimum requirements set out for an accurate reading by WinMLS software, and were thus discarded from further analysis. This highlights a previous point made. When using computer assisted methods of measuring and analysing acoustics, one needs to know the ins and outs of the software or else mistakes can creep in. If Figure 4.15 was used as an explanation of the reverberation decay for the small room, one may make incorrect conclusions regarding the 63Hz band as the figure shows considerable deviation for this band. In actuality, the software could not provide a valid result and thus the readings needed to be discarded. Thus, it is important firstly to have software that is accurate in what it says it can and cannot do, and secondly there is a responsibility of the user to be familiar with the software.

Figure 4.16: Plot of RT_{60} for all source-microphone positions in small room with EDR of 25dB applied.

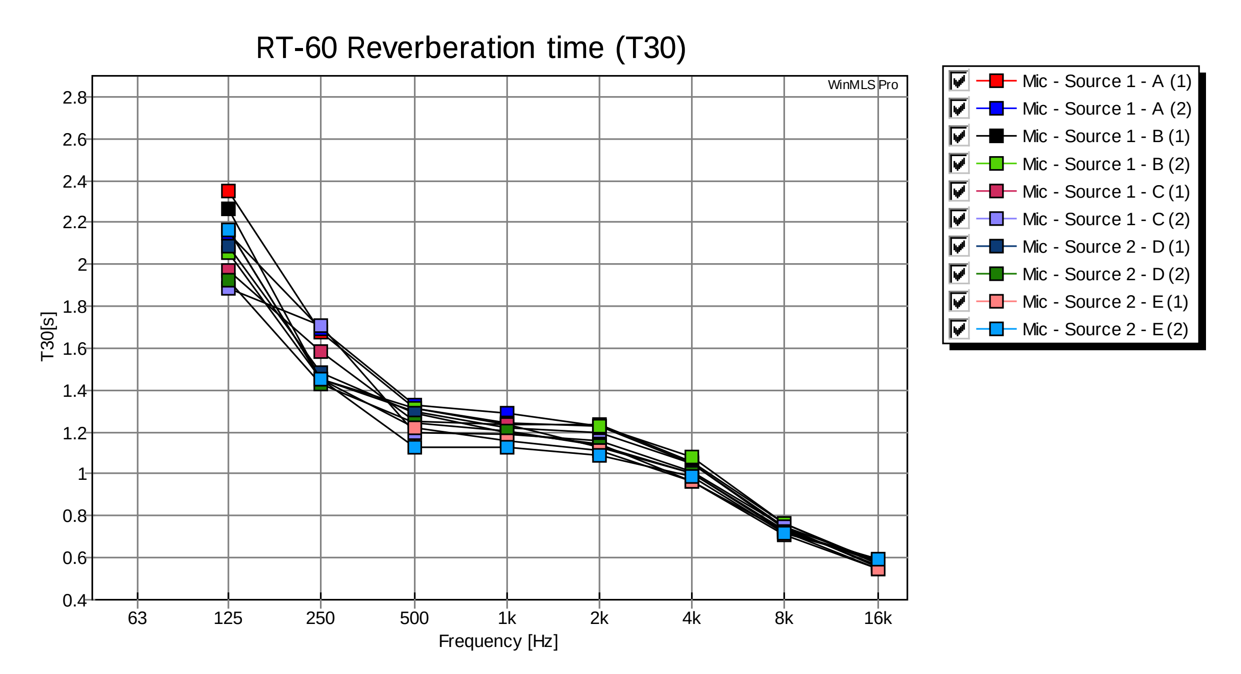


Table 4.7: Numerical RT_{60} for all test locations within the small room.

x-axis F [Hz]	Mic – Source 1–A(1) (s)	Mic – Source 1– A(2) (s)	Mic – Source 1–B(1) (s)	Mic – Source 1–B(2) (s)	Mic – Source 1–C(1) (s)	Mic – Source 1–C(2) (s)	Mic – Source 2–D(1) (s)	Mic – Source 2–D(2) (s)	Mic – Source 2–E(1) (s)	Mic – Source 2–E(2) (s)	Mean (s)
125	2.35	2.15	2.27	2.06	1.97	1.89	2.09	1.92	2.16	2.17	2.10
250	1.68	1.69	1.45	1.46	1.58	1.71	1.48	1.43	1.45	1.45	1.54
500	1.31	1.33	1.30	1.31	1.25	1.20	1.29	1.24	1.22	1.13	1.26
1000	1.24	1.29	1.22	1.25	1.24	1.19	1.20	1.20	1.16	1.13	1.21
2000	1.24	1.23	1.20	1.23	1.13	1.16	1.14	1.14	1.11	1.09	1.17
4000	1.05	1.05	1.05	1.08	1.01	1.01	0.97	1.01	0.96	0.99	1.02
8000	0.76	0.75	0.74	0.77	0.74	0.75	0.72	0.72	0.71	0.72	0.74
16000	0.58	0.56	0.56	0.57	0.58	0.58	0.55	0.59	0.55	0.60	0.57

The EDT of all test locations is shown in Figure 4.17 with its tabulated results following that.

Figure 4.17: Plot of EDT for all source-microphone positions in small room with EDR of 25dB applied.

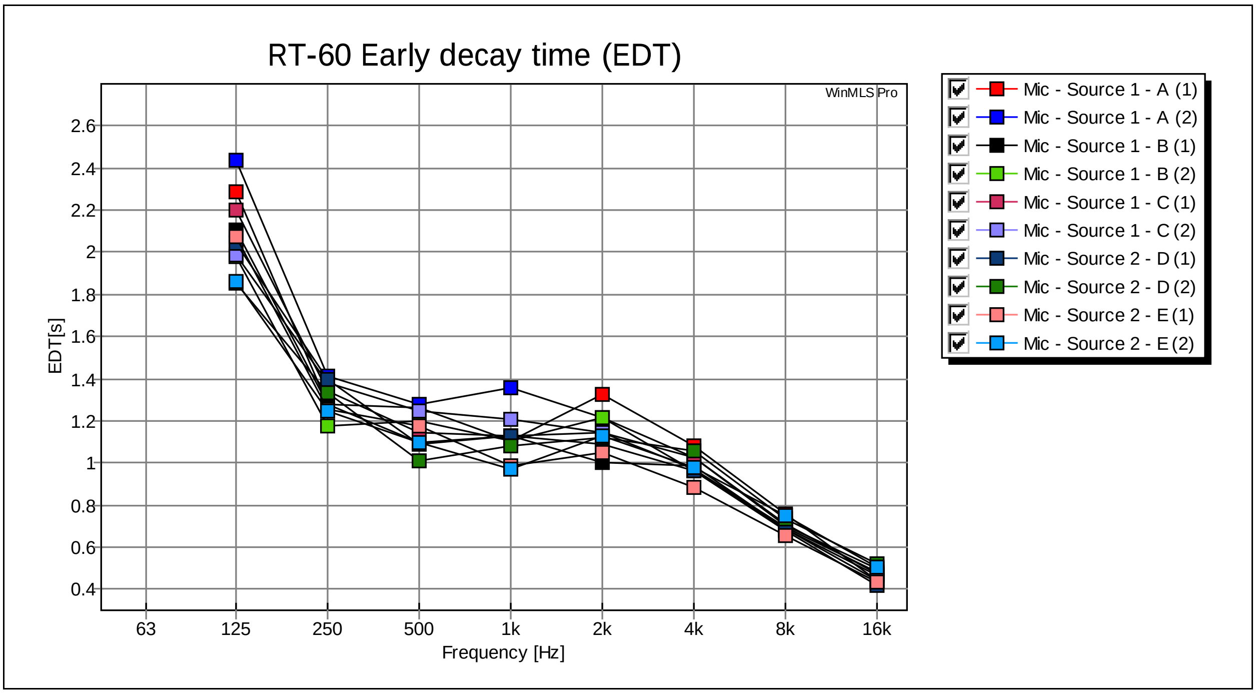


Table 4.8: Numerical EDT for all test locations within the small room.

x-axis F [Hz]	Mic – Source 1–A(1) (s)	Mic – Source 1– A(2) (s)	Mic – Source 1–B(1) (s)	Mic – Source 1–B(2) (s)	Mic – Source 1–C(1) (s)	Mic – Source 1–C(2) (s)	Mic – Source 2–D(1) (s)	Mic – Source 2–D(2) (s)	Mic – Source 2–E(1) (s)	Mic – Source 2–E(2) (s)	Mean (s)
125	2.29	2.44	2.10	1.98	2.20	1.98	2.04	1.85	2.07	1.86	2.08
250	1.28	1.41	1.28	1.18	1.34	1.38	1.40	1.33	1.25	1.25	1.31
500	1.26	1.28	1.10	1.20	1.14	1.25	1.09	1.01	1.17	1.09	1.16
1000	1.10	1.35	1.13	1.11	1.13	1.21	1.12	1.08	0.98	0.97	1.12
2000	1.33	1.21	1.00	1.21	1.14	1.15	1.09	1.12	1.05	1.13	1.14
4000	1.08	1.02	0.98	0.96	1.02	0.97	0.97	1.06	0.88	0.98	0.99
8000	0.76	0.71	0.68	0.69	0.70	0.70	0.68	0.74	0.66	0.75	0.71
16000	0.44	0.48	0.44	0.47	0.50	0.48	0.42	0.52	0.44	0.50	0.47

The octave mean RT for the small room is given in the next figure with the RT of each octave band noted above the respective centre frequency plot. The Mean EDT is plotted in Figure 4.19.

Figure 4.18: Octave band RT for the mean RT_{60} test for small room.

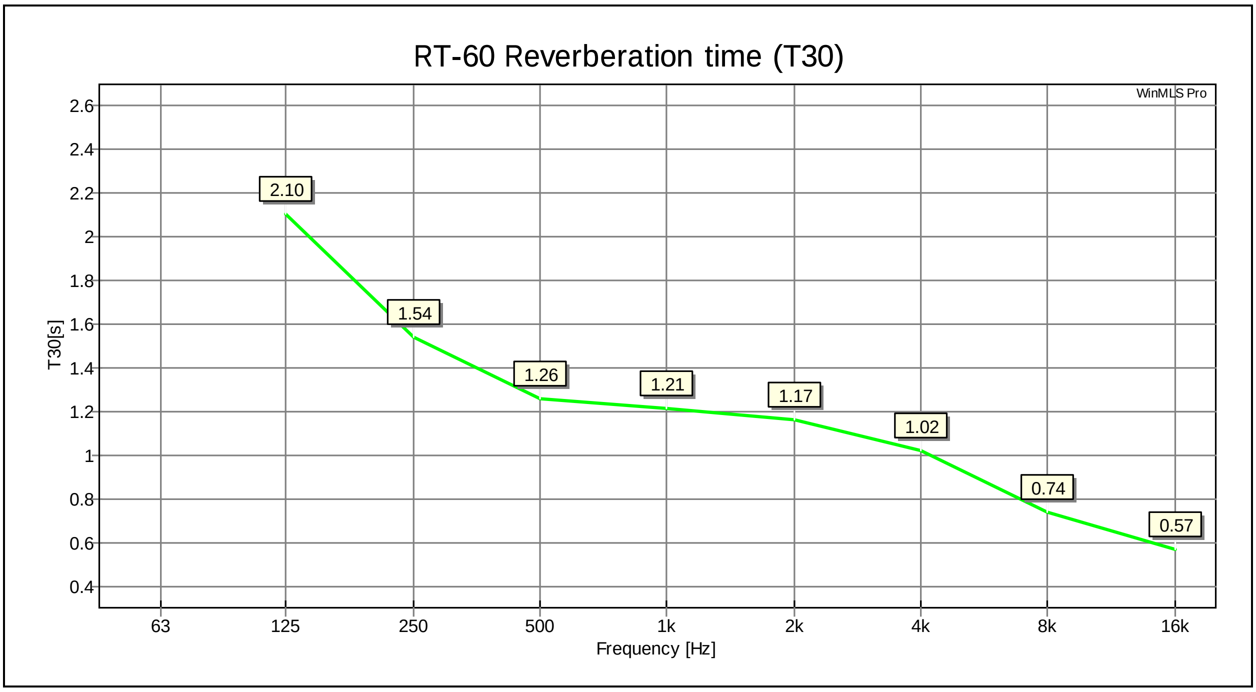
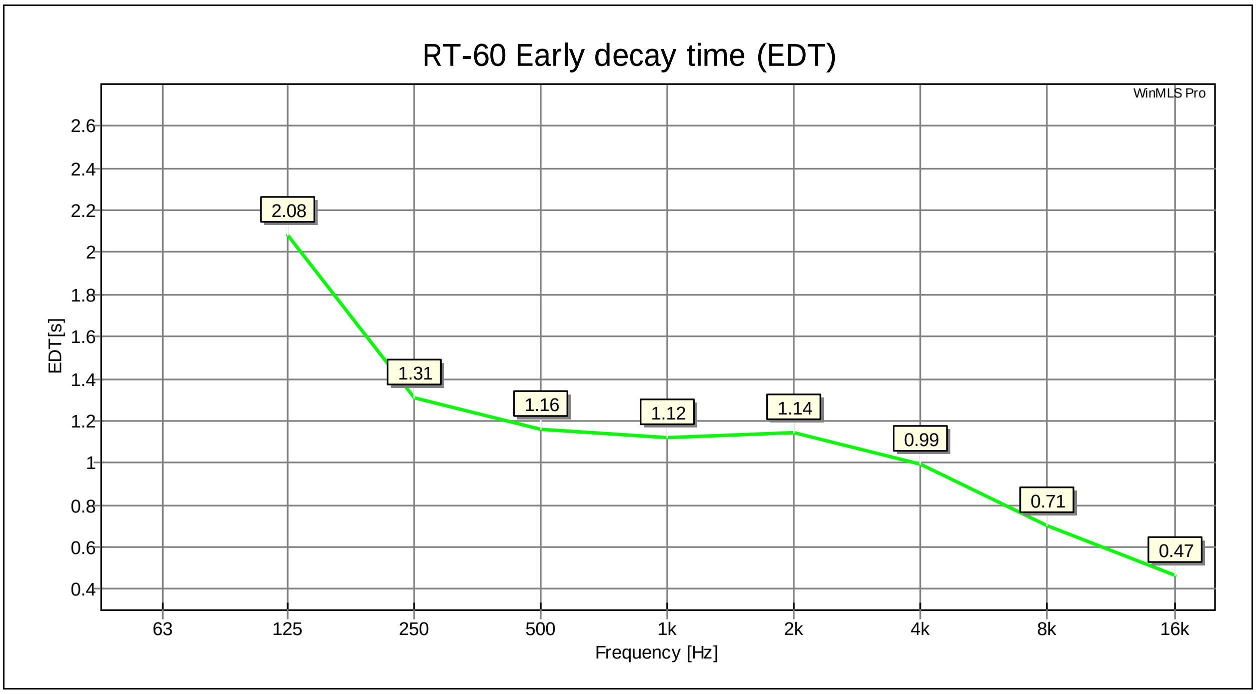
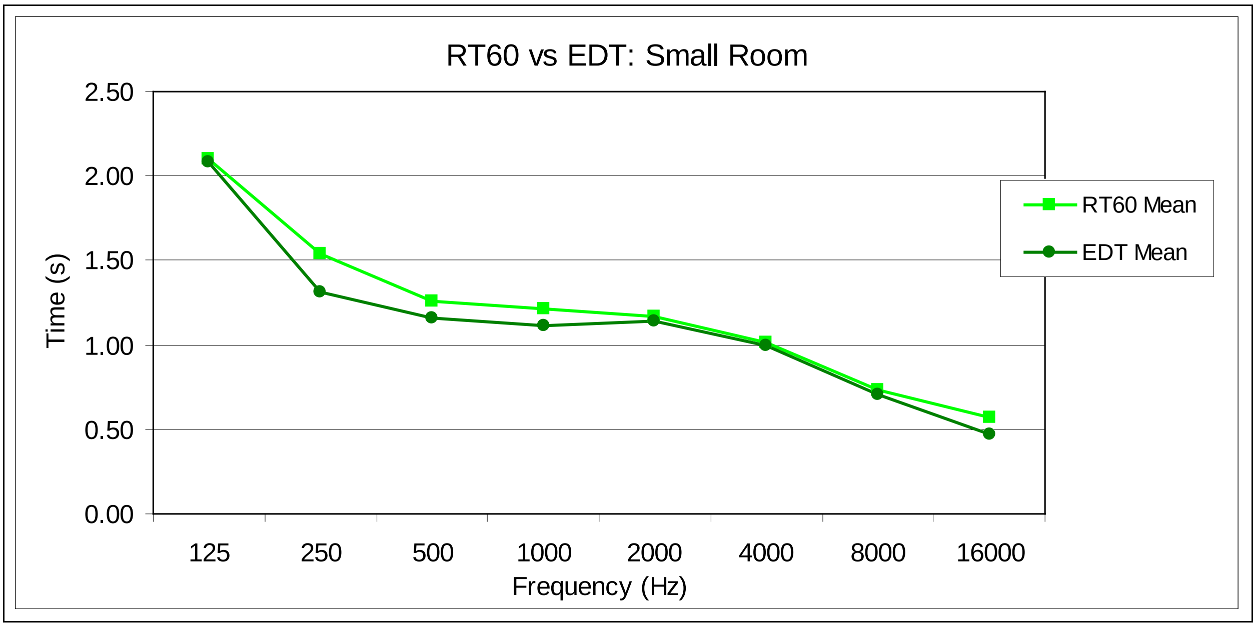


Figure 4.19: Octave band RT for the mean EDT test for small room.



The mean EDT and RT_{60} curves are similar in both shape and magnitude. There is some noteworthy deviation between 200Hz and 1kHz but the trend remains similar. Plotting them on the same axes provides a better view, which is shown in the next figure.

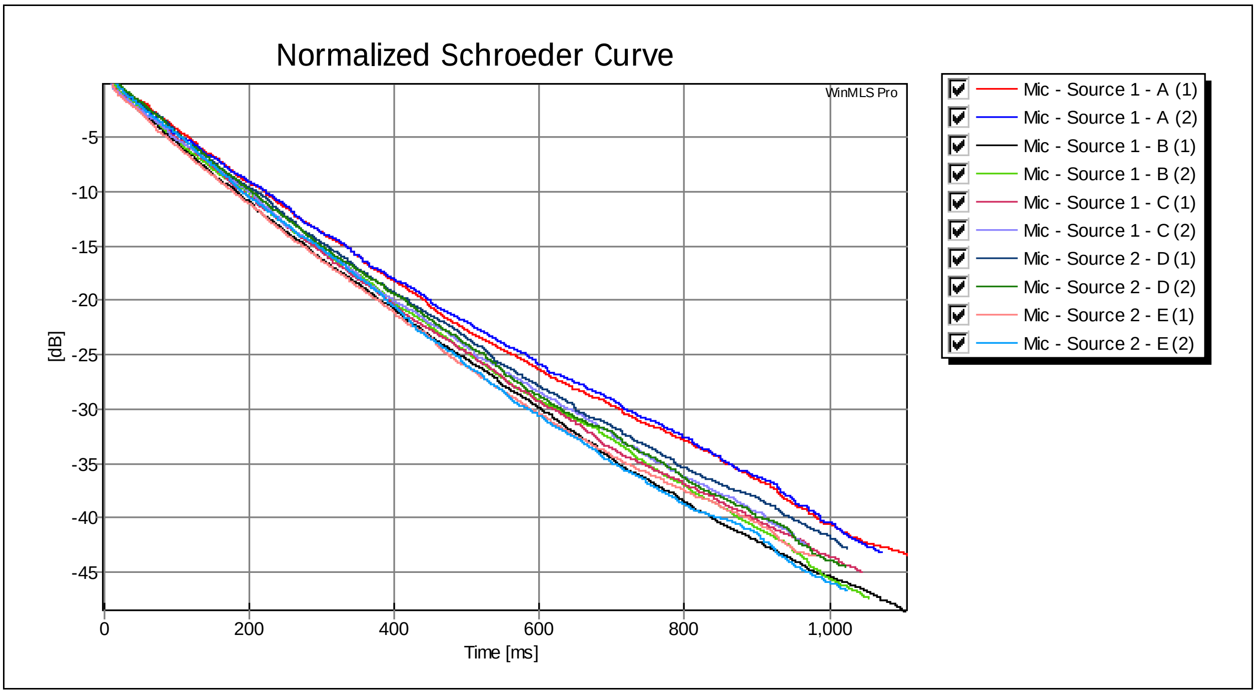
Figure 4.20: Comparison of RT_{60} and EDT for the small room shown on one set of axes.



4.2.3.4 Discussion of Results for Small Room

The lowest frequency components had the longest RT and EDT. Both curves have a plateau area between 400Hz and 2kHz after which they decrease almost linearly with increasing frequency. The room was rectangular and had a small volume. It was initially speculated that the room would exhibit a “muddy” sound, which was confirmed by the substantially higher RT and EDT in the 125Hz octave band. Although five source-microphones positions were used, the variance in RT decay slope was not significant for different positions as shown in the following Schroeder curve. The most notable source-microphone position was that of 1-A (red and blue plots) which shows a deviation from the group from -25dB onwards.

Figure 4.21: Schroeder curves plotted for all effective impulse measurements in small room.



The amount of data that can be compared and analysed for a single RT measurement is extensive. In this example, only a few comparisons were made. For 10 measurements, there are potentially 400 different combinations that could be analysed. For example, in this example only a single position was analysed in the beginning of this section even though there were 5 measuring positions and 10 impulses. While there is a lot of repetition, many results are visually viewed and only significant ones are used for analysis. If for example there was a set of octave bands that exhibited a longer RT than the other measuring positions, one could analyse that position with a statistical mean rather than comparing it to every single measurement. For purposes of completeness and for illustration of methodology, extra information was added to this experiment. The next test locations' tests will only highlight the results and problems that provided significant information.

4.3 Reverberation Time Analysis for Lecture Hall

This venue was chosen as it represents a completely different layout to the small room. The lecture hall has furniture, differing boundary surfaces and a shape that is unlike the small room.

Table 4.9: Summary of test setup data for second echoic location- lecture hall.

	Measurement 2: Lecture Hall
Date:	24 June 2009
Time:	18:10-20:45
Site:	Doornfontein, JHB (UJ) Lecture hall room 106 Outside John Orr Building
Room layout	See photos and top-view line drawing
Volume of room	Length 17m; width 18,5m; height 2,5 – 3,9m Overall volume = ±1229m^3
Condition of room	Open area in front. Tables and chairs staggered along an increasing floor plain.
Environmental conditions: Temperature, relative humidity	Clear cool night 14°C RH:36%
Degree of precision (coverage)	Engineering method (ISO, 2008)
Measuring height above ground/floor plane:	1,3m
Distance between source and microphone:	Variable, but >2m
States of occupancy	Unoccupied (maximum of two persons in the test room)
Spatial averaging	Arithmetic mean of RTs for all source-microphone positions.
Background noise level	32-34dB
Peak SPL of impulse	101dB

4.3.1 Room Description and Layout

The boundary surfaces were cement (three of the walls), wood (rear wall cladding), ceramic tile and suspended ceiling tiles (mineral board with no sound attenuation batts). The room had an almost square perimeter but had a semi-rounded front wall. The floor slopes upward toward the top seating row. The ceiling initially slopes upward and then levels out to create a lower height at the halls rear. There were no windows. There were four doors, which were closed during all measurements. The room has not had any intentional acoustic treatment. There were numerous desks installed along the sloping floor. The chairs were plastic and have no absorptive material on them. The ceiling was of a suspended type with the cement roof having at least 150cm gap above. Some of the ceiling boards were missing. The majority of surfaces within the room were reflective.

Figure 4.22 a-e: Room layout – Lecture hall.

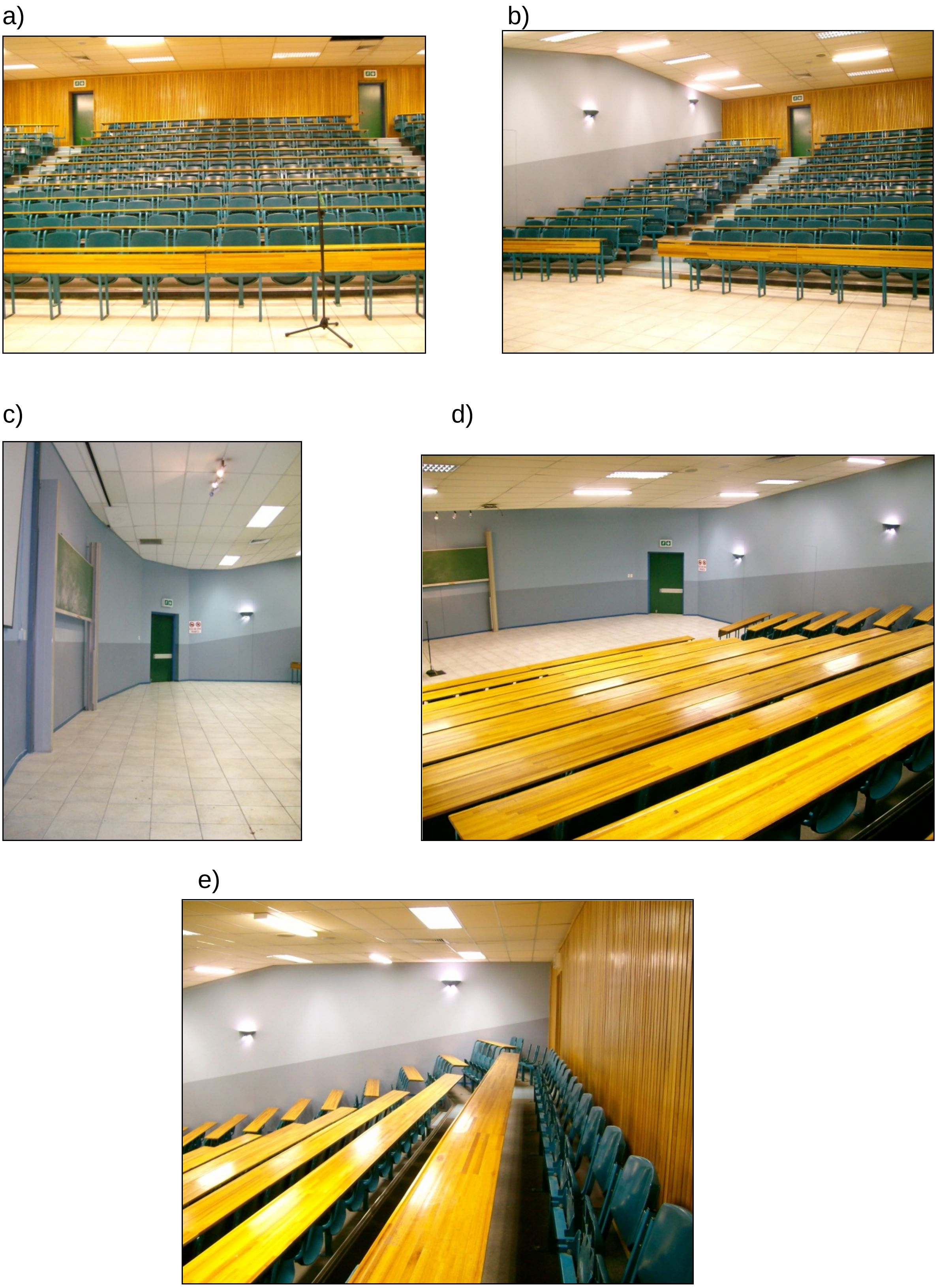
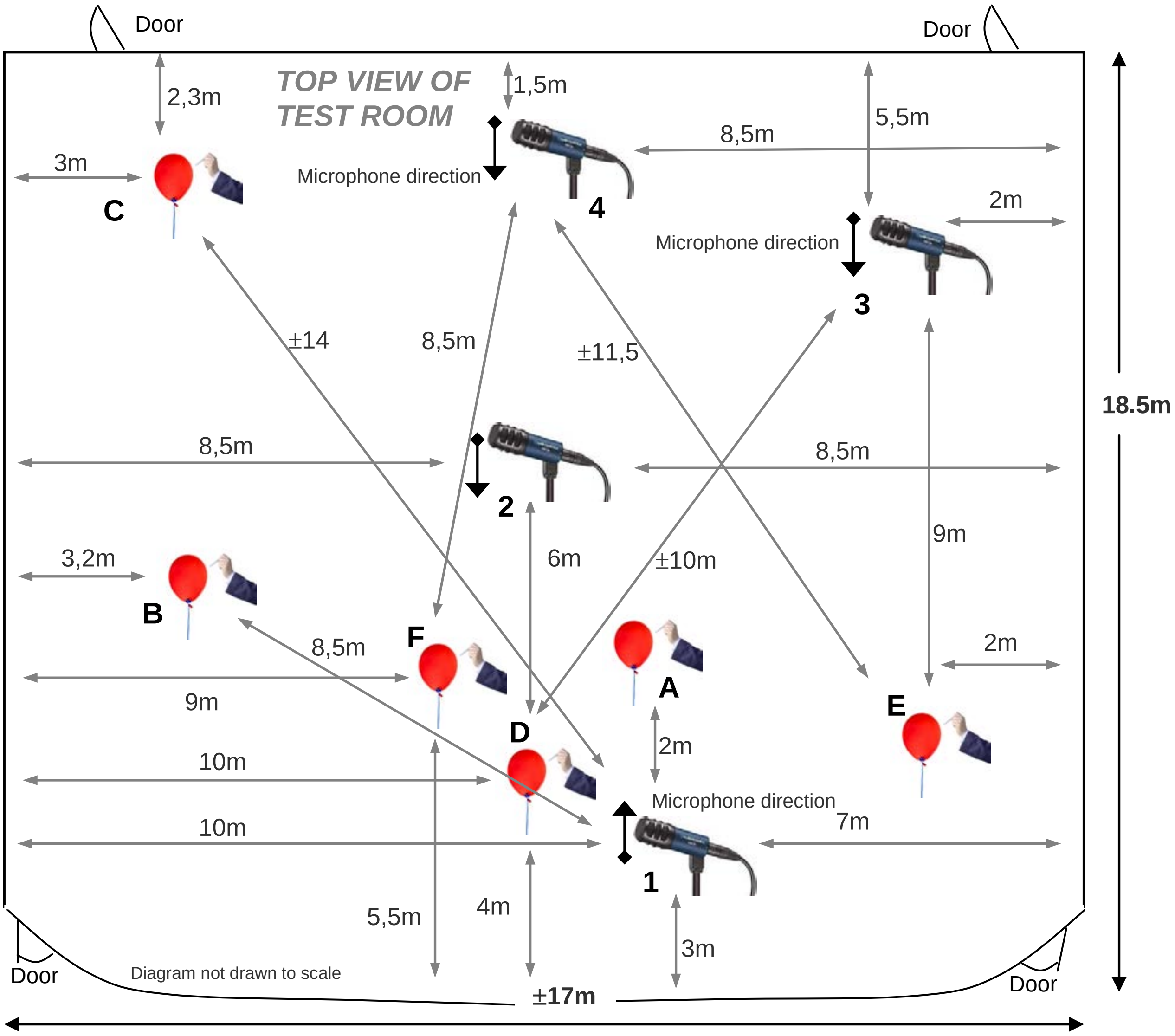
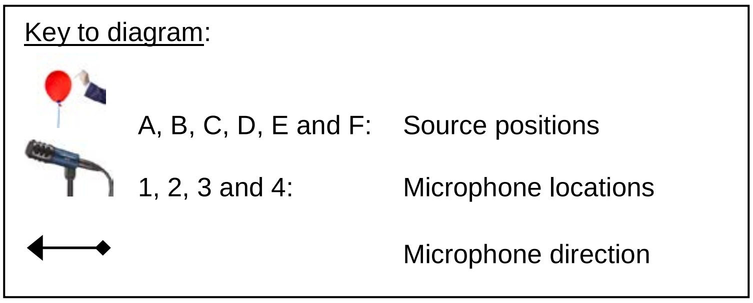


Figure 4.23: Top view of test room showing room layout and source-microphone positions for lecture hall.





4.3.2 Measurements

A preliminary survey of the room was undertaken where the researcher walked around the room clapping loudly to get familiar with the room's response. The researcher also requested his assistant to stand in various locations and talk loudly while the researcher walked about the room. This is a common method and an acceptable method of preliminary room surveying (Eargle & Foreman, 2002:203; Gervais, 2006:141).

Eight measurement configurations were undertaken with each source-microphone position having at least two impulse tests. In total 17 effective[27] impulse tests were recorded. The readings were taken for the different source-microphone positions. The minimum distance for this room was calculated using Equation (4.1):

$$(\hat{T} = 1{,}7\text{s})$$
$$V = \pm 1229\text{m}^3$$
$$d_{min} = 2{,}91\text{m}$$

Table 4.10: Measurement position combinations for lecture hall.

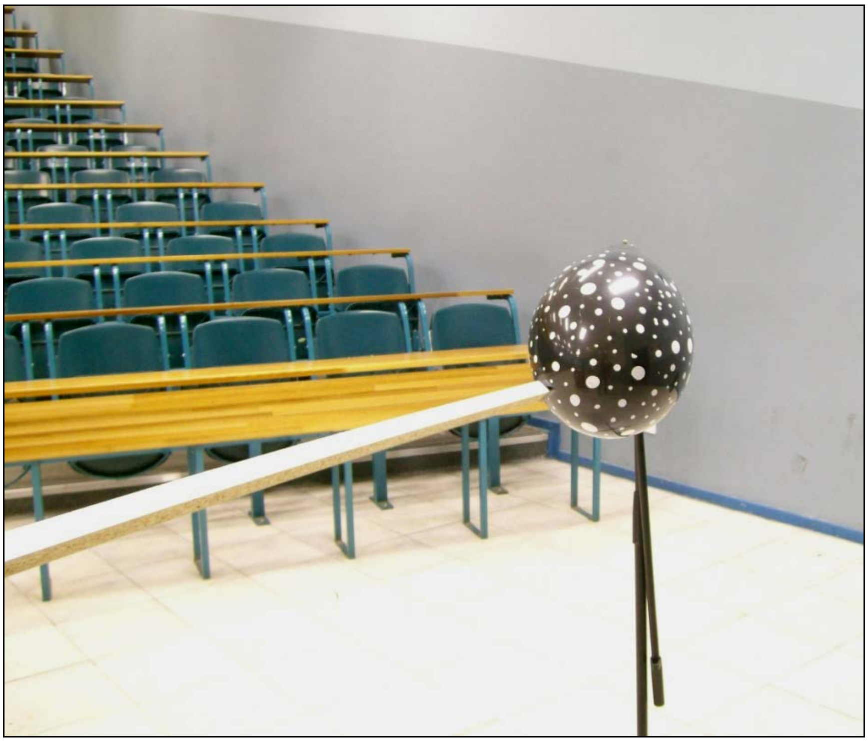
Measurement Combinations (with reference to the top-view diagram)	Distance between source and microphone (m)	Nearest reflecting surface from source (m)	Nearest reflecting surface from microphone (m)	Comments
1 - A	2 (see note[28])	5	3	Angle of incidence 0°
1 – B	8,5	3,2	3	Angle of incidence not 0°
1 – C	±14	2,3	3	Angle of incidence not 0°
2 – D	6	4	8,5	Angle of incidence 0°
3 – D	±10	4	2	Angle of incidence not 0°
3 – E	9	2	2	Angle of incidence 0°
4 – E	±11,5	2	1,5	Angle of incidence not 0°
4 – F	8,5	5,5	1,5	Angle of incidence almost 0°

The balloon was popped using an extension arm as shown in the next photo. This method was followed for all impulse tests in all locations.

[27] While numerous impulse tests were conducted, some were discarded. Some reasons for discarding were because the portable battery power turned off and no trace was captured, or the balloon was popped but the program to record the impulse was not turned on. Another factor was the occurrence of noise. It occurred on a few occasions where someone walked into the test venue, or an aeroplane flew by which voided the measurement. The main reasons were user mistakes and sporadic environmental noise. Thus, the term “effective” was used to describe measurements that were successfully carried out without alteration or manipulation.

[28] Source-microphone distance is less than the minimum distance recommended by the equation. However, this location was the closest approximation to the subjective tests conducted in chapter 6 and was included for that reason.

Figure 4.24: Example of extension arm for balloon pop impulse.



4.3.3 Results

All 17 impulse tests were plotted into an RT_{60} and EDT plot shown next. The numbers in the boxes represent the EDR value and not the RT. There are limitations in importing highly populated plots into an A4 document. The original plot is much larger and has further details, which have been omitted owing to space limitations of paper size and aspect ratio. Although the plot is somewhat squashed, a quick analysis shows the focal points that need more attention. For example, it seems that all the 63Hz octave band impulse responses were unacceptable as they were all below 25dB, which is the lower limit for a reliable reading (minimum EDR setting of WinMLS). One of the EDRs is as low as 6,1dB (the one that is off the chart towards the top left position, EDR not shown). Thus, all RT measurements[29] in the 63Hz octave band have been excluded and the revised plot is shown in Figure 4.26.

[29] It should be noted that there were different source-microphone positions and that the following plots just clumps them all together for a total effective room RT and EDT value. It is possible to obtain RT and EDT for different areas within an enclosed space as well as obtain the total room average. For this location (lecture hall) only the latter applies, while in the next location specific areas were identified and had their RT and EDT measured for each.

Figure 4.25: RT for all seventeen impulse tests conducted for large room with their EDRs shown.

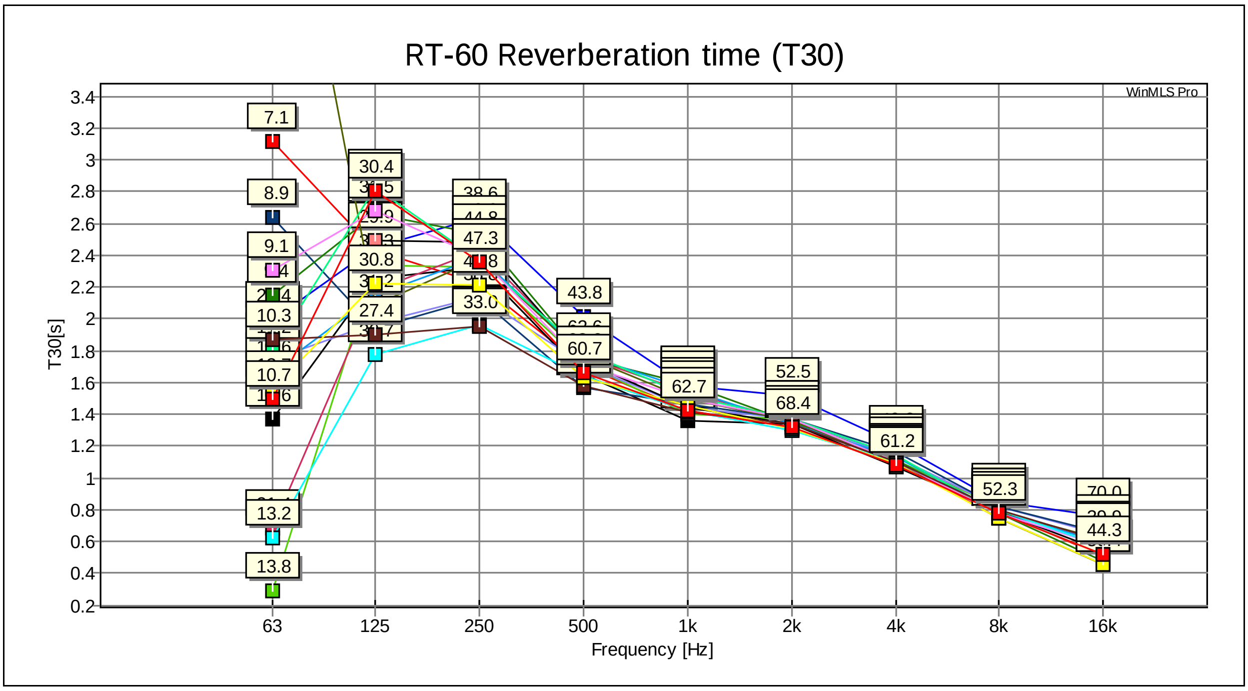
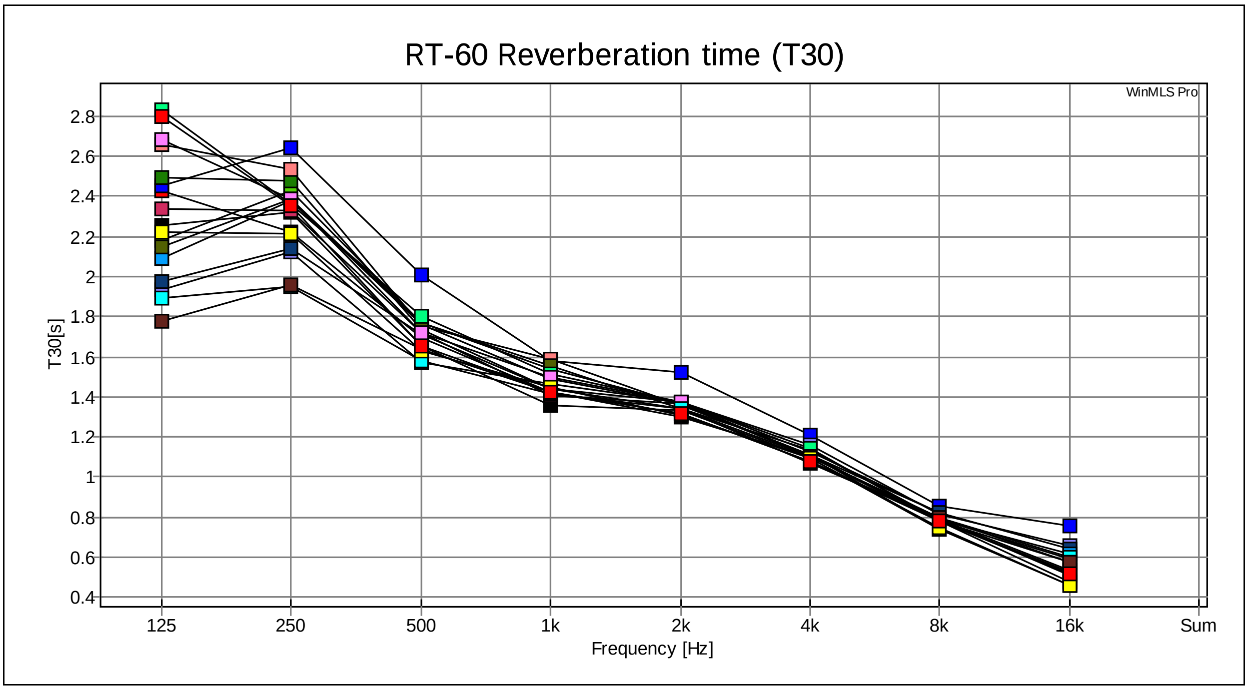
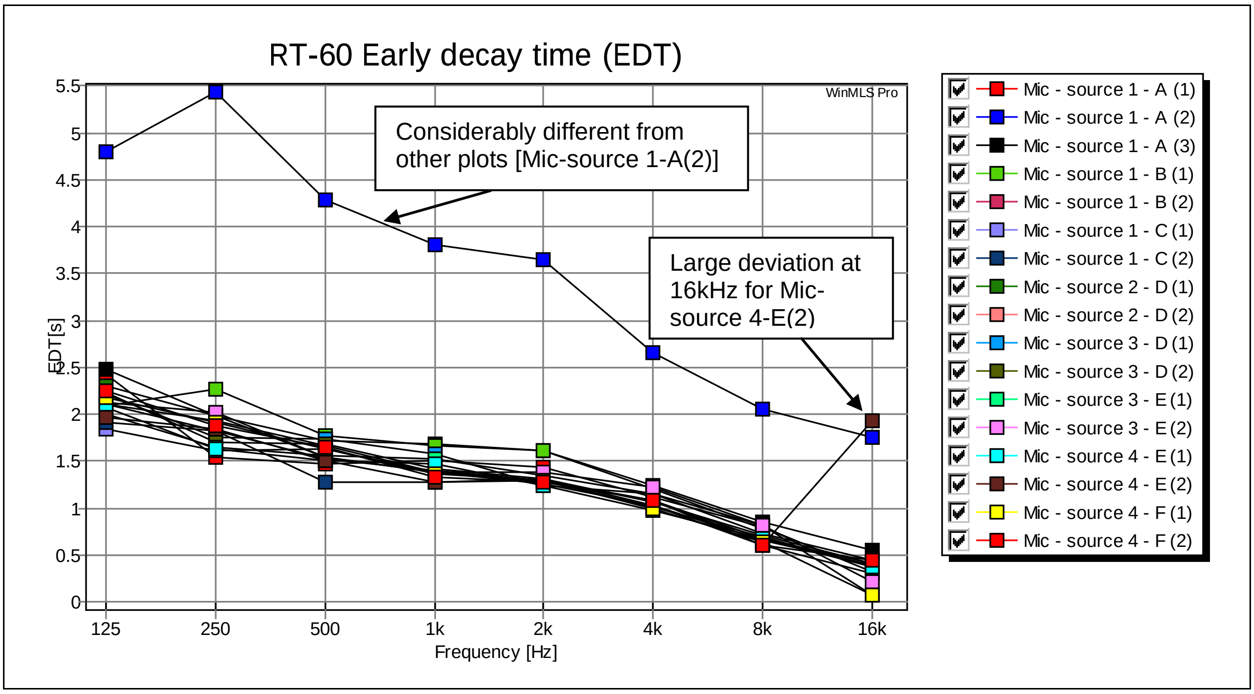


Figure 4.26: Revised Octave RT_{60} plot excluding any points with EDR less than 25dB.



The EDT plot is shown in the next figure. The plot has been set to display only points that had an EDR above 25dB. There was an impulse plot that differed considerably from the rest, namely, mic-source 1-A (2) which is plotted in blue.

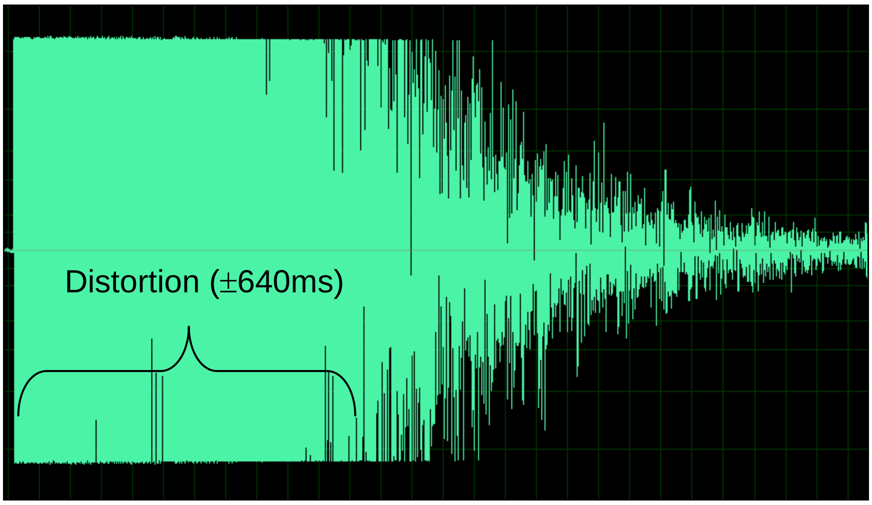
Figure 4.27: Revised Octave EDT plot excluding any points with EDR less than 25dB.



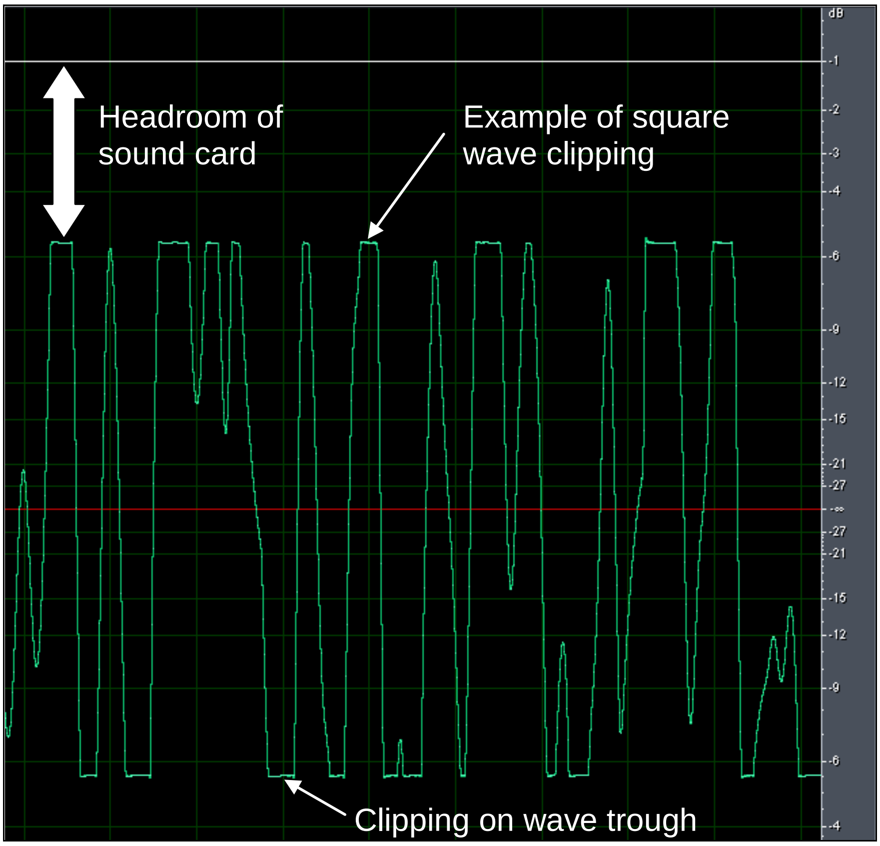
Looking at the revised octave RT plots more closely, it can be deduced that the plot labelled as 1-A(2) stands out from the rest especially in the EDT plot. While there should be some deviation between the plots especially seeing that they are taken in different positions, plot 1-A(2) differs from its partners 1-A(1) and 1-A(3) which were recorded in the same location. It was decided to inspect the original wave files for that position. On closer inspection, it was found that this position was the first position used during the tests. The first position generally requires more time for measurements, as the setup needs to be fine-tuned and calibrated. It was also found that there were three impulses taken and that one of them had extensive distortion. The distortion can be seen as a square-wave shaped clipping (hard clipping) on the SPL crests and troughs of the impulse as shown in the next two figures. The impulse sound distorted before entering the sound card. The sound card still had 5dB headroom plus its overdrive capacity of ±1,5dB. Upon listening to the impulse, distortion could be heard for the first part of the sound of about 640ms.

Figure 4.28 a & b: Distortion on impulse test.

a) Impulse response.



b) Zoomed view of overdriven microphone response.



The reason for the distorted signal was that the microphone circuit was adjusted for the wrong range and the accompanying pressure-electric voltage circuit was overdriven. The original wave file for position A-1 had three individual impulse tests conducted. The RT_{60} and EDT are shown in the next two figures. It can be seen that it was the second impulse that was incorrect. This was the same impulse that exhibited the distortion when analysed from its original wave file. This is another reason why it is important to conduct duplicate and in this case triplicate tests in each source-microphone location.

Figure 4.29: Measurement inaccuracy for RT_{60} owing to user error.

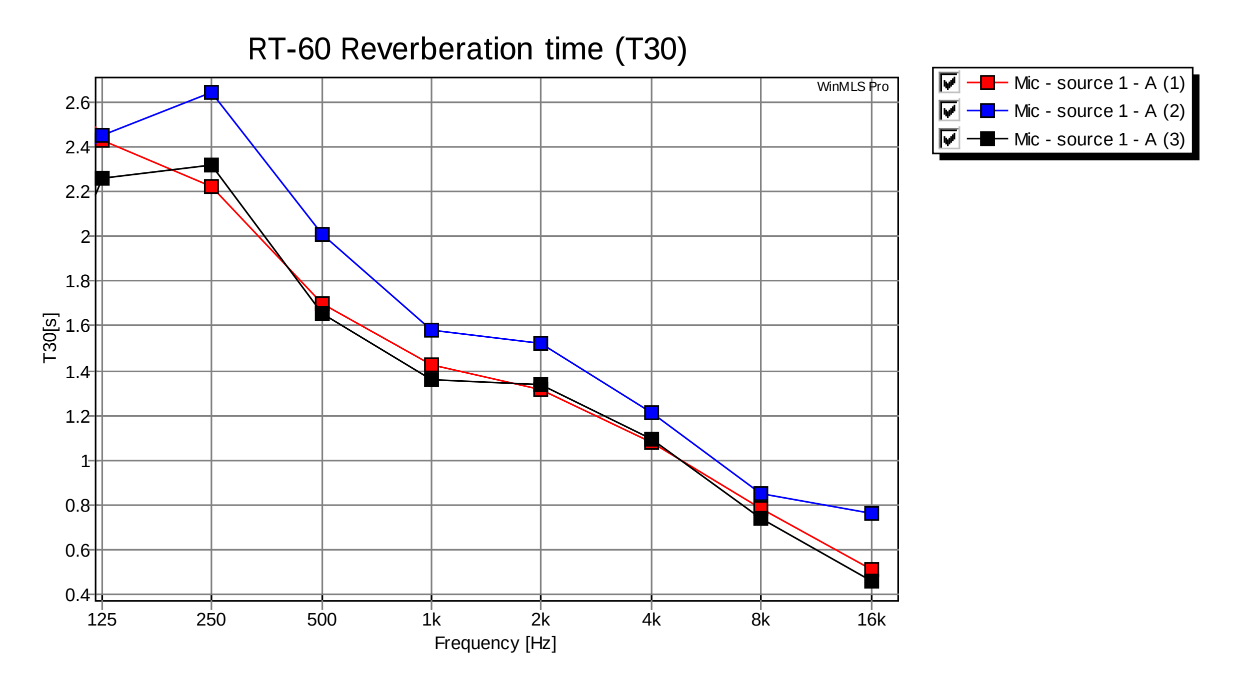
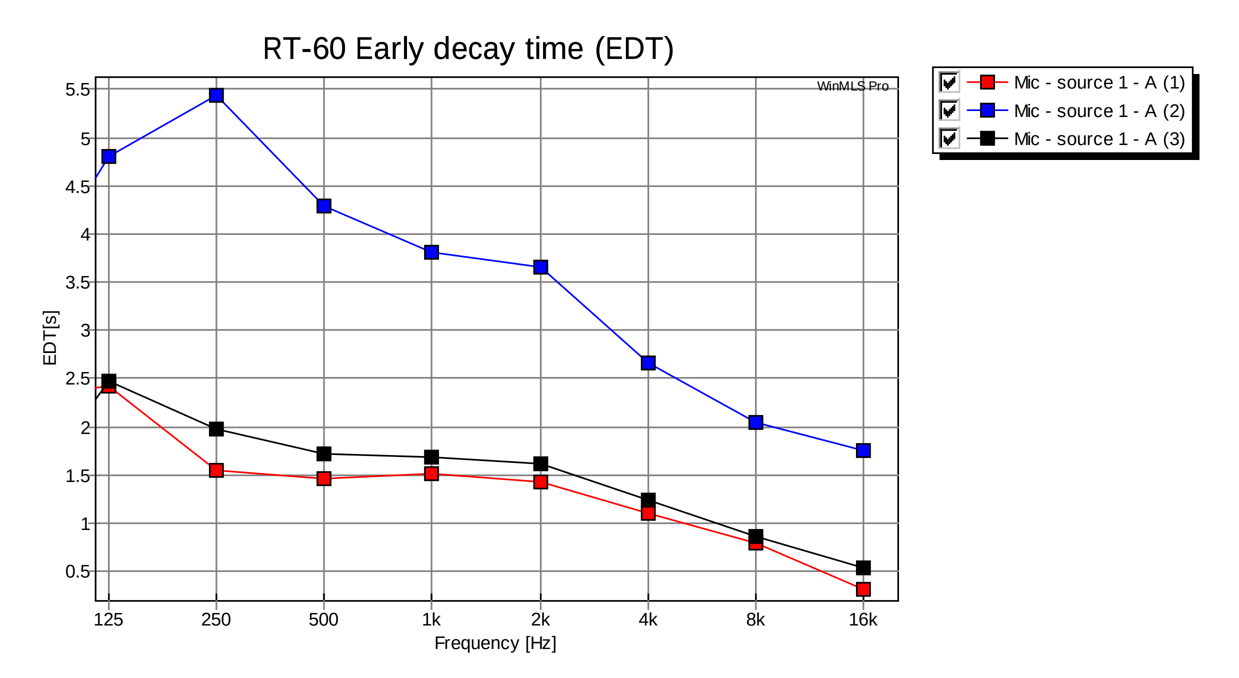


Figure 4.30: Measurement inaccuracy of EDT owing to user error.



From the two preceding figures it can be deduced that the second measurement, 1-A(2), is the problematic one. This was especially evident for the EDT test. Two points need highlighting, firstly, distorted impulse tests can offset an RT analysis and lead to inaccurate readings. Listening to the impulse is a quick method for determining whether the sound was distorted. Secondly, computer programs process the information and are limited in their ability to classify the signal as good or bad. In this case the SNR was good for all three measurements, the correlation coefficients were also good and the EDRs were similar for all three impulse tests, yet one of the measures was completely distorted. The following table summarises the data and includes the percentage deviations for the distorted measurement. The distorted measurement is italicised.

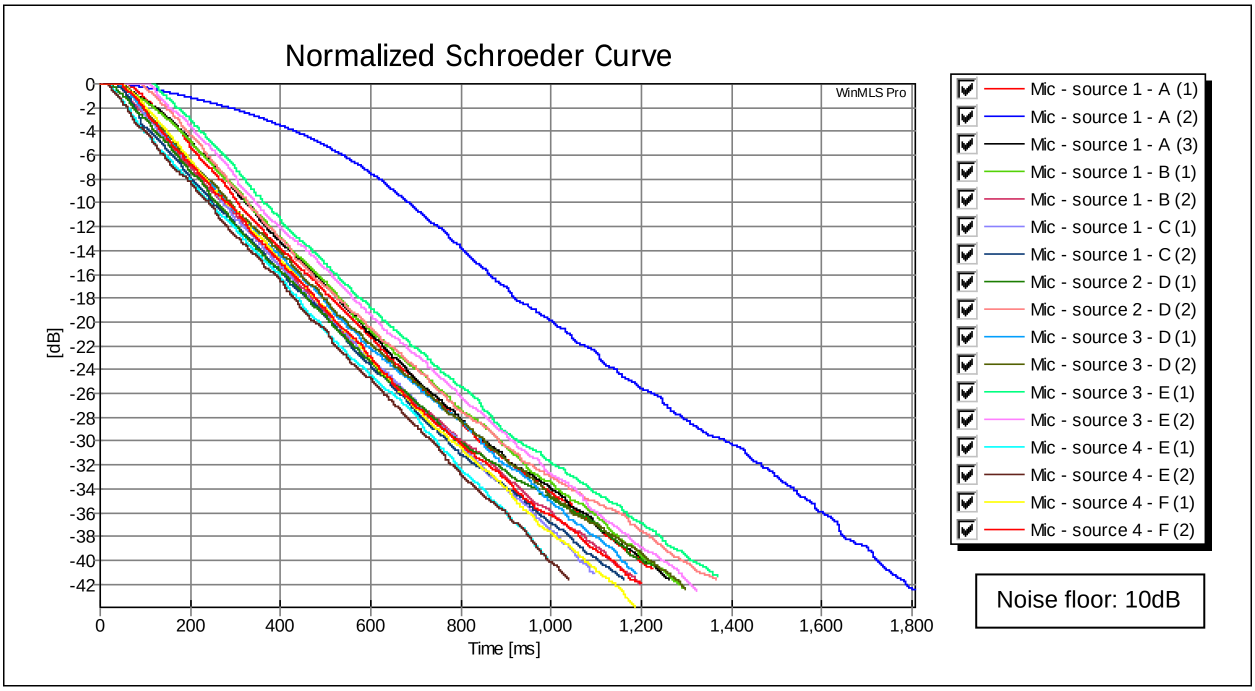
Table 4.11: Comparison of two acceptable measures with that of the error measurement.

RT_{60}	F[Hz]	125	250	500	1000	2000	4000	8000	16000
Mic – source 1 – A (1)	**RT30 [s]**	2.43	2.22	1.70	1.42	1.31	1.08	0.79	0.51
Mic – source 1 – A (3)	**RT30 [s]**	2.26	2.32	1.65	1.36	1.34	1.10	0.74	0.46
Mean	**RT30 [s]**	2.35	2.27	1.68	1.39	1.33	1.09	0.77	0.49
Mic – source 1 – A (2)	***RT30 [s]***	*2.45*	*2.64*	*2.01*	*1.58*	*1.52*	*1.21*	*0.85*	*0.76*
Percentage error	**%Error**	**4.48**	**16.30**	**20.00**	**13.67**	**14.72**	**11.01**	**11.11**	**56.70**

EDT	F[Hz]	125	250	500	1000	2000	4000	8000	16000
Mic – source 1 – A (1)	**EDT [s]**	2.42	1.54	1.47	1.51	1.43	1.11	0.8	0.32
Mic – source 1 – A (3)	**EDT [s]**	2.47	1.98	1.72	1.68	1.61	1.24	0.85	0.54
Mean	**EDT [s]**	2.445	1.76	1.595	1.595	1.52	1.175	0.825	0.43
Mic – source 1 – A (2)	***EDT [s]***	*4.8*	*5.43*	*4.28*	*3.81*	*3.65*	*2.66*	*2.05*	*1.76*
Percentage error	**%Error**	**96.32**	**208.52**	**168.34**	**138.87**	**140.13**	**126.38**	**148.48**	**309.30**

The resultant EDT percentage error was as high as 309% for the 16kHz octave band. All the octaves below had at least a 90% error. The RT_{60} error was a mixed result with a minimum error of 4,48% and a maximum of 56,7% which is still significant. It is however accepted that the octave band of 16kHz consists of the highest audio frequency components and that other factors are at play as well. One such factor would be the bandwidth of the microphone as well as the microphone circuit, which although specified to handle 20kHz, does exhibit more inaccuracy towards its upper bandwidth limit[30]. The error is shown clearly in the following Schroeder plot.

Figure 4.31: Schroeder plot of all impulse tests including the error measurement.



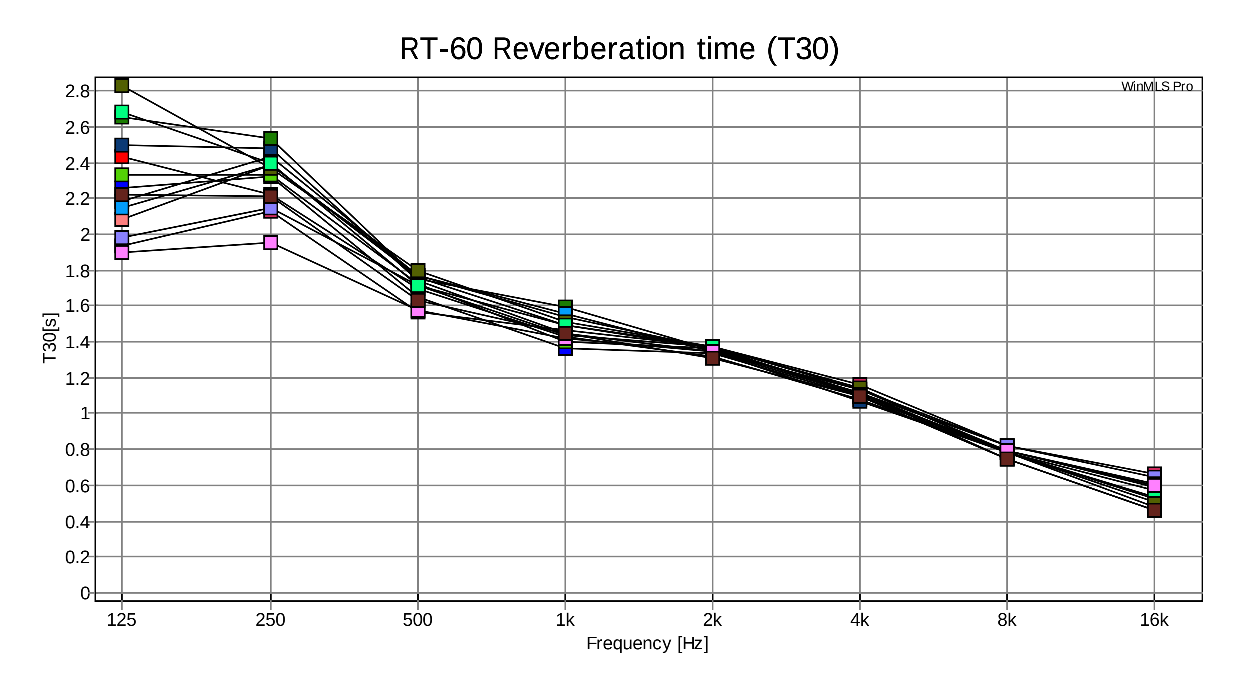
[30] The subject of the microphone's response is discussed in further detail in Appendix C.

All further analysis took place with the erroneous measurements omitted.

4.3.3.1 Comparing the Different Source-Microphone Positions

There were eight microphone-source combinations. The plot of all RT_{60} are shown in the next figure. The top view floor plan shown earlier maps out the chosen source-microphone positions. The EDT is shown in Figure 4.33.

Figure 4.32: Plot of RT_{60} for all source-microphone positions in lecture hall.



There is consistency in RT for frequencies above 400Hz. Octave bands centred at 2kHz, 4kHz and 8kHz are very close in their RT index. Frequencies below 300Hz show a wider range of RT difference. The first octave band is not plotted as the EDRs for those points were below the minimum acceptable level. The correlation coefficients of all the plots were close to or equal to negative one and the SNRs were generally high. The numerical values are shown in the next table including the mean values. The EDT values are shown in Figure 4.33.

Table 4.12: Numerical summary of RT_{60} for the lecture hall test.

x-axis F [Hz]	Mic – source 1-A (1)	Mic – source 1-A (3)	Mic – source 1-B (1)	Mic – source 1-B (2)	Mic – source 1-C (1)	Mic – source 1-C (2)	Mic – source 2-D (1)	Mic – source 2-D (2)	Mic – source 3-D (1)	Mic – source 3-D (2)	Mic – source 3-E (1)	Mic – source 3-E (2)	Mic – source 4-E (1)	Mic – source 4-E (2)	Mic – source 4-F (1)	Mic – source 4-F (2)	Mean (s)
125	2.43	2.26	2.18	2.33	1.94	1.98	2.50	2.66	2.09	2.15	2.83	2.68	1.90	1.78	2.22	2.80	2.29
250	2.22	2.32	2.43	2.33	2.13	2.14	2.48	2.54	2.38	2.39	2.36	2.39	1.95	1.96	2.21	2.35	2.29
500	1.70	1.65	1.77	1.72	1.57	1.71	1.74	1.75	1.76	1.77	1.80	1.72	1.58	1.64	1.63	1.66	1.70
1000	1.42	1.36	1.54	1.40	1.46	1.44	1.44	1.59	1.49	1.56	1.52	1.50	1.41	1.42	1.45	1.42	1.46
2000	1.31	1.34	1.37	1.37	1.37	1.37	1.34	1.35	1.36	1.34	1.37	1.37	1.34	1.30	1.31	1.32	1.35
4000	1.08	1.10	1.11	1.11	1.16	1.13	1.07	1.11	1.10	1.13	1.15	1.10	1.10	1.10	1.09	1.08	1.11
8000	0.79	0.74	0.74	0.82	0.82	0.82	0.79	0.78	0.79	0.79	0.78	0.78	0.80	0.79	0.74	0.78	0.78
16000	0.51	0.46	---	---	0.66	0.64	0.54	0.48	0.61	0.59	0.52	0.57	0.60	0.57	0.46	0.52	0.55

Figure 4.33: Plot of EDT for all effective source-microphone positions in lecture hall.

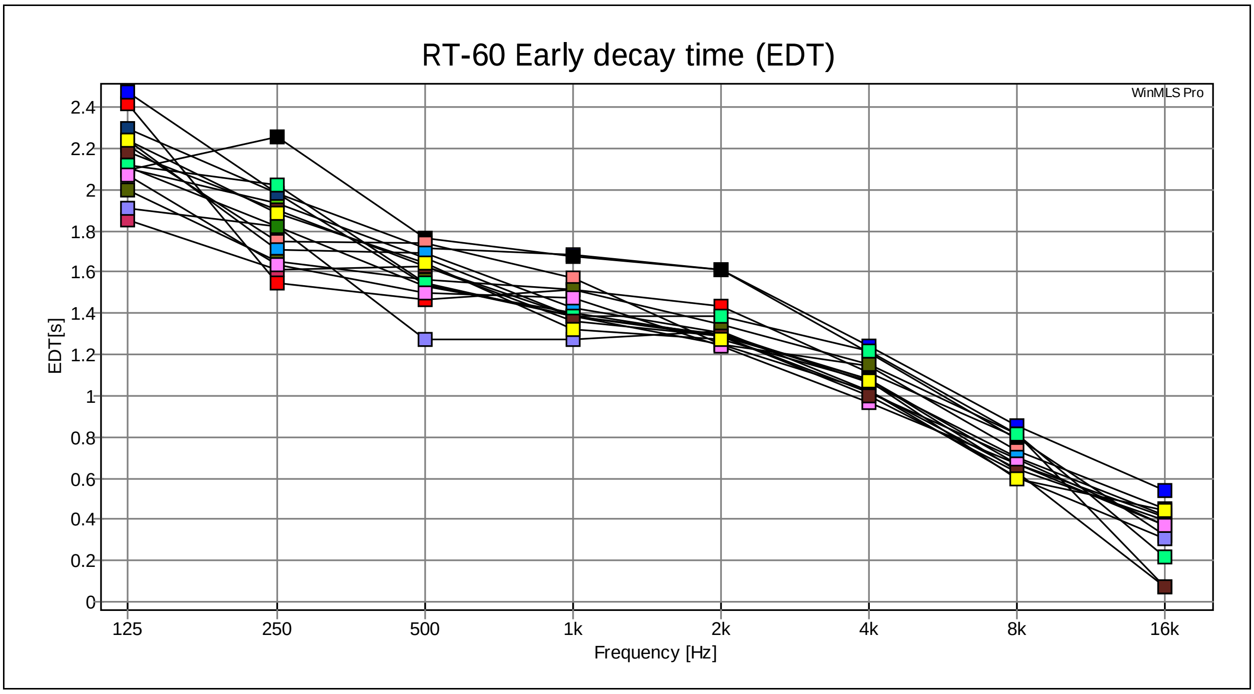
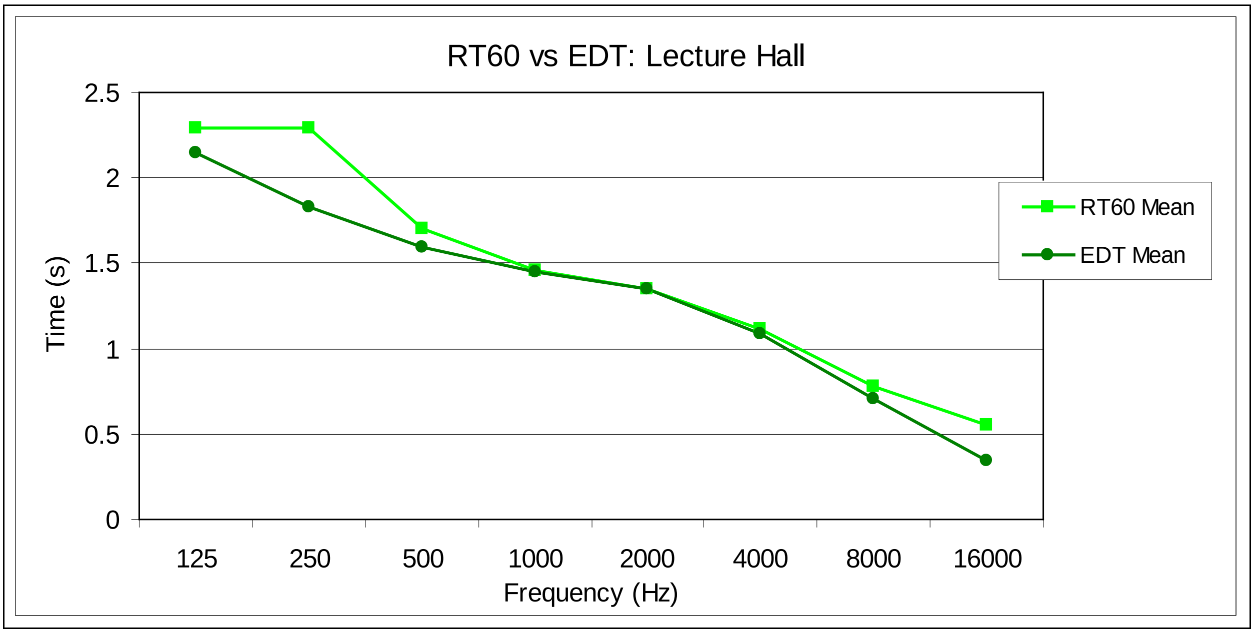


Table 4.13: Numerical summary of EDT for the lecture hall test.

x-axis F [Hz]	Mic – source 1-A (1)	Mic – source 1-A (3)	Mic – source 1-B (1)	Mic – source 1-B (2)	Mic – source 1-C (1)	Mic – source 1-C (2)	Mic – source 2-D (1)	Mic – source 2-D (2)	Mic – source 3-D (1)	Mic – source 3-D (2)	Mic – source 3-E (1)	Mic – source 3-E (2)	Mic – source 4-E (1)	Mic – source 4-F (1)	Mic – source 4-F (2)	Mean (s)
125	2.42	2.47	2.10	2.11	1.85	1.91	2.30	2.11	2.22	2.24	1.99	2.12	2.07	2.19	2.24	2.15
250	1.54	1.98	2.26	1.93	1.61	1.82	1.98	1.82	1.75	1.70	1.65	2.02	1.63	1.90	1.88	1.83
500	1.47	1.72	1.76	1.66	1.62	1.27	1.54	1.53	1.74	1.69	1.57	1.55	1.50	1.63	1.64	1.59
1000	1.51	1.68	1.67	1.38	1.38	1.27	1.39	1.40	1.57	1.42	1.52	1.39	1.47	1.36	1.32	1.45
2000	1.43	1.61	1.61	1.30	1.25	1.32	1.28	1.29	1.24	1.31	1.35	1.38	1.24	1.29	1.27	1.35
4000	1.11	1.24	1.21	1.08	1.02	1.03	1.01	1.08	1.14	1.07	1.15	1.22	0.97	1.00	1.07	1.09
8000	0.80	0.85	0.79	0.64	0.65	0.61	0.69	0.67	0.73	0.70	0.81	0.82	0.67	0.63	0.60	0.71
16000	0.32	0.54	---	---	0.39	0.31	0.37	0.41	0.45	0.42	0.08	0.22	0.37	0.07	0.45	0.34

Using the mean values and the octave centre frequencies, the effective RT and EDT for the large lecture hall can be plotted as shown in the next figure.

Figure 4.34: Resultant RT_{60} and EDT for the lecture hall shown on one set of axes for comparison.

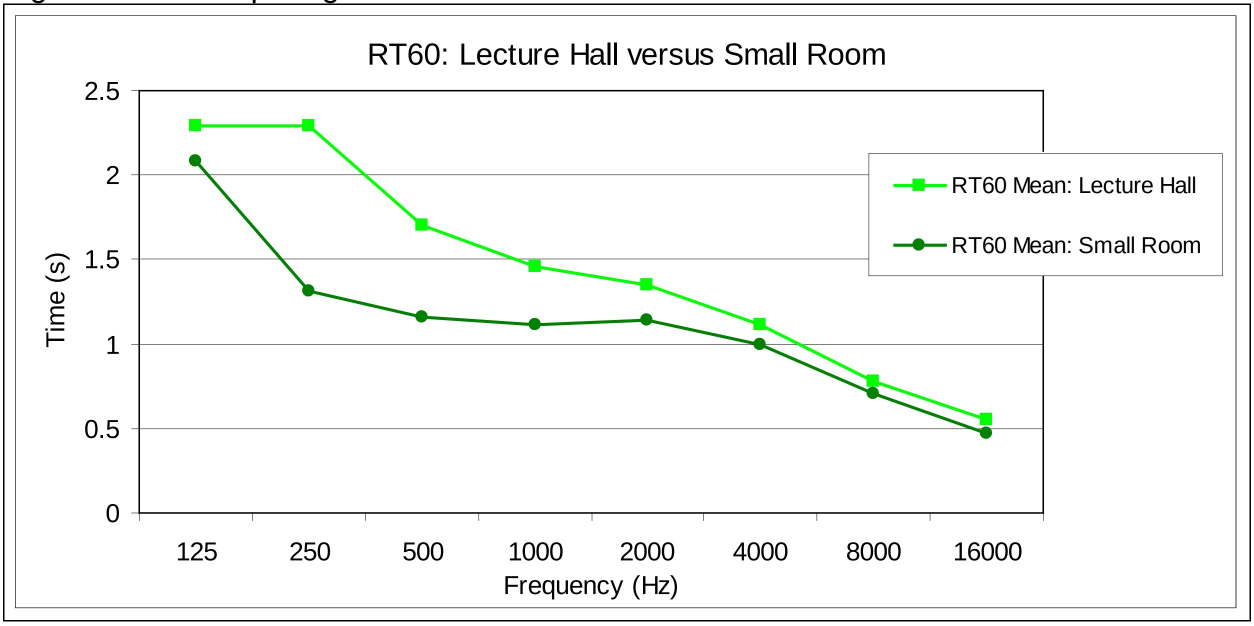


4.3.3.2 Discussion of Results for Lecture Hall

The mean EDT and RT track each other quite neatly for most octave bands excepting at 250Hz and 16kHz. The octave band RT and EDT show a negative slope as the frequency increases. The most noticeable feature of this large room was that the majority of readings for RT_{60} were similar for differing source-microphone locations, but they did start to deviate towards the two lower octave bands though. Between 2kHz and 8kHz, the RTs for all source-microphone positions were very close. It is assumed that the balloon may have not been strong enough to excite the lowest frequencies, which may account for the deviations in that range. The higher the octave band, the lower the RT and EDT for the lecture hall, which is a predictable result.

Figure 4.35 compares the RT of the lecture hall with the small room. The lecture hall and small room's RT and EDT were only similar in the 125Hz band, and the bands above 4kHz. The curvature of the octave slope between small room and lecture hall differed significantly in the low to mid bands. The lecture hall's overall reverberation and EDT are longer than that of the small room. The EDT graph is similar to that of the RT_{60} and was not included.

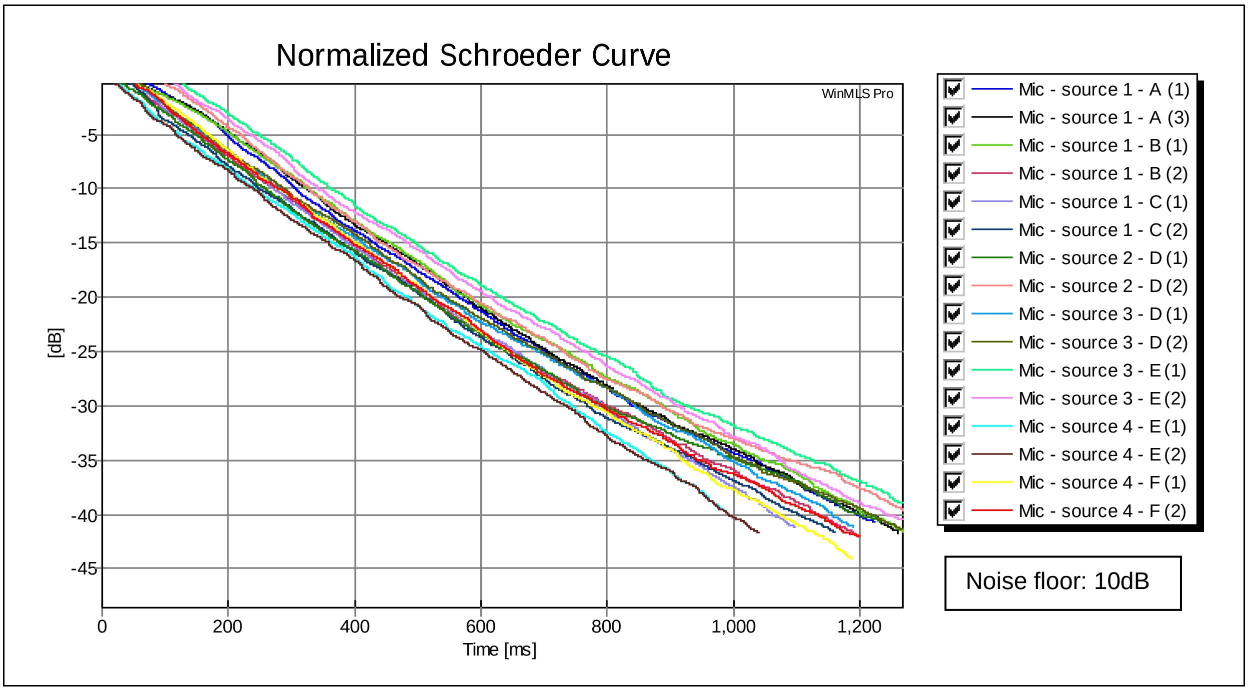
Figure 4.35: Comparing the RT of the lecture hall with the small room.



As this test was conducted in a lecture hall, it would be interesting to find out the STI and AL_{cons} for this hall. A comparison could be made between the speech intelligibility and the RT.

The Schroeder plot summarises the RT for this room shown next.

Figure 4.36: Schroeder curves plotted for all effective impulse measurements in lecture hall.



4.4 Reverberation Time Analysis for Arts Theatre

This type of location was chosen as it is an active performing music theatre that has undergone sound treatments.

Table 4.14: Summary of test setup data for third echoic location – Arts Theatre.

	Measurement 3: Arts Theatre
Date:	25 June 2009
Time:	12:15-16:30
Site:	Auckland Park, JHB (UJ) Arts Theatre Kingsway Campus
Room layout	See photos and top-view line drawing – Fig.4.37 & 4.38
Volume of room	±6800m^3 (measured by walking along the widths of the hall, estimating the height by counting cement slabs and bricks, then subtracting the protruding areas)
Condition of room	Large stage area with audience seated in half-moon curvature around the stage front. Balcony with 5 seating rows. Seats covered with absorptive material. Stage curtains were up and the side curtains tied out of the stage area. The stage area was clear with some props packed against the far wall.
Environmental conditions: Temperature, relative humidity	Clear cool day 16°C RH:42%
Degree of precision (Coverage)	Engineering method (ISO, 2008)
Measuring height above ground/floor plane:	1,3m
Distance between source and SPL meter:	Variable, but ≥2m
States of occupancy	Unoccupied (less than four persons in the test room)[31] (seats are covered with absorptive material) Stage: completely unoccupied
Spatial averaging	Arithmetic mean of RTs for all source-microphone positions for the room as a whole as well as separate areas of the enclosed space identified and averaged for area specific RT data.
Background noise level	43-45dB
Peak SPL of impulse	99dB

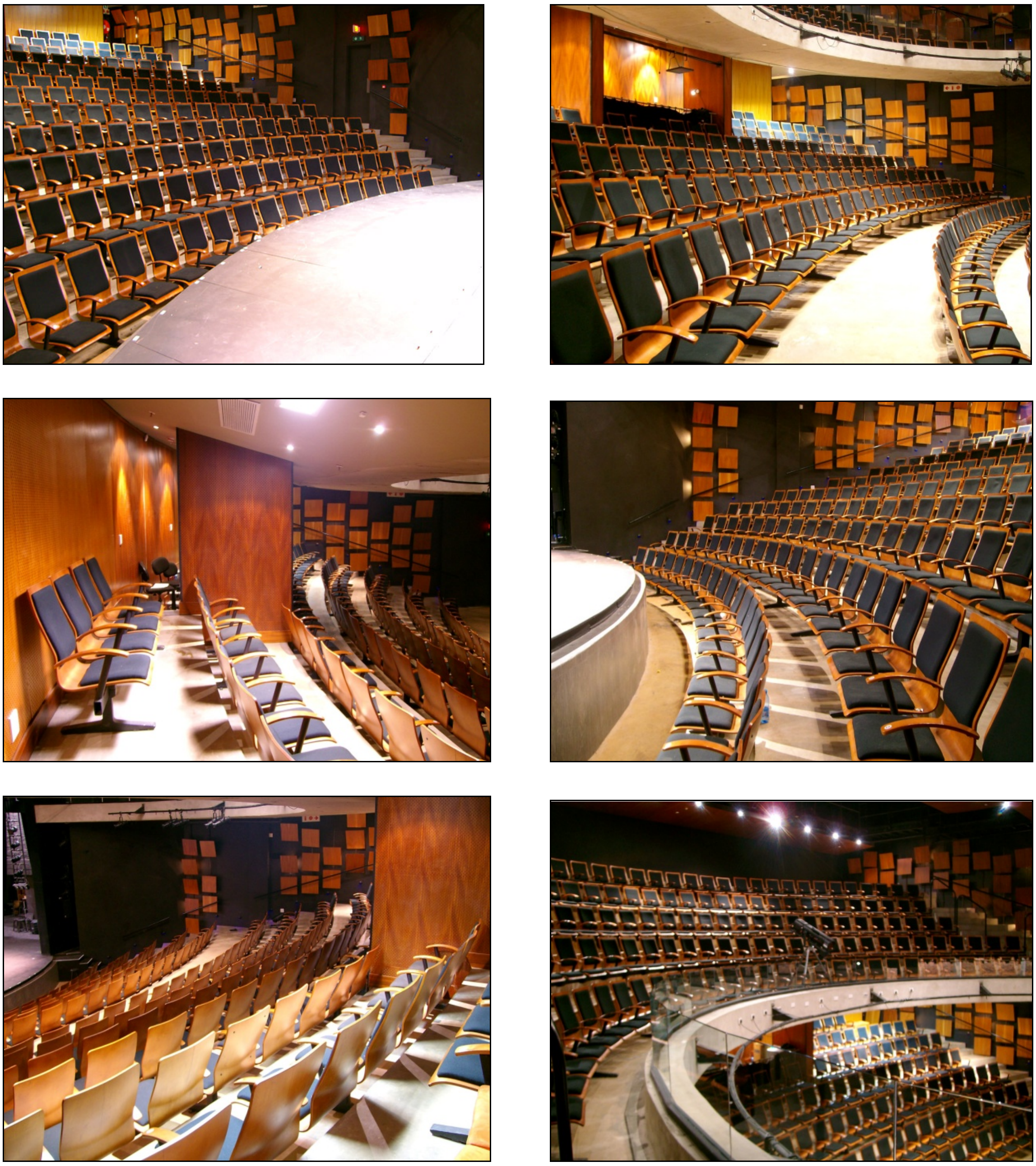
4.4.1 Room Description and Layout

The room was specifically designed to be used as an arts theatre and thus has been equipped with the necessary acoustic and staging equipment. The walls are made of cement but have acoustic director boards positioned along the sidewalls. The stage area has a notably high roof. The room layout of the seating area exhibits a slight horse shoe. The audience's seats are fastened to a cement floor that inclines. There is a balcony overhang with a similar seating pattern to that of the lower level seats. There are four rows, which are under the balcony overhang.

[31] The ISO 3382:2008 standard allow for a maximum of two persons present within the room for the label unoccupied to be applied (p:2). As the room was large and only three people were present, the label unoccupied was still chosen.

Figure 4.37 a & b: Photos of room layout for arts theatre.

a) View of seating area.

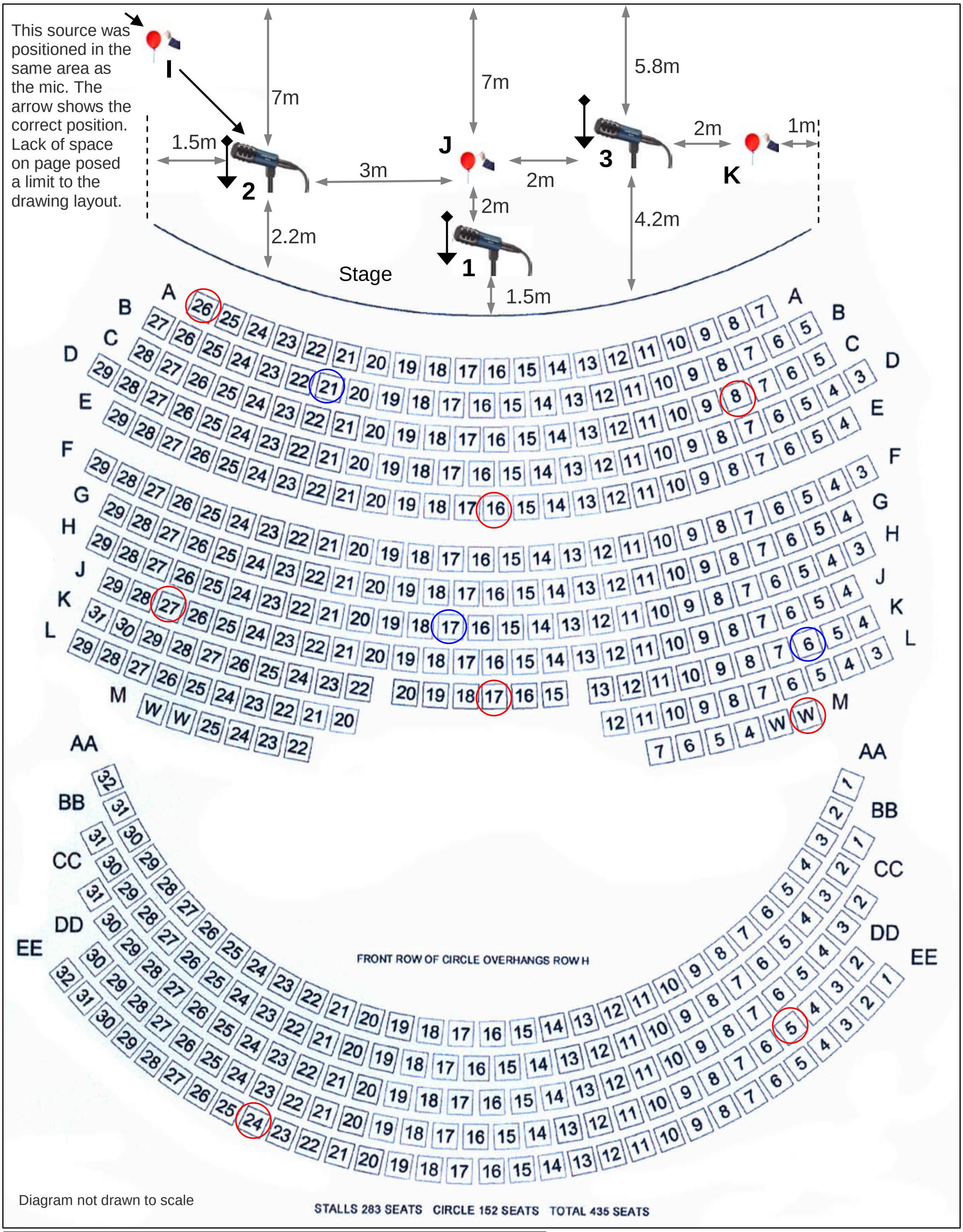


b) View of stage area.

The stage was empty for this study. When doing RT tests one should try to get the test location as close to the performance layout as possible. Thus ideally, if the performances were orchestral, then the RT measurements should be conducted with all the chairs and musicians present. In terms of this study, it was not necessary as the results from this chapter are to be compared with a vocal recording where the stage and hall occupancy will remain unchanged.

The source-microphone positions have been plotted using the floor-seating plan. The stage area shows the microphone and balloon icons as used in the previous two rooms; however, as there was insufficient space on the page, blue and red circles were used as equivalents for the audience area instead of the icons. The positions were coordinated using the seat and row numbers and thus if a microphone was placed in row B seat number 21, then that seat was circled in blue. All the positions are tabulated in Table 4.15 and are needed together with the top view layout figure to locate the individual test positions.

Figure 4.38: Top view of test room showing room layout for arts theatre.



Key to diagram:

I,J and K: Stage source positions

1,2, and 3: Stage microphone-locations

Source locations in the audience area

Microphone locations in the audience area

Microphone direction

The source-microphone positions are plotted in the table below and are referenced to the seat numbers. Using the seat number, the source-microphone positions can be located.

Table 4.15: Measurement position combinations for Arts theatre.

Measurement combinations	**Row in audience**	**Seat number in audience**	**Stage position**	**Measurements identifier (filename)**	**Test part a, b or c** (See note[32])
1 – A	A	26	1	1 – A (A26)	b
1 – B	E	16	1	1 – B (E16)	b
1 – C	C	8	1	1 – C (C8)	b
1 – D	M	W	1	1 – D (MW)	b
1 – E	K	17	1	1 – E (K17)	b
1 – F	J	27	1	1 – F (J27)	b
1 – G	DD	5	1	1 – G (DD5)	b
1 – H	EE	24	1	1 – H (EE24)	b
1 – I	---	---	1 – I	1 – I (stage)	a
1 – J	---	---	1 – J	1 – J (stage)	a
1 – K	---	---	1 – K	1 – K (stage)	a
2 – A	A	26	2	2 – A (A26)	b
2 – B	E	16	2	2 – B (E16)	b
2 – C	C	8	2	2 – C (C8)	b
2 – D	M	W	2	2 – D (MW)	b
2 – E	K	17	2	2 – E (K17)	b
2 – F	J	27	2	2 – F (J27)	b
2 – G	DD	5	2	2 – G (DD5)	b
2 – H	EE	24	2	2 – H (EE24)	b
2 – J	---	---	2 – J	2 – J (stage)	a
2 – K	---	---	2 – K	2 – K (stage)	a
3 – A	A	26	3	3 – A (A26)	b
3 – B	E	16	3	3 – B (E16)	b
3 – C	C	8	3	3 – C (C8)	b
3 – D	M	W	3	3 – D (MW)	b
3 – E	K	17	3	3 – E (K17)	b
3 – F	J	27	3	3 – F (J27)	b
3 – I	---	---	3 – I	3 – I (stage)	a
3 – J	---	---	3 – J	3 – J (stage)	a
3 – K	---	---	3 – K	3 – K (stage)	a
4 – K	H	17	K	4 – K (H17)	c
5 – K	K	6	K	5 – K (K6)	c
6 – K	B	21	K	6 – K (B21)	c

4.4.2 Measurements

A survey of the room was conducted. A total of 33 measurement position configurations were undertaken. Each position had at least two impulse tests conducted, thus a total of

[32] Reverberation tested from different perceptual areas, namely "a" – stage only; "b" – stage (sound source in audience and microphone on stage); "c" – audience (sound source from stage area but microphone in various audience positions). Explained further under the results section.

at least 66 impulse tests were conducted in the theatre. The two impulses per source-microphone position were averaged and thus each position only shows one plot/result in the following analysis. The majority of impulse tests were conducted with the sound source being on the stage. According to Bradley (2005:173), the EDT often exhibits a different decay curve for different locations within a hall. While Bradley found that the average EDTs tended to be lower than the RT_{60}, they both are often found to differ from the theoretical design when compared to the measured values. The RT_{60} has a more stable response than the EDT across the hall. This difference would probably relate to different reverberance experienced by the audience in the different locations. This occurrence can be explained partially by the architectural structure such as the ceiling and walls, which may introduce focusing or diffusing effects, and partially by the nature of the absorbent materials within the hall. Bradley (2005: 173-174) believes that by plotting the relationship between EDT, sound strength G and source-receiver distance values we can better understand the effects of the shape of a hall on early-arriving reflections and highlight atypical features that may be present. It was decided to incorporate some of this data into this experiment in order to obtain a comprehensive understanding of the EDT across the arts theatre. For completeness, the RT_{60} has also been studied in the same way as the EDT, however sound strength was excluded in this analysis. Thus, for this test the measurements were divided into three sections, which has been explained further under the next heading.

4.4.3 Results

The results were split into three parts:

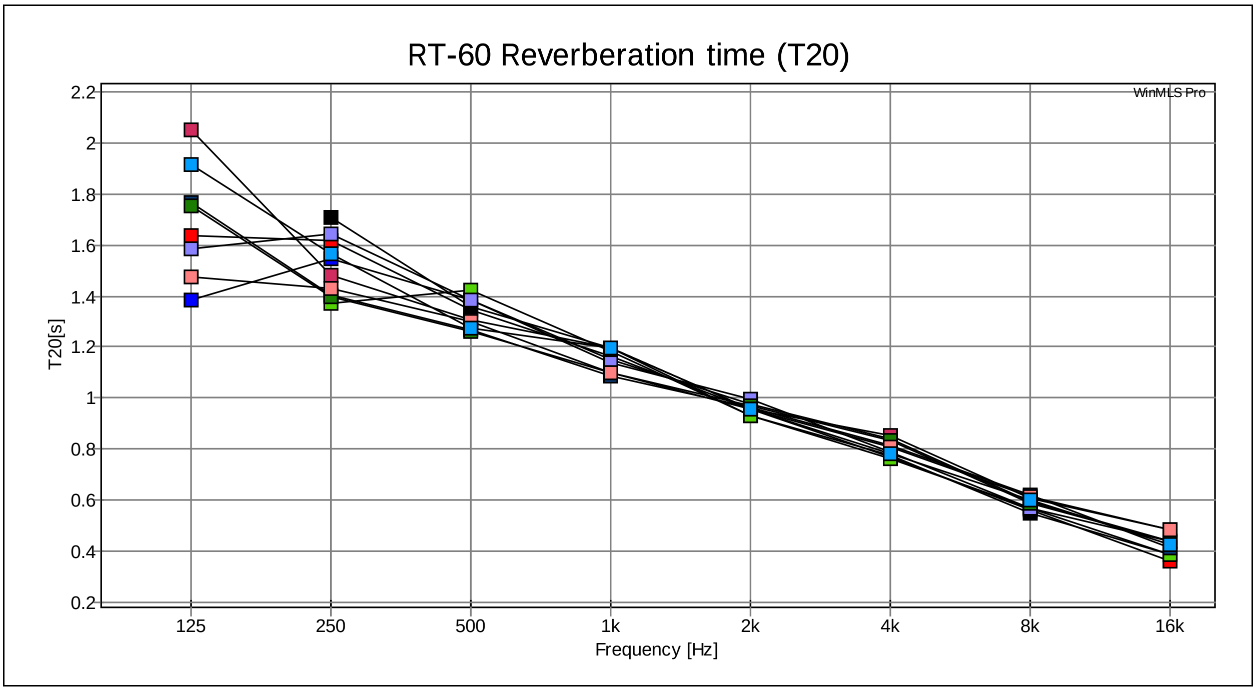
a) The RT_{60} and EDT of the stage area (sound source and microphone on stage)
b) The RT_{60} and EDT of the theatre from the perspective of stage (sound source in audience and microphone on stage)
c) The RT_{60} and EDT of the audience area (sound source from stage area but microphone in various audience positions)

4.4.3.1 Reverberation and EDT taken from stage area

The stage source-microphone combinations are labelled in the final column of Table 4.15. These positions are labelled with an “a”.

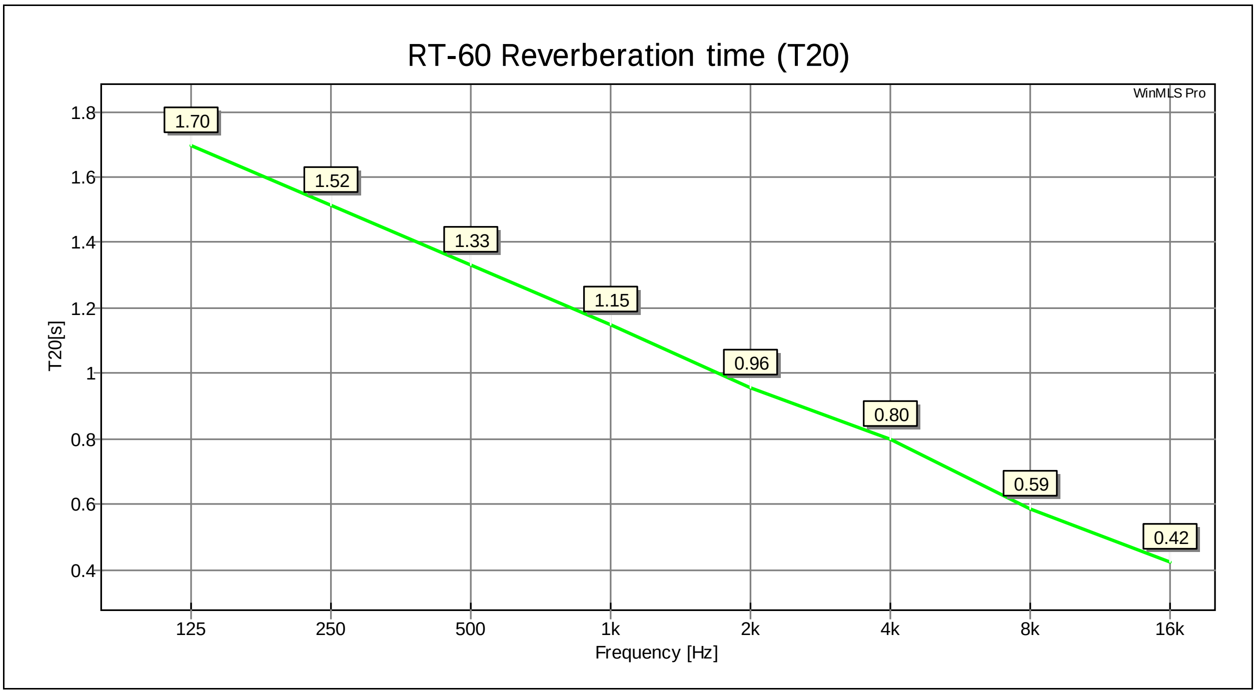
Starting with the RT_{60}, the responses of all the stage source-microphone positions were plotted and follows next.

Figure 4.39: Octave-specific RT_{60} plots of all stage source-microphone combinations.



The figure exhibits a negative sloped almost completely linear response. Owing to the larger hall size, the RT_{60} was derived from the RT_{20} as the SNR obtained was lower. There were no acceptable 63Hz plots and some 125Hz plots had their EDRs too low to consider. The mean RT curve for the stage area is shown in the next figure with the RT_{60} values labelled in the boxes above each octave band.

Figure 4.40: Octave-specific mean RT_{60} plot for stage source-microphone combinations.



The EDT plots exhibited a larger variation for the different source-microphone combinations as expected. The following figure shows the individual plots which is then followed by their mean plot in Figure 4.42.

Figure 4.41: Octave-specific EDT plots of all stage source-microphone combinations.

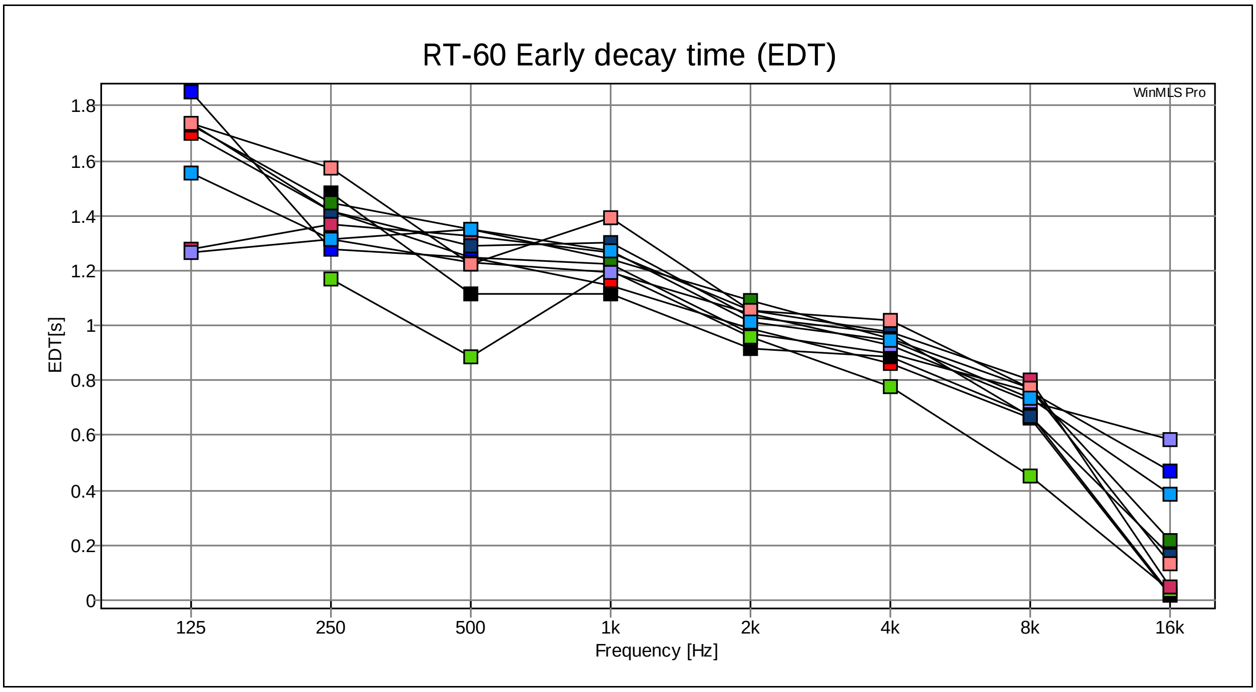
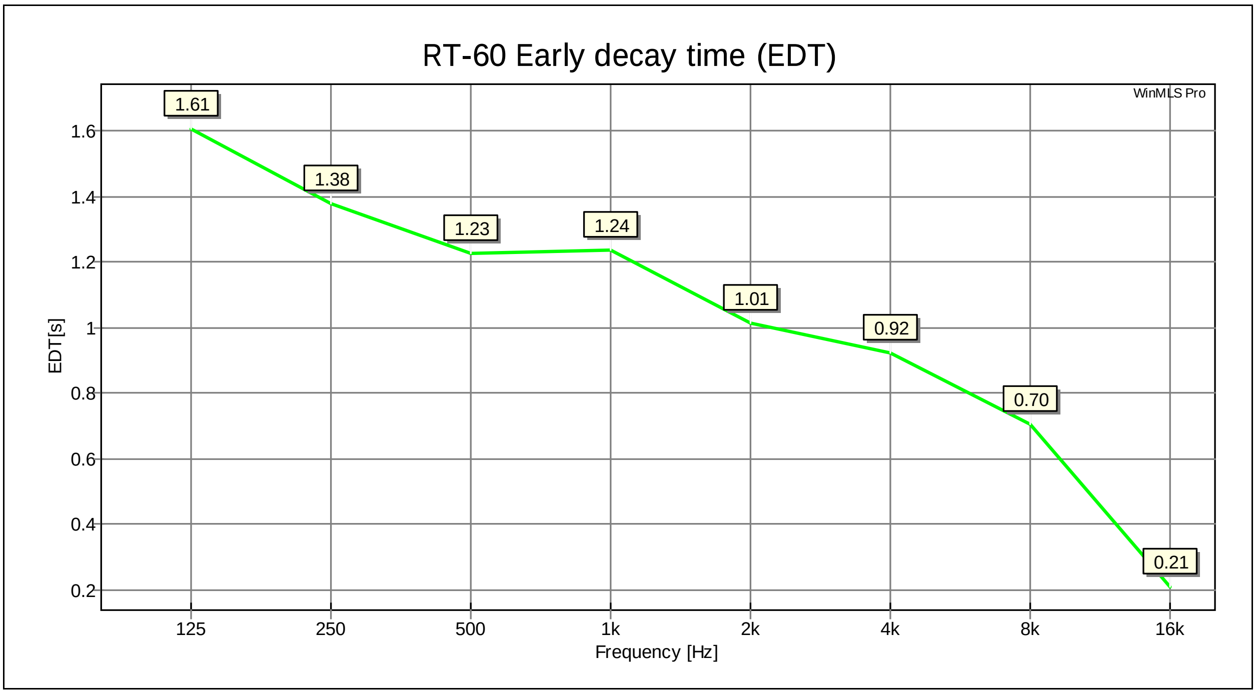


Figure 4.42: Octave-specific mean EDT plot for stage source-microphone combinations.



Comparing the RT_{60} with the EDT for the stage area follows in the next plot with the tabulated results in Table 4.16.

Figure 4.43: Comparison of the RT_{60} and EDT for stage source-microphone combinations.

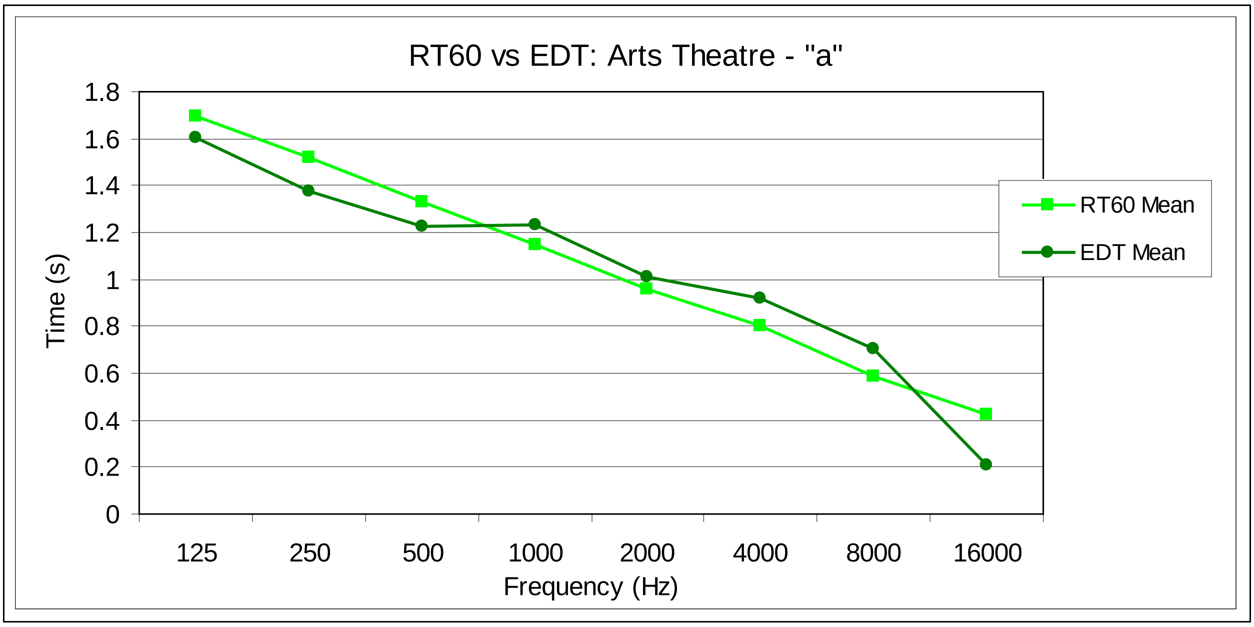


Table 4.16: Summary of mean RT_{60} and EDT results for source-microphone combinations.

F [Hz]	Mean EDT (s)	Mean RT60 (s)
125	1.61	1.70
250	1.38	1.52
500	1.23	1.33
1000	1.24	1.15
2000	1.01	0.96
4000	0.92	0.80
8000	0.70	0.59
16000	0.21	0.42

4.4.3.2 Reverberation and EDT taken from the perspective of stage (sound source in audience and microphone on stage)

The same format as the previous test has been followed. The resulting plots are shown in the next few figures. The test positions correlate to the "b" label as indicated in the final column of Table 4.15. The figures are self-explanatory and thus no narration was included.

Figure 4.44: Octave-specific RT_{60} plots recorded from the perspective of stage.

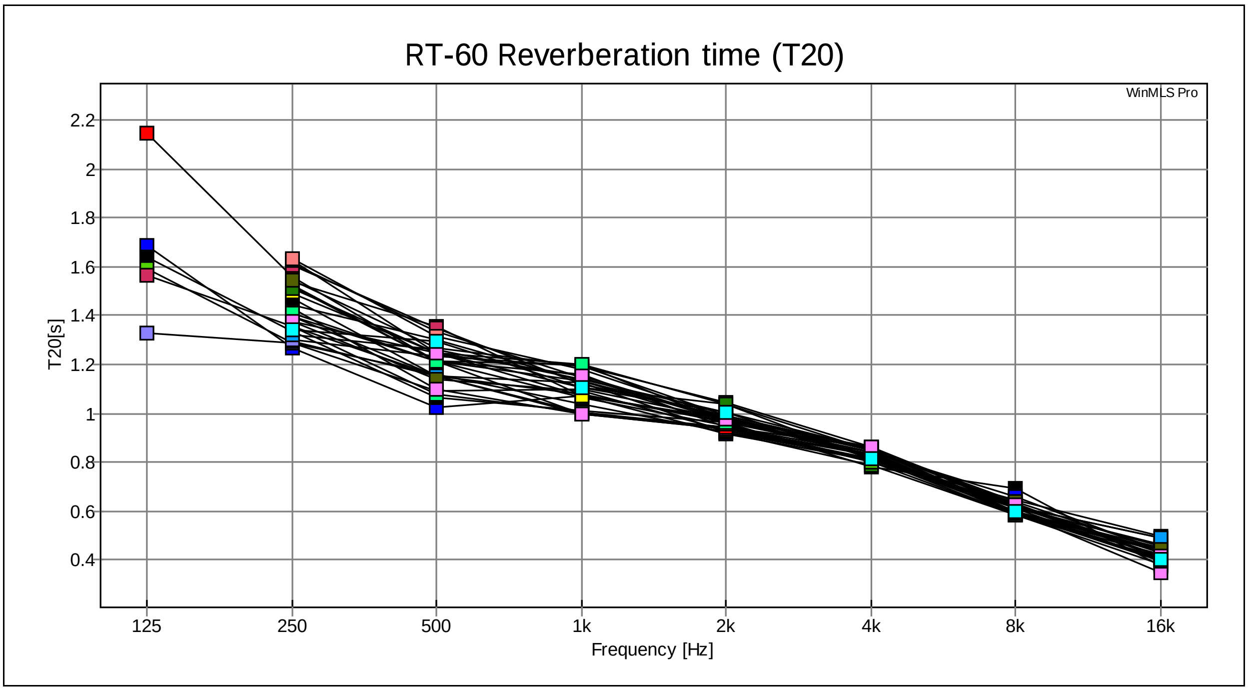


Figure 4.45: Octave-specific mean RT_{60} plots recorded from the perspective of stage.

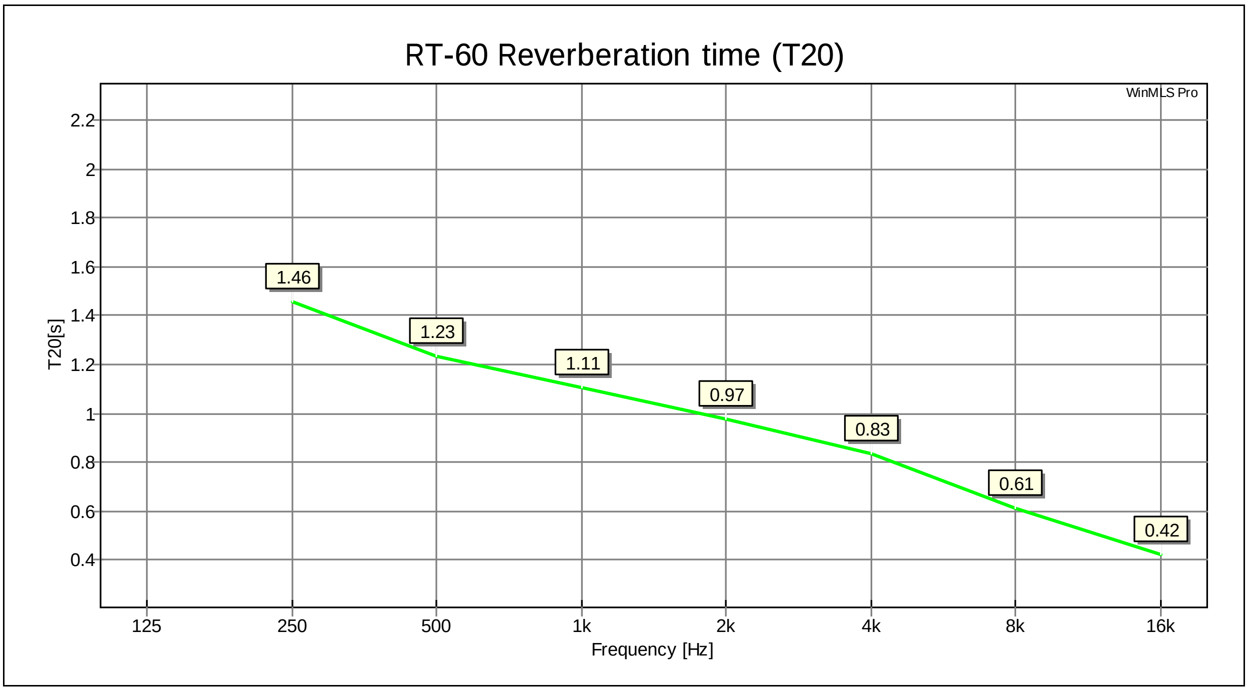


Figure 4.46: Octave-specific EDT plots recorded from the perspective of stage.

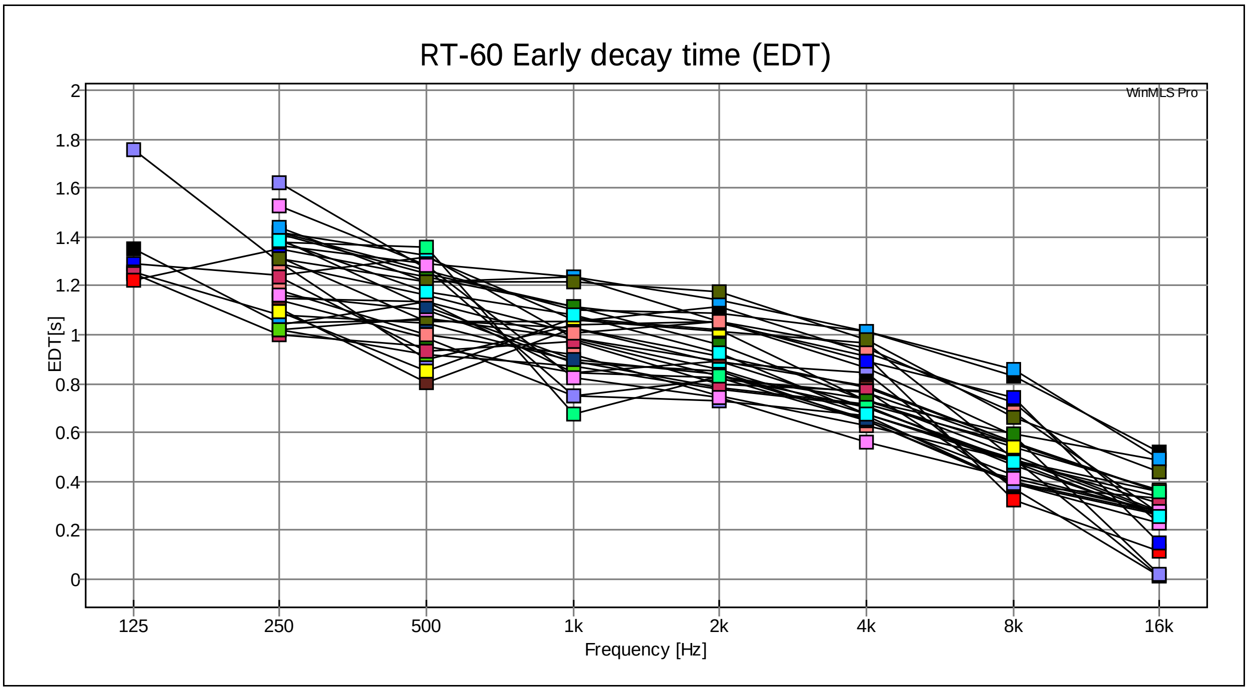


Figure 4.47: Octave-specific mean EDT plot from the perspective of stage.

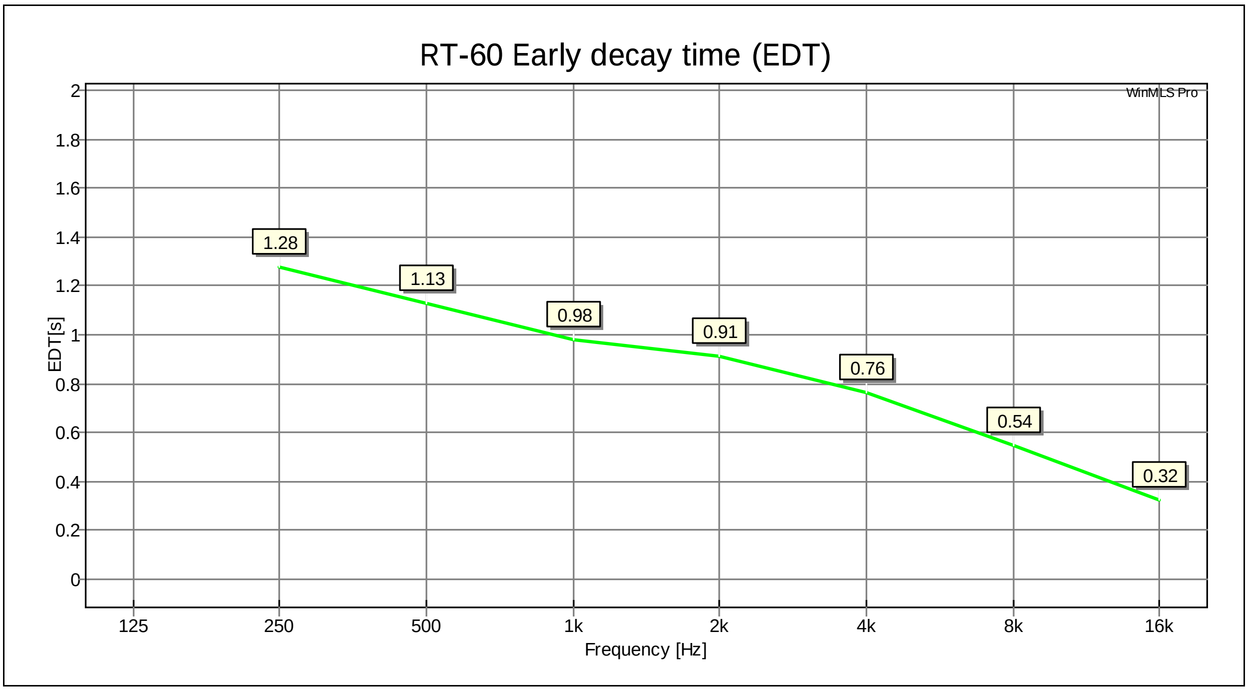
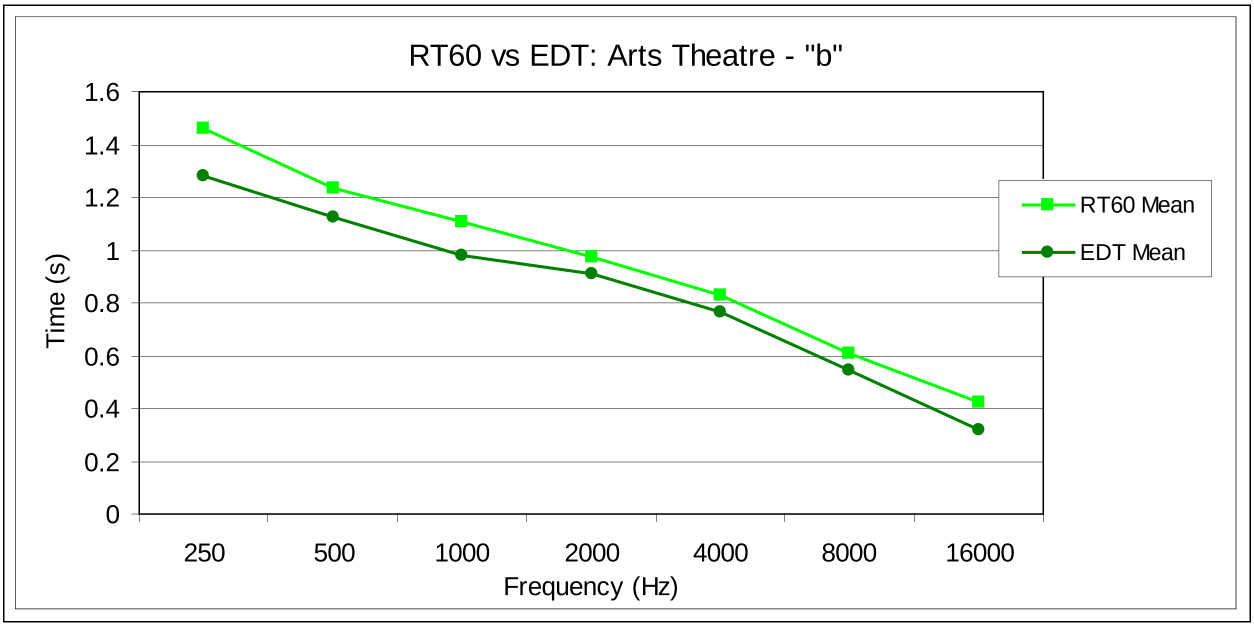


Figure 4.48: Comparison of the RT_{60} and EDT plots recorded from the perspective of stage.



4.4.3.3 Reverberation and EDT of the audience area (sound source from stage area, microphone in various audience positions)

Following the same procedure again excepting that now the test positions correlate to position “c”. The results are plotted as shown. Fewer measurements were conducted as this location was not used in the AR application.

Figure 4.49: Octave-specific RT_{60} plots recorded from the perspective of the audience.

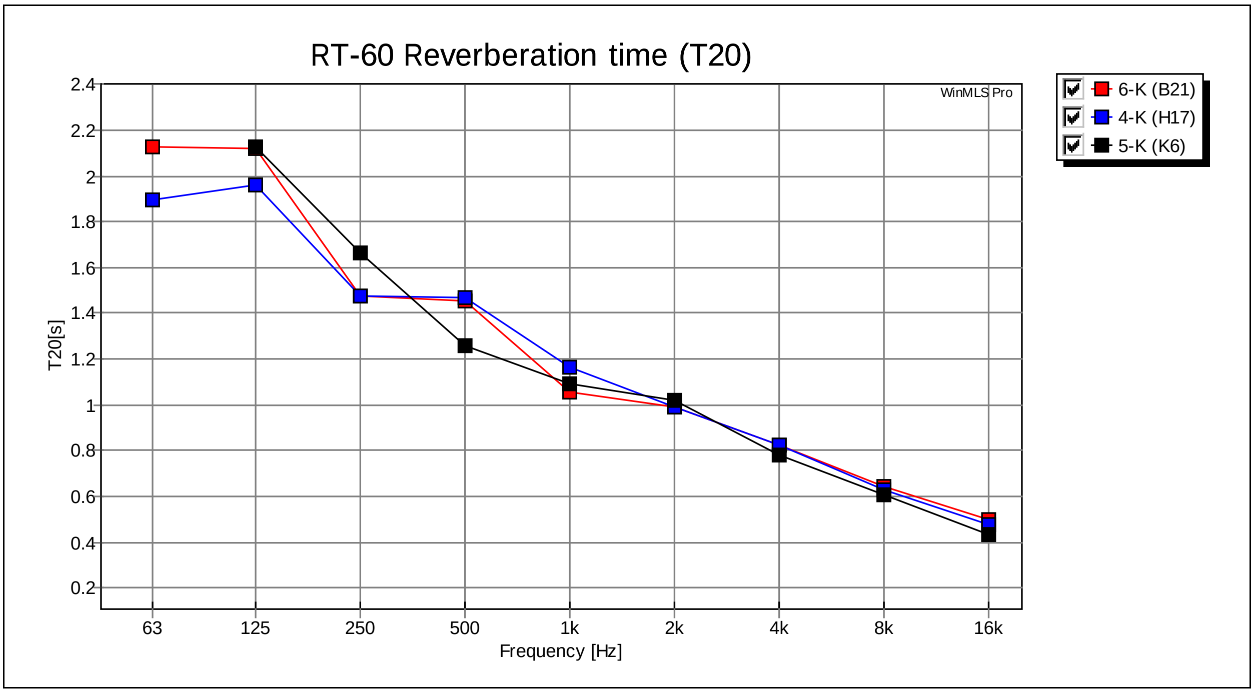


Figure 4.50: Octave-specific mean RT_{60} plots recorded from the perspective of the audience.

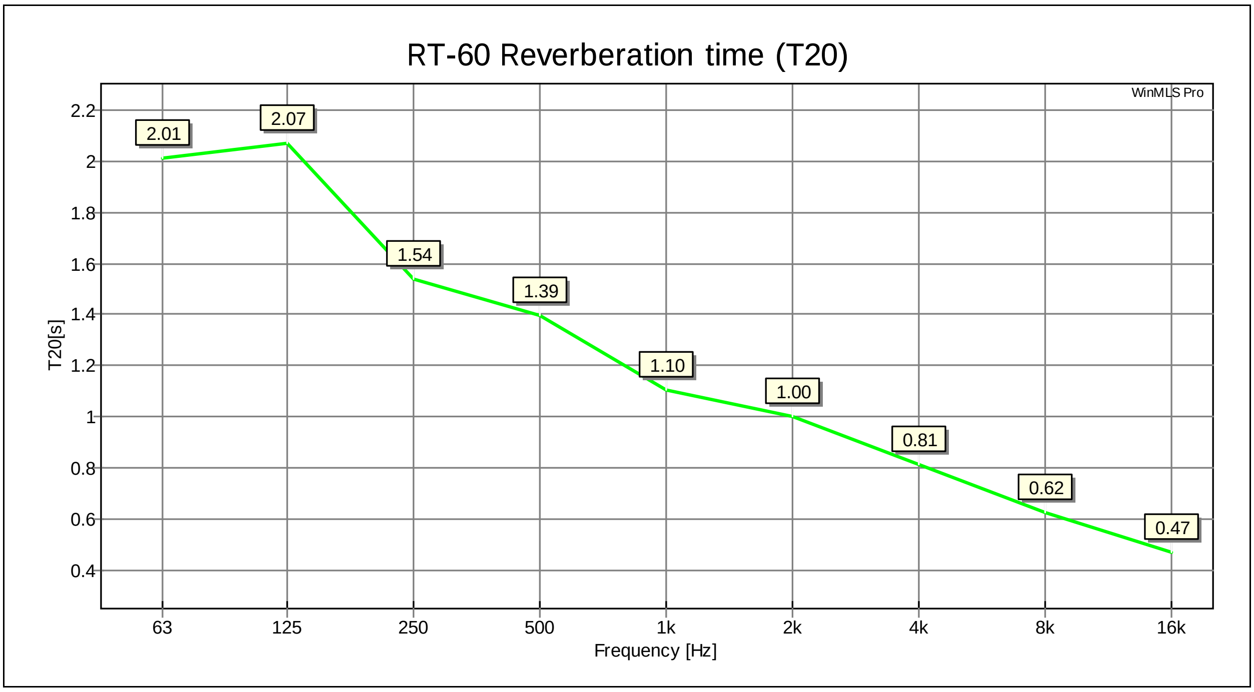


Figure 4.51: Octave-specific EDT plots recorded from the perspective of the audience.

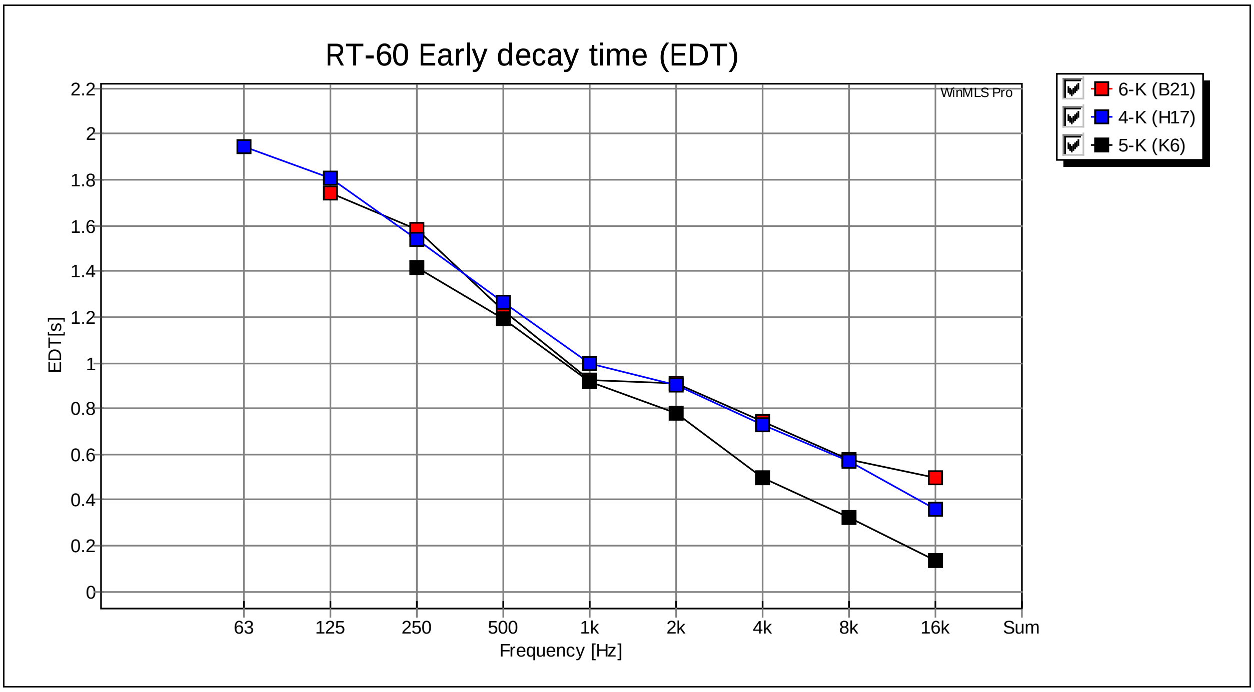


Figure 4.52: Octave-specific mean EDT plot recorded from the perspective of the audience.

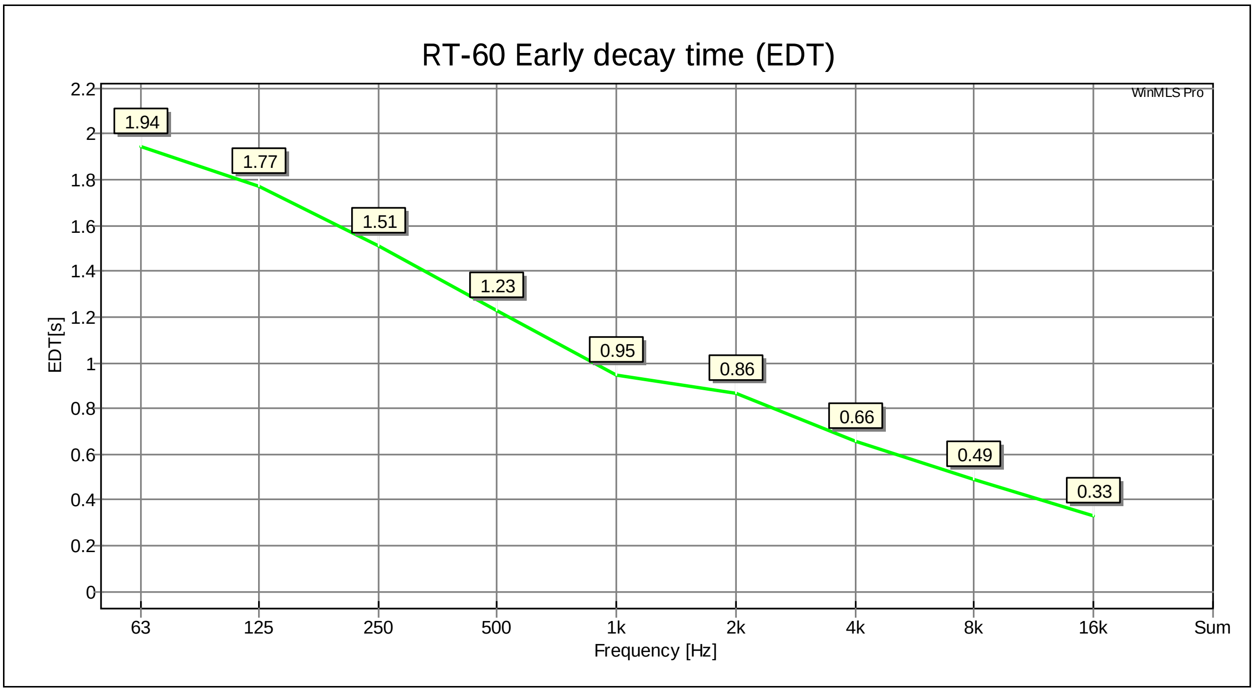
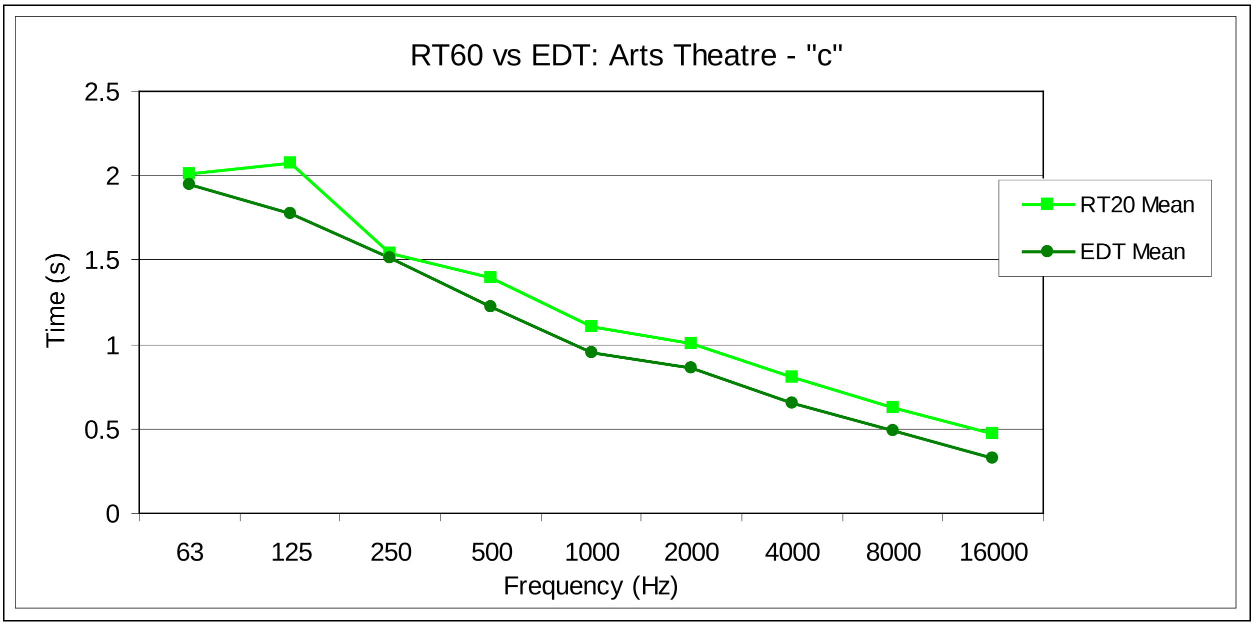


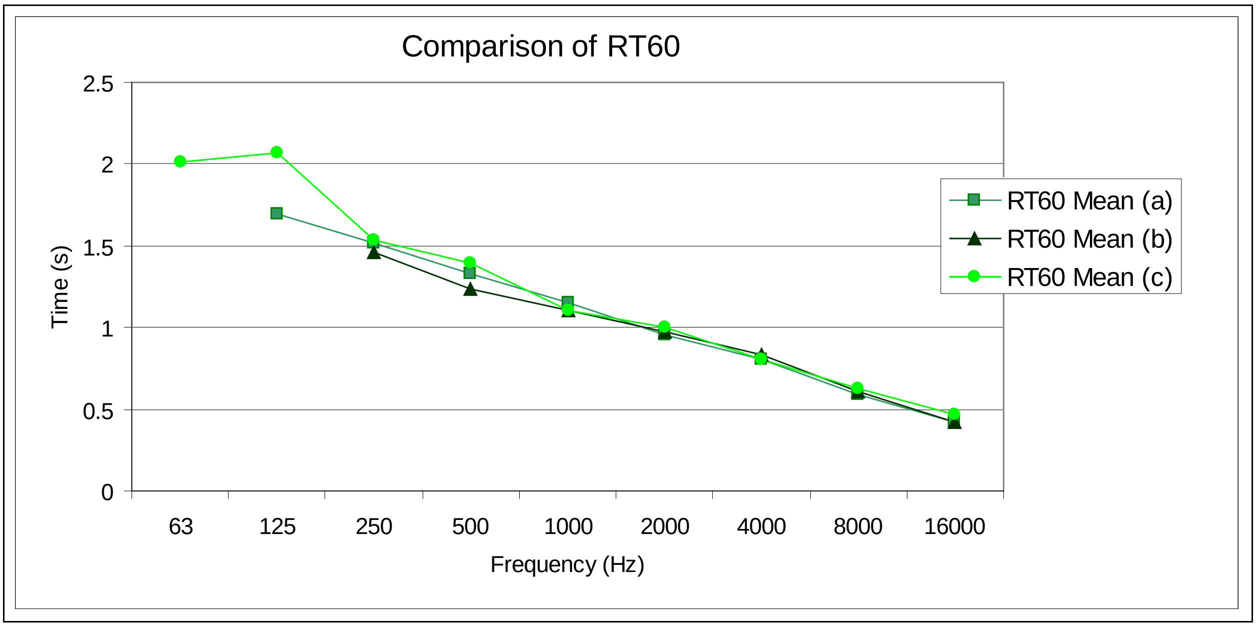
Figure 4.53: Comparison of the RT_{60} and EDT plots recorded from the perspective of the audience.



4.4.3.4 Discussion of Results for Arts Theatre

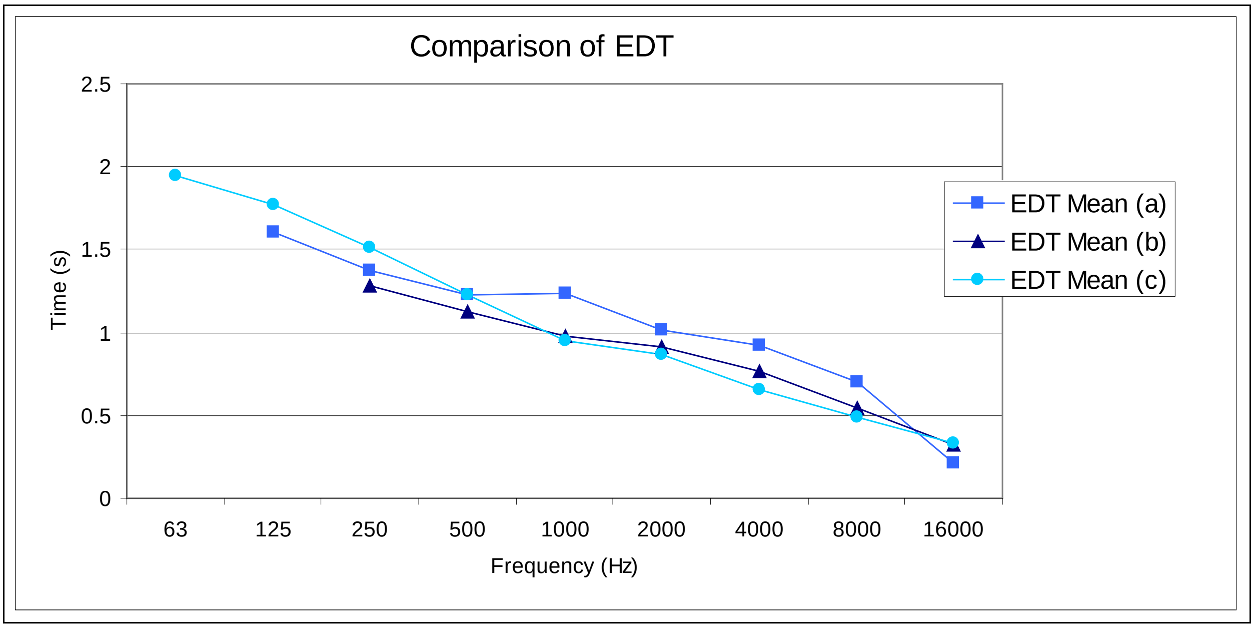
The three source-microphone test location combinations (“a”, “b” and “c”) are compared with each other in the following two plots; starting with the mean RT_{60} for each of the three theatre sections, then followed by a comparison of the mean EDT for the three sections.

Figure 4.54: Comparison of mean RT_{60} for sections a, b, and c of arts theatre.



The octave decays for all three perceptual locations were almost linear. The only noticeable deviation was found for plot "c" where the lower octave band of 125Hz shows a higher RT value than for that of "a", as well as "b" if it was extrapolated along its trajectory. The 63Hz band results seems like they would be similar as the linear slope of the three curves would statistically meet in a similar location to that of plot "c". Overall, the theatre had a similar RT the different locations. Unfortunately, the lower octave bands could not be analysed and are still important in making comments about the type of reverberant sound. It is highlighted that in large rooms, a stronger impulse source would be needed.

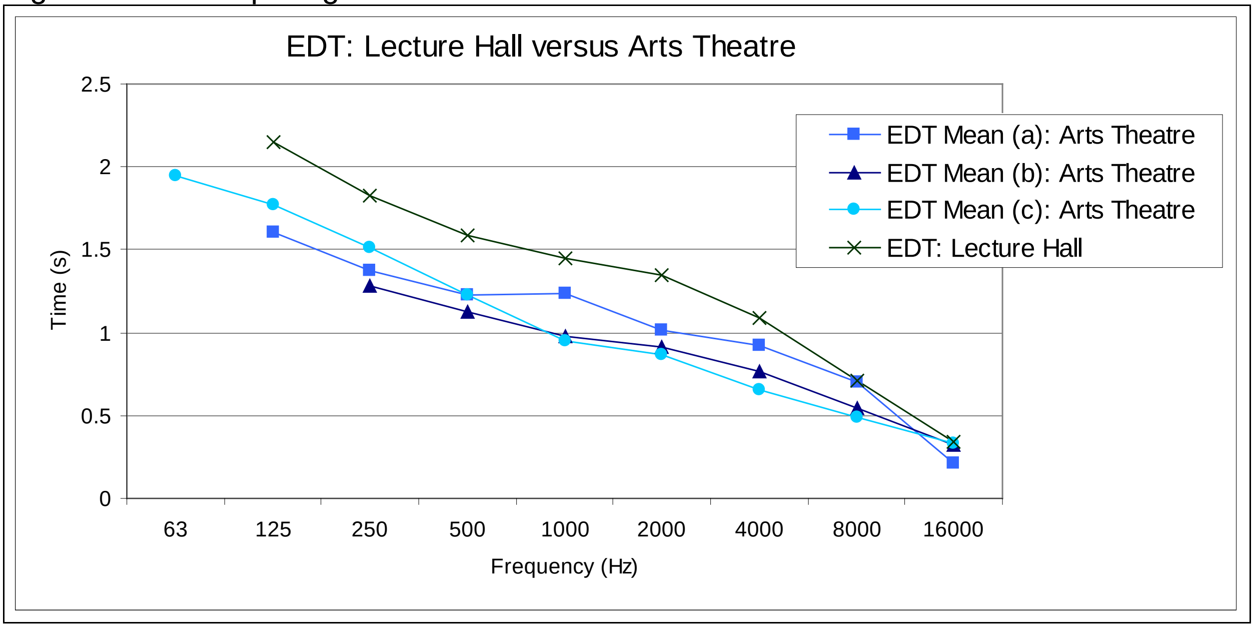
Figure 4.55: Comparison of mean EDT for sections a, b, and c for the arts theatre.



There was some difference in EDTs between the different sections, a, b and c respectively. For example, section a (stage only) exhibited longer EDTs than the other two sections, but only for frequencies above 500Hz. There were no excessive deviations for EDT in terms of the averaged comparisons of the three perceptual locations. The EDT of the three locations had a similar trend.

The arts theatre had a different reverberation and EDT to that of the lecture hall. Although the trend of the octave slope had similar parts to the lecture hall, the individual reverberation times for each octave were lower for the arts theatre. This was a surprising result. If the lecture hall and art theatre's EDT are compared on one set of axes, one can see this point clearly. The theatre was sectionalised thus there are three EDT plots for the arts theatre as shown in Figure 4.56.

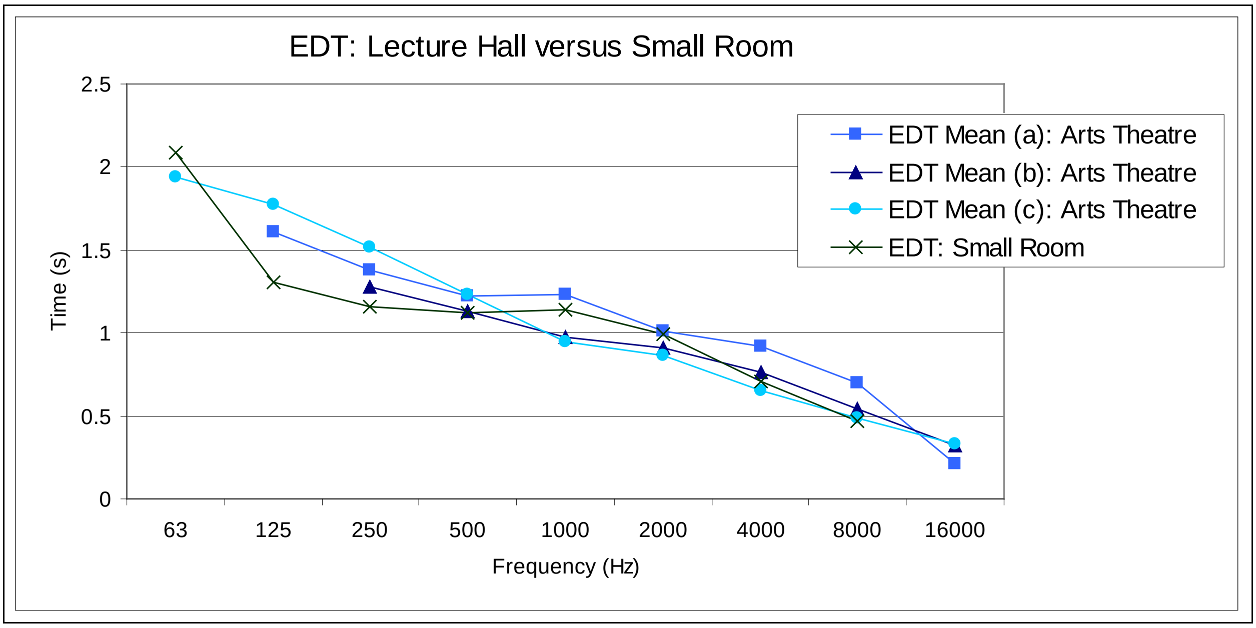
Figure 4.56: Comparing the EDT for the lecture hall and the arts theatre.



Comparing the arts theatre with the small room provided an unexpected result. The EDT for the small room fell in the same range to that of the arts theatre for octave bands including and above 500Hz. Thus, in terms of upper mid to high frequencies, there was similarity between the two venues even though the physical layout was different. Unfortunately, making assumptions based only on EDT would lead to premature conclusions and is highlighted by Bradley (2005) where he recommends that many more indexes be studied such as: sound level strength, *G*; early to late arriving sound ratio, C_{80}; lateral energy fraction, LF; and spatial effect measurements such as early and late inter-aural cross colouration (IACC).

The EDT for the small room plateaued during the lower octaves while the EDT for the arts theatre climbed steadily. The EDT for the small room had a sudden increase in EDT at the lowest octave and thus exhibited a sharp deviation in its trajectory. It was expected that the EDT of the arts theatre would exhibit a more "stable" response while the small room's EDT "bounced" around.

Figure 4.57: Comparing the EDT for the small room and the arts theatre.



4.5 Reverberation Time Analysis for Anechoic Environment (Reference)

A location that exhibited the closest approximation to an anechoic environment was selected. An open piece of farmland was used for this test. The idea of using an outdoor location for the anechoic test was found in Everest's (1981:200) *Master Handbook of Acoustics* where he attempted to mimic an echo free area and used an outdoor test site that overlooked a valley. While outdoor areas are not truly echo free, as there still exist reflections from the ground, however, in terms of this study, it provided a close enough approximation. A completely dead response was not required, as this study attempts to highlight the benefits of using a moderately dead room with compensation from artificial reverberation methods to provide a variable reverberant response instead. A pure echo free room is a rare facility and costly, while a moderately dead room is easier to obtain.

As this experiment forms the reference measurement to the next two chapters, it was covered in more detail.

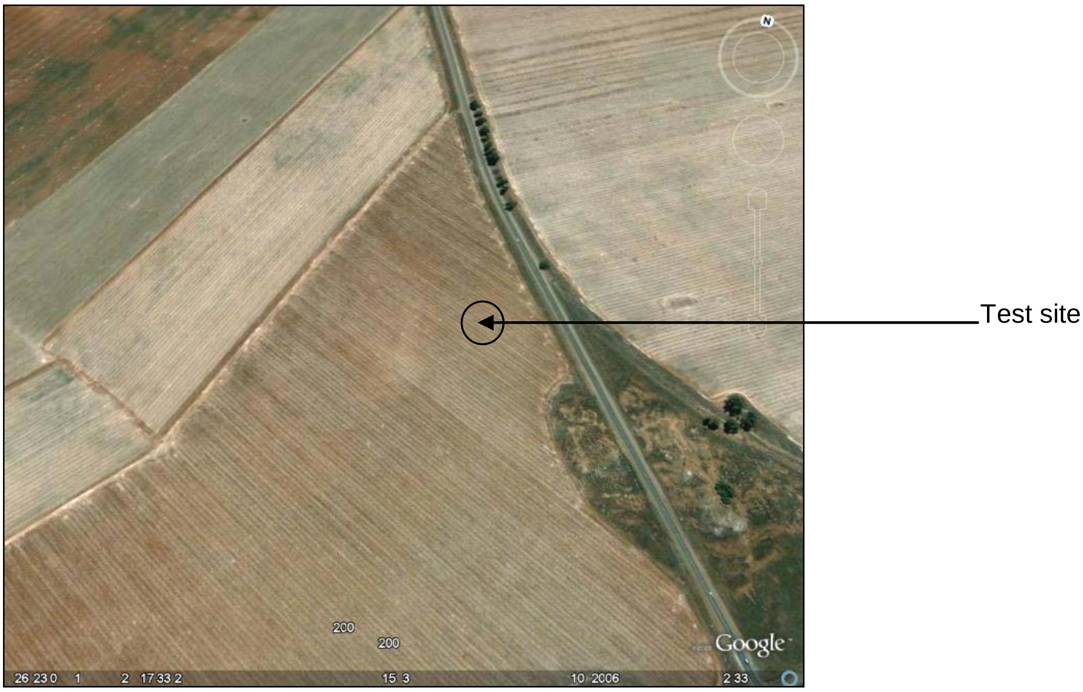
Table 4.17: Summary of test setup data for anechoic location – Farm in Heidelberg.

	Measurement 4: Anechoic
Date:	19 June 2009
Time:	15:05-16:25
Site:	Farm in Heidelberg (*Rooikraal Landgoed*)
Environment plan	See aerial and establishing photos
Volume of room	NA
Condition of nearby environment	Flat grassy plain. Corn fields nearby. No nearby structures or trees. Closest vertical object at least 2km away. Ground was dry with medium height wild grass.
Environmental conditions: Temperature, relative humidity	Partly cloudy day (Some signs of thunderstorm activity in the horizon). 19°C RH:49%
Degree of precision	Engineering method (ISO, 2008)
Measuring height above ground plane:	1,3m
Distance between source and SPL meter:	2m (1m and 1,5m taken as well)
States of occupancy	unoccupied
Background noise level	34-36dB
Peak SPL of impulse	103dB

4.5.1 Environmental Description and Layout

A maize farm was located south of Germiston 10km off the N3 highway in Heidelberg (±20km outside of Nigel). The maize crop had recently been cut and there was a large open flat space available. The following image was taken from *Google Earth*.

Figure 4.58: Google Earth picture of the anechoic location.



A spot was located that had no trees or other obstacles in the way of the horizontal plane. A road was nearby but was not busy, a car or two passed every 3-6 minutes. The next two figures show establishing shots of the chosen area.

Figure 4.59: Establishing shot of the chosen area.

Figure 4.60: Establishing shot of the chosen area - zoomed.

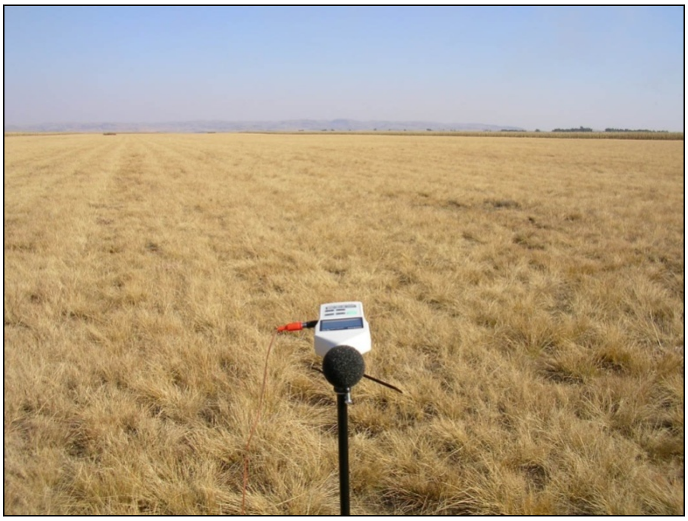
4.5.2 Measurements

Reverberation time tests as well as SPL tests were conducted on this site. The SPL tests were more detailed as they also were used for a test to prove that a reverberant room has a higher SPL level than an anechoic area while all other variables are kept the same for both locations. That test is presented in Appendix B. From this point on, only the RT tests are discussed.

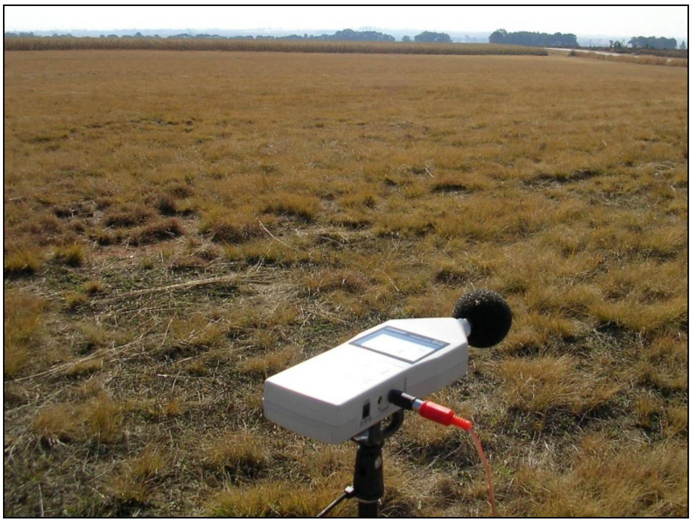
The surrounding environment from all angles of the SPL meter was included to show that there was very few reflecting surfaces near the device, excluding the ground, the SPL meter [and microphone] itself, some test gear and the stand.

Figure 4.61 a-c: Three-sixty directional view from the reference position of the SPL meter.

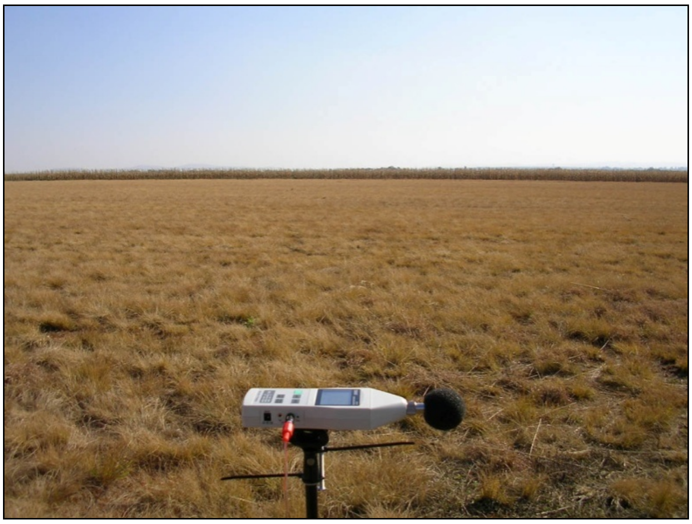
a)



b)



c)



The setup used for this experiment is shown in Figure 4.62. The figure shows a speaker on one of the upright stands. When the impulse tests were conducted, the speaker was removed and the balloon was positioned in its place. The distance between the two stands was increased as well. The speaker was used for an additional test, which can be found in Appendix B. A blanket is not shown, but was placed over the equipment to reduce reflections from the equipment on the ground.

Figure 4.62: The experimental setup for anechoic measurements.

The minimum source-microphone distance cannot be calculated as the volume is undetermined. Thus as per ISO standard 2m distance was used.

4.5.2.1 The Influence of the Environment on the Microphone

Wind blowing across a microphone can produce a lot of extraneous noise. As this was an outdoor measurement, an assessment of the environmental factors that could affect the microphone's performance was conducted. There was an intermittent breeze blowing at this test location. A test to determine the influence of the breeze on the SPL was conducted; the change in SPL was minimal. The porous sponge that accompanied the microphone was used. The humidity was about 50%, which has little influence on the performance of the microphone and recording equipment. However, sudden temperature changes may lead to condensation in the microphone. Singing into a microphone that has recently been stored in a cold location may generate condensation. To counteract this effect, the microphone temperature was monitored and only used when it was above room temperature. The impulse tests were recorded first and when the microphone warmed up, the voice recordings used in the subjective testing were recorded (Chapter 6).

There were signs of an upcoming thunderstorm on the horizon, which means that there was probably a reduction in pressure during the afternoon of this experiment. According to Brüel and Kjær (1984:25) variations in atmospheric pressure of ±10% will have a negligible influence (less than ±0,2 dB) on microphone sensitivity. Added to this, the pressure variation is too slow to have any effect on this experiment as the recording time per test was short.

4.5.3 Results

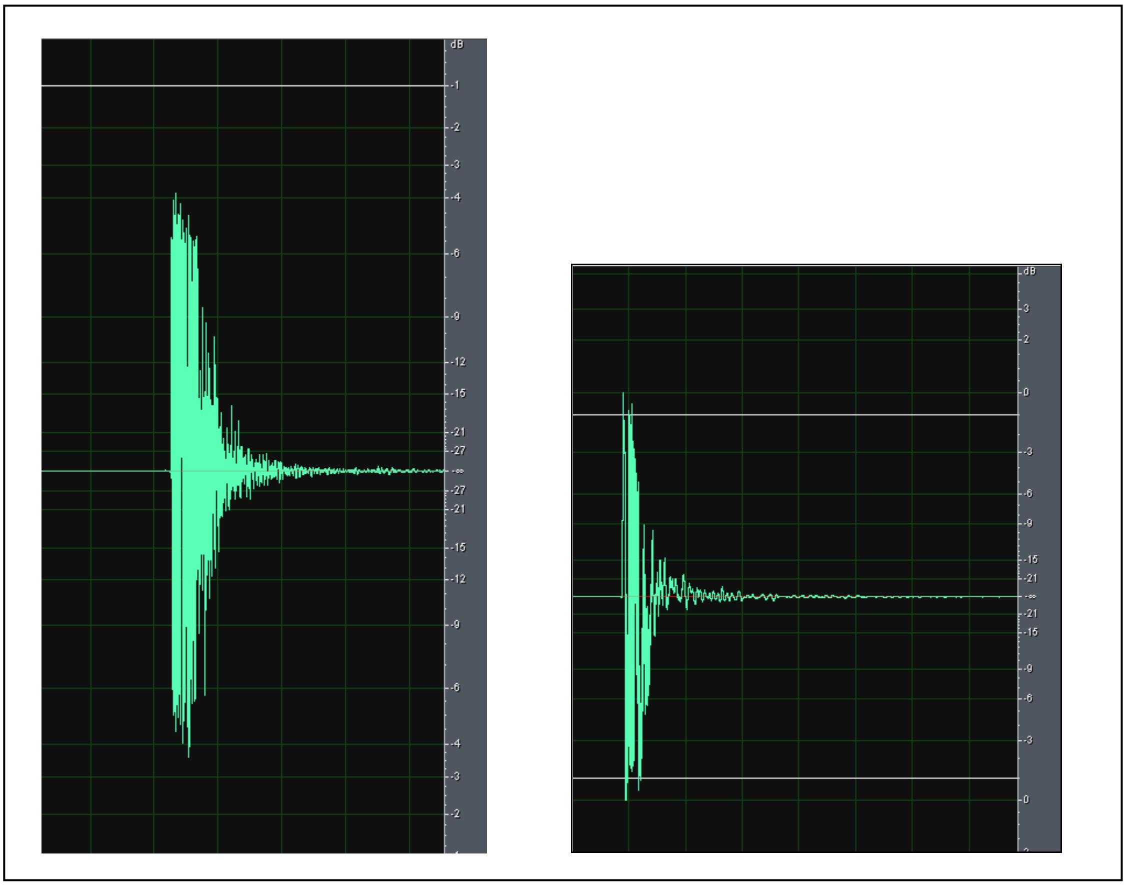
Owing to the environment being "*dead*[33]" the RT is very low and thus the time values that are shown are small. In such cases, it is important to establish whether the bandwidth product of the sound card can handle such a rapid decay slope (Lee, 2003). For example, Lee and Yu (2005) have studied the RT of passenger cars to ascertain their absorption ratio in order to measure noise control in the design of the audio system. By using the RT as obtained from inside of the vehicle, one can determine some of the audio systems parameters. The absorption of the materials in the vehicle is important in how much reverberation is measured. Cars generally have a remarkably short RT, which is so short that it is difficult to measure due to limitations in the traditional BPF bandwidth. Band-pass filters are used to target the individual frequency bands such as what was shown in the octave filtering plots earlier. At low frequencies, the bandwidth is narrow, while at high frequencies the bandwidth is much larger (Lee & Yu, 2005). The RT in a car is short, thus the product of bandwidth (B) multiplied with the RT is small. The delay response of the BPF imposes a limit on the short RT that is to be measured. Lee and Yu (2005) were able to measure the very short RTs using a wavelet filter bank in order to overcome the problem of the BPF response curve. They found that the RT time was generally 0,07s for passenger cars.

Keeping the above in mind, an assessment of the test gear was conducted. The sound card, microphone and recording software were found to handle low RTs[34]. Secondly, the tolerances are smaller when working with fast RTs. While an analysis has still been shown for this environment type, longer RT displays are better visually represented than shorter impulse decay slopes.

[33] The test area is considered "dead" in terms of this study but it is acknowledged that the environment chosen does not represent a pure echo free test area.

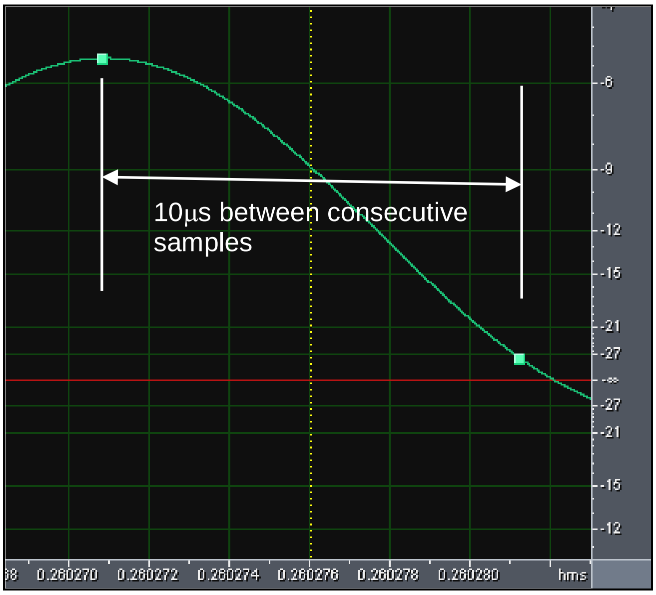
[34] The recording software can handle a sample rate of up to 196kHz. The software was set to 96kHz to match the sample rate of the sound card. The sound card could successfully sample a signal of 21μs. The bandwidth of the sound card is in excess of 25kHz. The sound impulse was of the order of 70 000μs. There was no problem found in terms of recording a low RT in the milli-second range.

Figure 4.63: Sample Graphical representation of impulse decay waveforms from Cool Edit Pro Software[35].



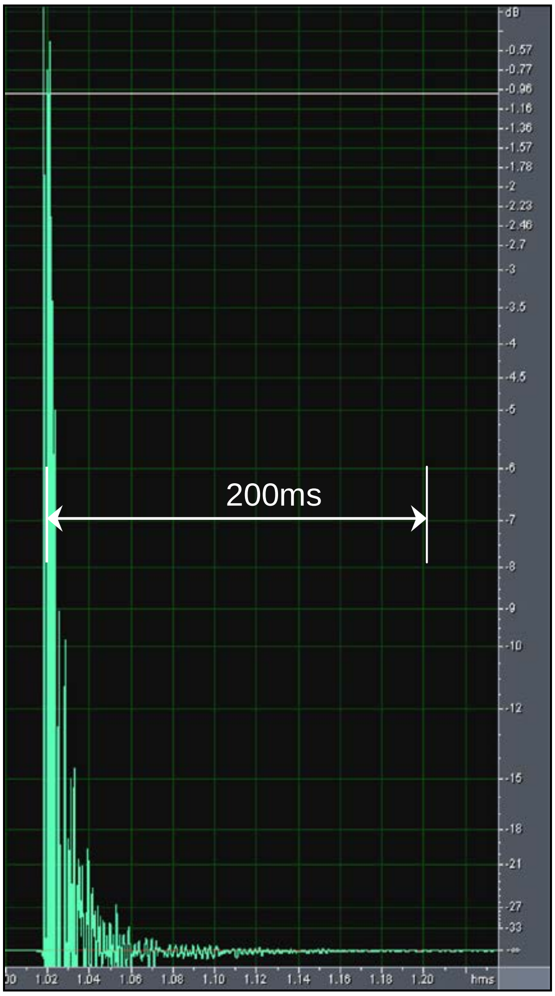
An example of the sample rate is shown in the next figure.

Figure 4.64: Sample rate view showing the time base and sample points in Cool Edit Pro.



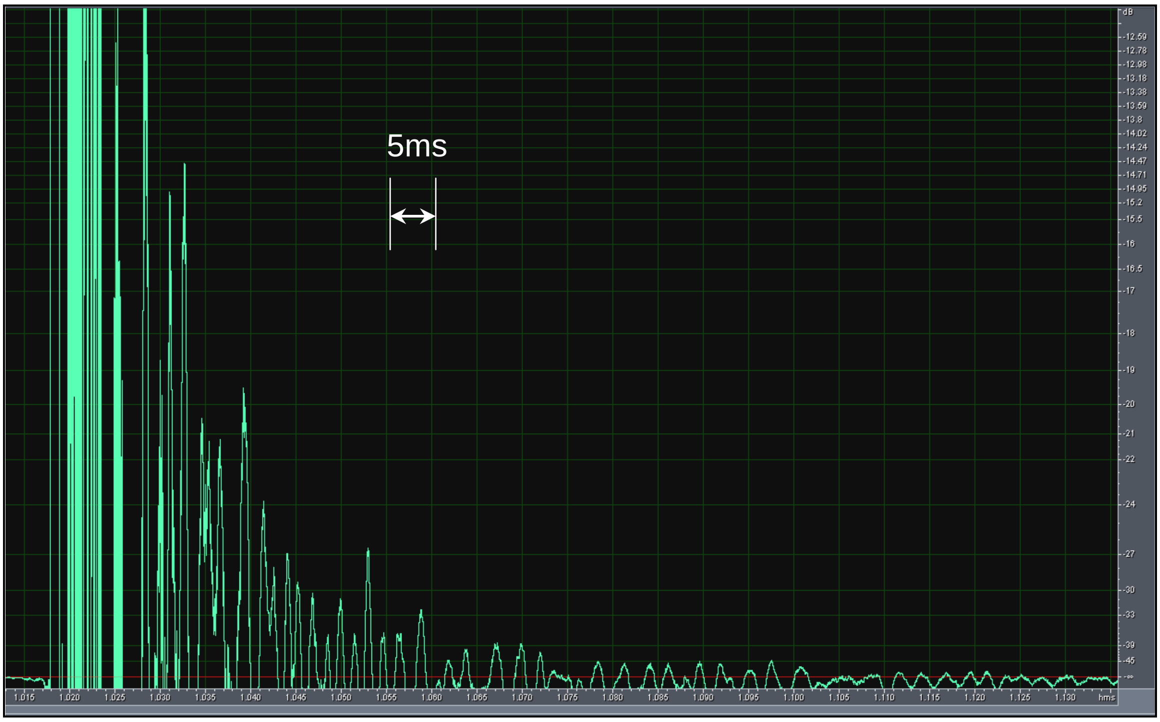
[35] The right hand picture in the figure exhibits a few overdriven peaks. Distortion of the impulses was checked and only impulses that showed no over-driving were used.

Figure 4.65: Decay slope (2m from source, angle of incidence zero degrees).



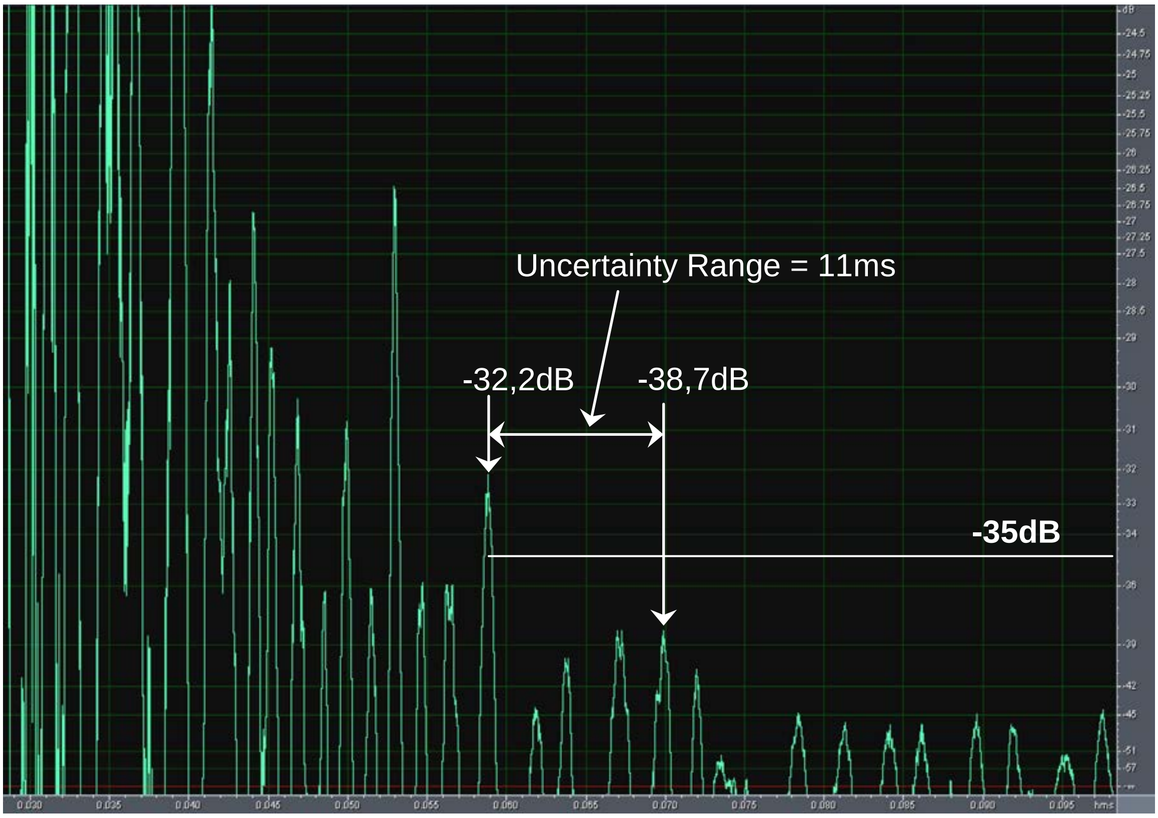
The next screen shot shows the rapid decay slope of one of the impulses. Note the time base is in increments of 5ms.

Figure 4.66: Screen shot showing the lower dB range: Zoomed view of tail.



Using Cool Edit Pro to determine the RT posed a problem in the lower dB cut-off point. The -35dB point was not easily obtained and a best-fit approximation was required. A graphical explanation as to why it is not easy to locate the lower point is shown next. This highlights one of the problems with using the direct analysis from level recorder method to obtain the RT.

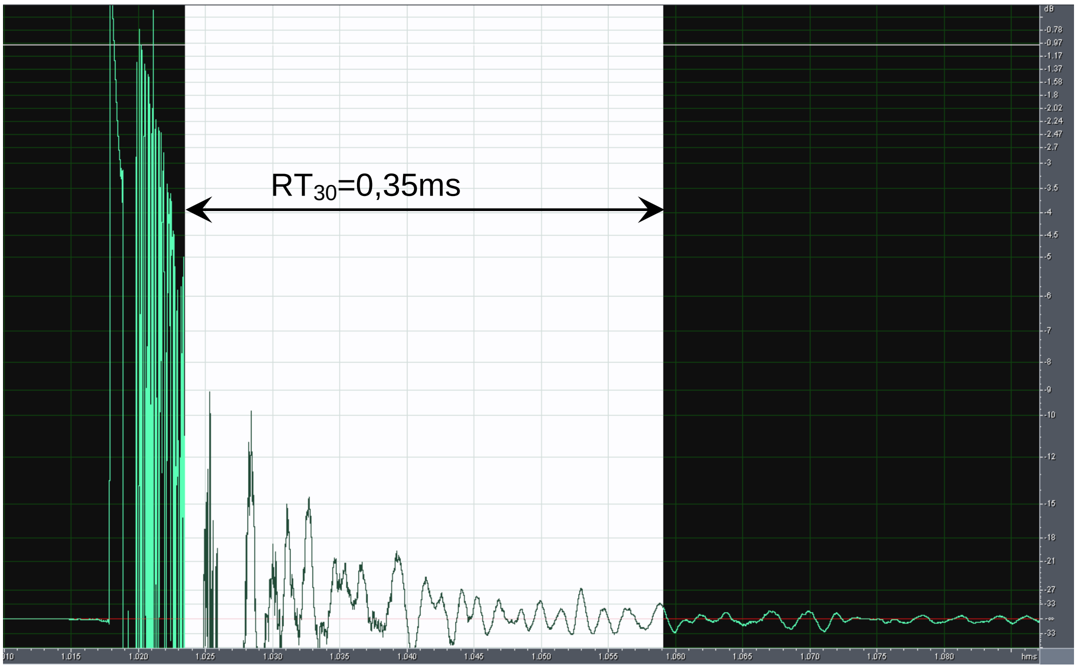
Figure 4.67: Uncertainty range at -35dB.



From the previous figure, a common problem is exhibited. The problem is that it is not always clear where exactly the lower cut-off point would be for RT measurement. It has been labelled as an uncertainty range as the -35dB (or -25dB where applicable) is not clearly shown. An approximation is required in order to determine the correct point. In such a situation, a best-fit line is drawn and the lower limit point selected from the line. However, this method is best achieved when the RT curve exhibits linearity such as when the time base is converted to a logarithmic scale. The Schroeder method would simplify this problem (still to come), however it is also beneficial to familiarise oneself with the actual raw wave file as was done in this example.

An attempt was made to determine the RT_{60} from the RT_{30} using a direct read off from the level recording as shown in the next figure.

Figure 4.68: Obtaining RT_{60} by direct read off from displayed waveform recording.



Measurement	RT_{30}	RT_{20}
Method: Direct read off. (Raw wave analysis)	70ms	69ms

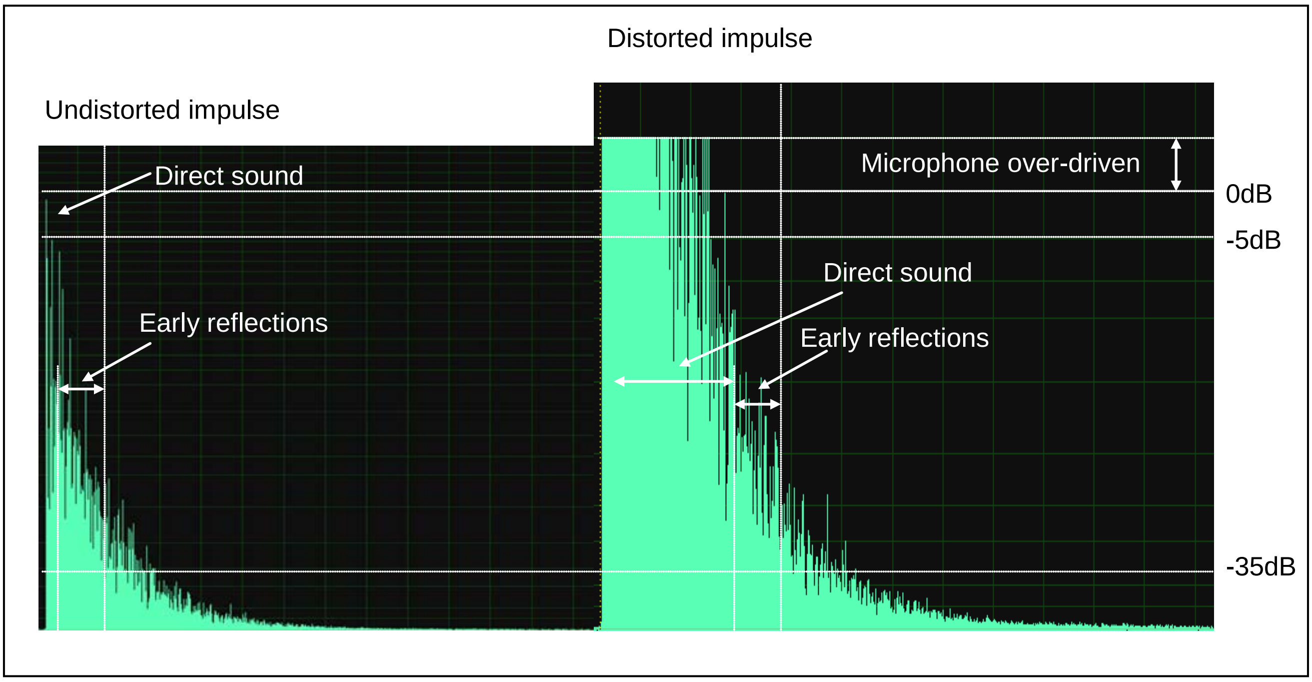
4.5.3.1 A Look Into the Octave Band Specific RTs

As in the previous three experiments, WinMLS software has been used for further analysis.

4.5.3.1.1 Notes Regarding Distortion

The SPL meter microphone had two distorted waveforms out of the five. Both were discarded. There was slight distortion at the very beginning of the impulse waveform on the third impulse test but when compared to the other undistorted samples there was no difference in the RT characteristic of the plot. The distortion consisted of a single wave peak clip. This situation is different to the previous case of distortion shown earlier in the chapter. The RT is only taken from a -5dB point so the single distorted wave crest had a negligible effect, while in the earlier distortion analysis, many wave peaks were distorted. In a normal impulse, the sound decays exponentially and analysis takes place from -5dB onwards. However, if the signal was very distorted, then where it would normally start decaying, it is still distorted and thus the software “sees” it as still part of the direct sound as there is no decay shown. Thus, the software cannot process that segment of sound and the RT reading is incorrect especially from an octave perspective. The following figure illustrates this problem.

Figure 4.69: Illustration of problem with distorted impulse sound.



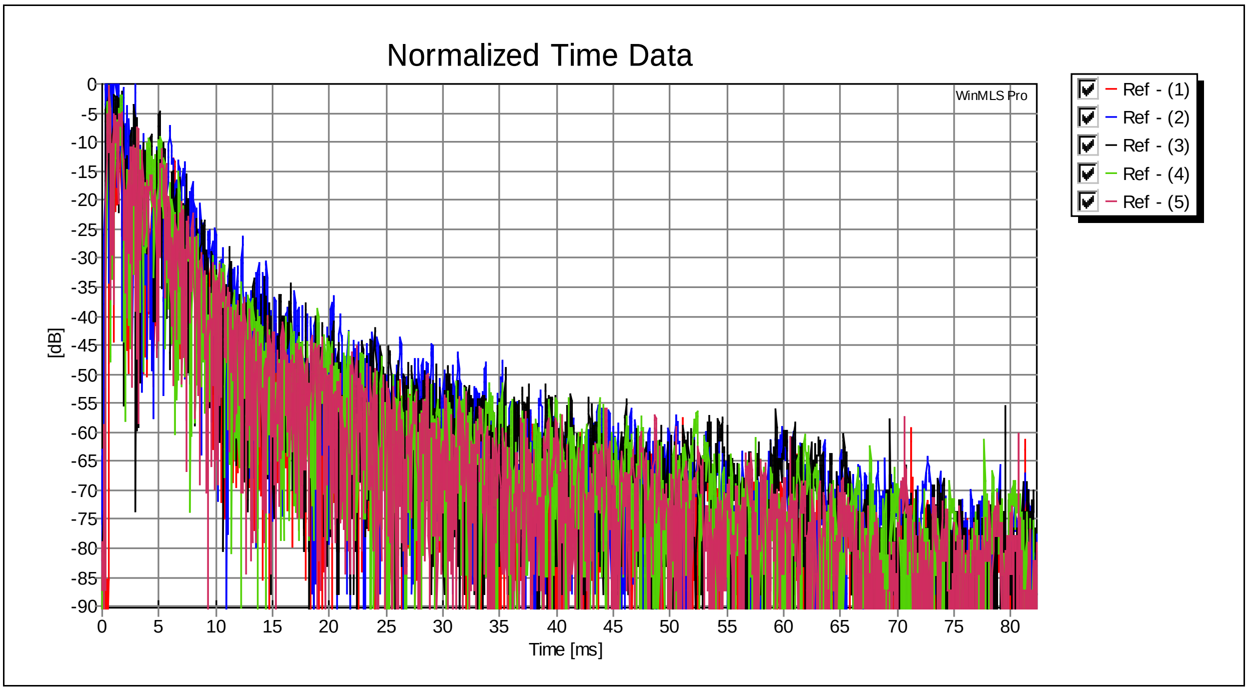
From the above figure, a few inferences can be made. It is assumed that both impulses were taken in the same place with the same environmental layout with the only difference being that the second one's microphone was overdriven. The points are as follows:

- The direct sound on the undistorted impulse is much shorter in period than the distorted one.
- The early reflections of both impulses are not in the same part of their respective waveforms. The distorted wave's early reflections would be counted as later reflections on the undistorted waveform.
- The relative position of the -5dB to -35dB RT sample size are constituted of different waveform parts as in the previous point.
- The loudness or sound strength differs between the two, although this may not necessarily change the RT characteristic.
- The noise floor may be part of the distorted waveform's lower reflected wave component if the SNR is low. Thus, noise components may be "seen" as intelligible signal.

After discarding any distorted samples, the effective sample measurements were undertaken. A few source-microphone distances were used and compared to check for microphone over-loading. The effective impulses used for this test area were all at the ISO specified distance of 2m. All measurements were checked for distortion. The impulse tests are all labelled "ref" as they were the reference tests. As there was more than one impulse test conducted, they are labelled accordingly to the impulse test number. Thus, all in all five reference measurements were used for this RT analysis. They were quite close in their repeatability.

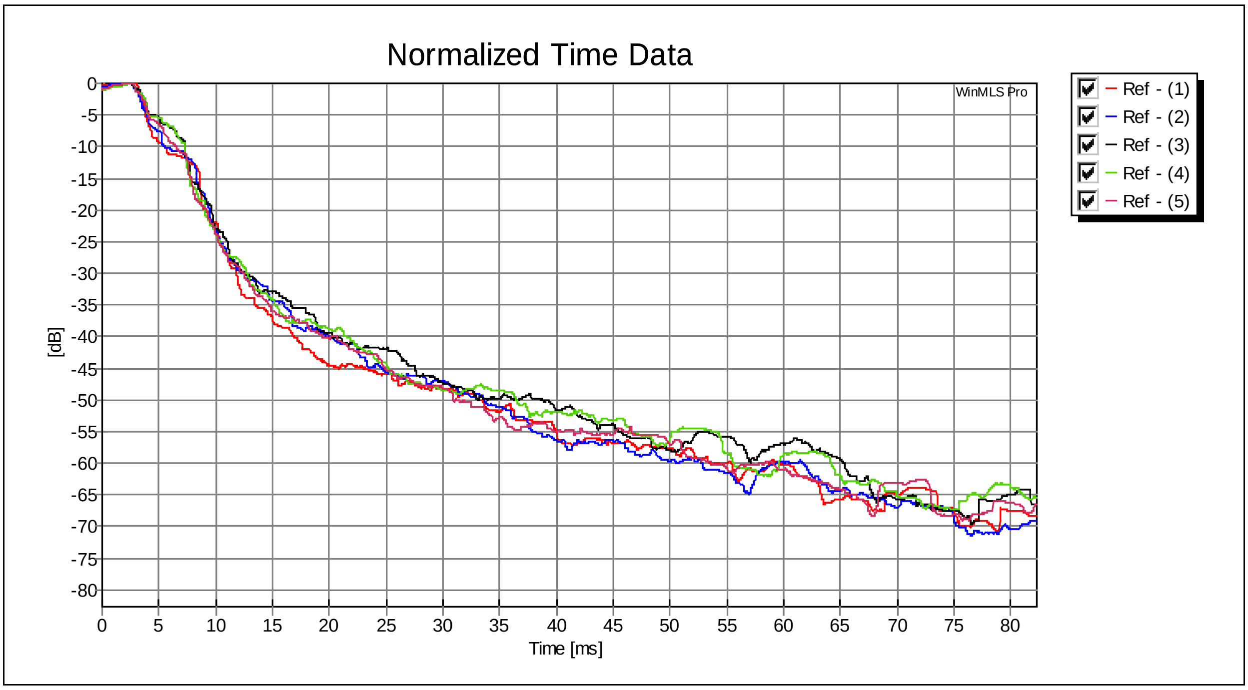
Plotting the impulse response for all five measurements results in a messy display as shown in the next figure.

Figure 4.70: Impulse response for the five test samples without any processing applied.



The plot looks too busy when all five impulses are shown with their full peaks and troughs. If the "energy in bins" is applied to the decay slopes, a cleaner view is obtained. The five impulses track each other quite closely for the full time period as shown in the next figure.

Figure 4.71: Impulse response for the reference test samples with "energy in bins" applied.



From the preceding figure, it can be shown that by 75ms, all the waveforms have reduced by 65dB and that from -50dB onwards the responses become more noisy. It may seem like a gradual decay, but it is in fact rapid as the time base is in milliseconds.

The unfiltered Schroeder curve is shown next. There is some deviation in the low dB range, in particular the green plot Ref-(4) drifts away from the group. There are three labels A, B and C that have been used to illustrate the effect of calculating the RT_{60} from three different decay ranges shown in the next table.

Figure 4.72: Schroeder curves of the five impulse tests for anechoic area.

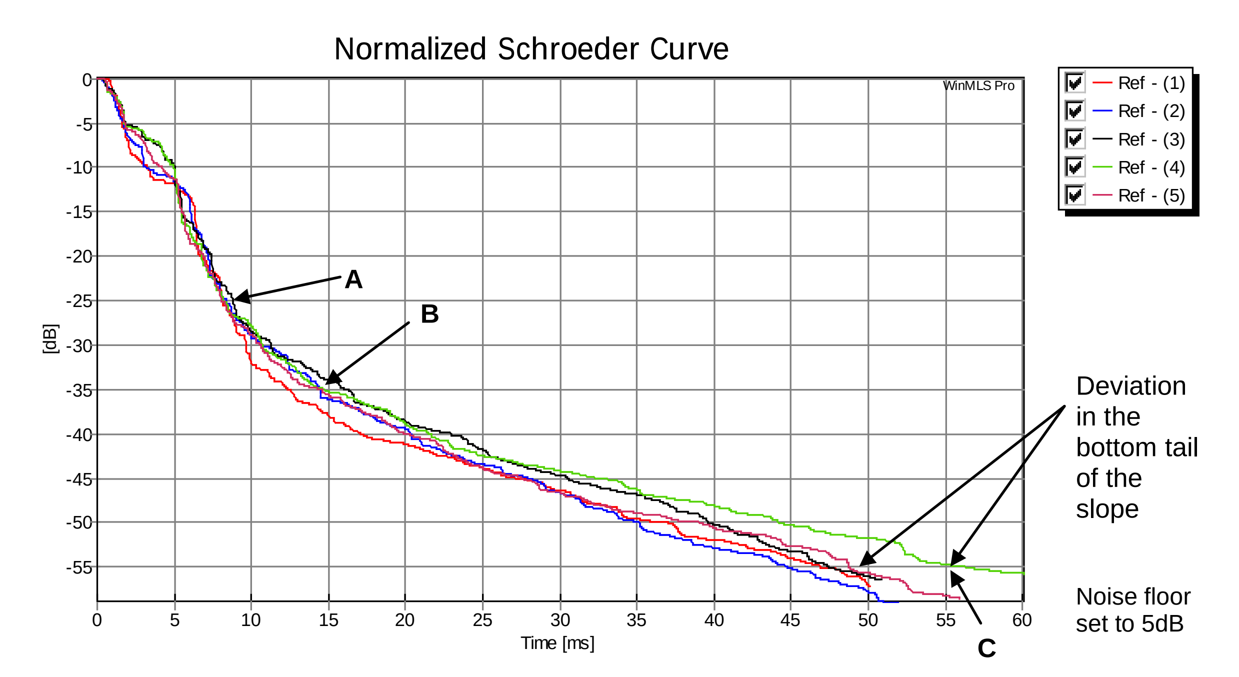
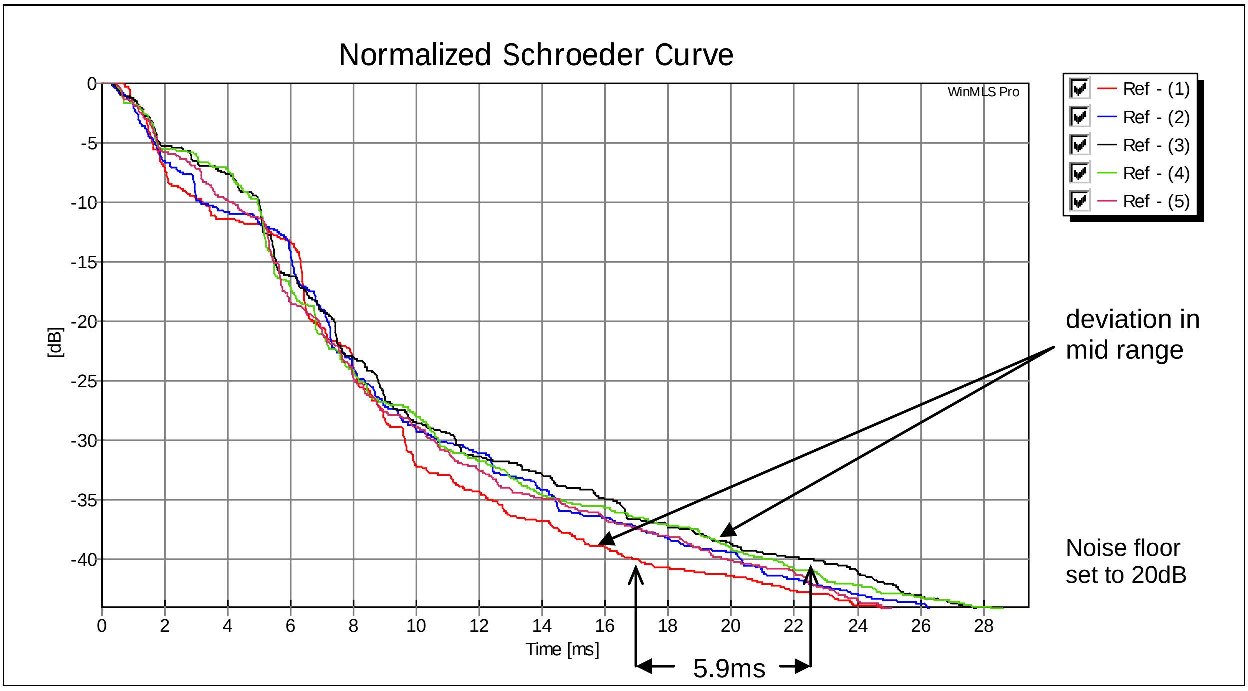


Table 4.18: The effect of applying different Schroeder decay ranges for computing RT_{60}.

Label	Decay range	Effective RT_{60}
A	-5 to -25dB	24ms
B	-5 to -35dB	28ms
C	-5 to -55dB	66ms

Two points can be highlighted here; firstly there are differences in RT from different decay ranges used, secondly, there is some difference in the RT between the five impulses, and in particular, the Ref-(4) impulse. If the RT was calculated from the full 65dB decay then there would be a difference in Ref-(4) compared to the group. In most practical conditions, it is not possible to have a signal gain of 65dB above the ambient SPL and thus most RT measures are conducted over a shorter dB range. The minimum RT at -55dB was found to be plot Ref-(2) which exhibited an RT of 44ms while the maximum RT was 55ms for Ref-(4). Converting that to RT_{60}, the former would be 54ms while the latter would measure 66ms, which translates to a 20% difference between them. Owing to the high SNR obtained in these measures, the noise floor can be increased and the next figure shows the response of an exaggerated noise floor level of 20dB.

Figure 4.73: Schroeder curves of the five impulse tests for anechoic area with 20dB noise floor applied.

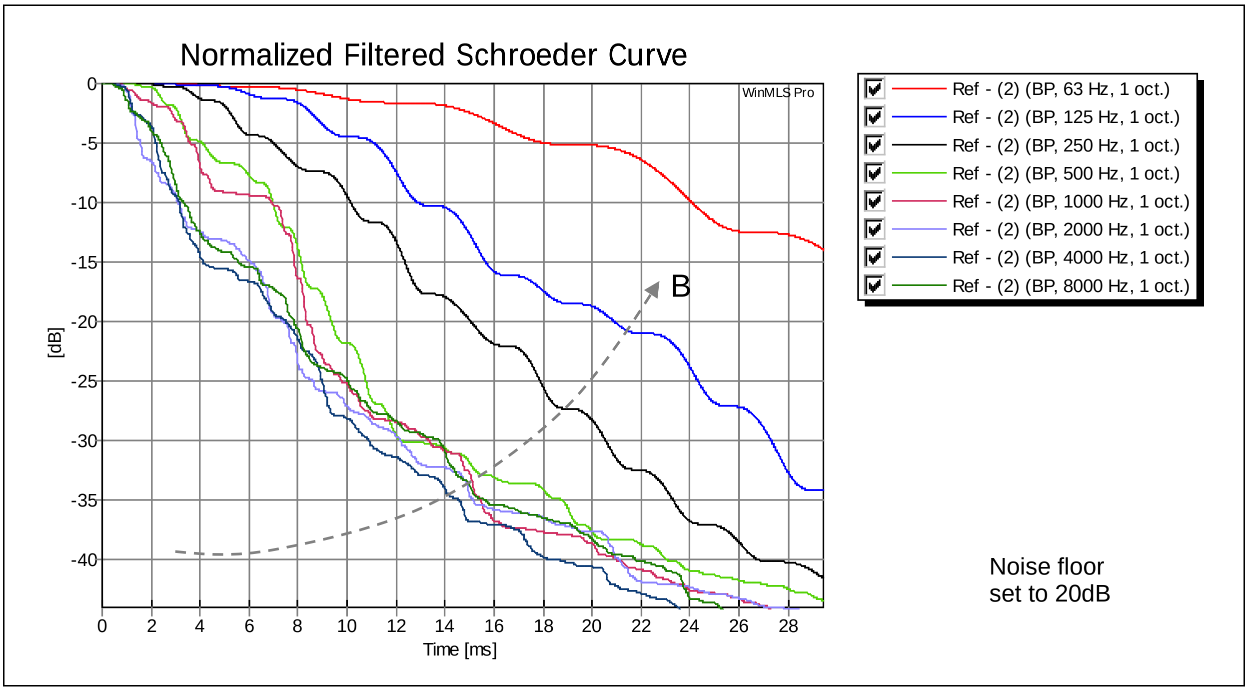


From Figure 4.73, it can be seen that critical analysis of the entire Schroeder curve is necessary as the RT value changes at least 10% for different positions. For example, the red plot [Ref-(1)] is now the deviator from the group in the range of -30dB to -45dB. Many RT measures do not allow for such a large dB range of analysis and thus inaccuracies may creep in merely by the choice of bottom dB limit for the RT value. Taking this example further, if the RT is calculated from the red plot, the equivalent RT_{60} was found to be 29,1ms while the green/black plot exhibited an RT_{60} of 37,7ms. Thus, there is a significant difference. In this testing case study, all the plots are of the same location and same test and thus will be averaged so the deviation will not be as significant.

It is known that high frequency components decay faster than lower frequencies. This is shown in the next octave specific Schroder curve. The analysis is only of a single impulse as the plot would be too busy if all were shown. The second impulse was used[36].

[36] The impulses for the octave specific Schroeder plots all gave a similar trend. There were some differences in the gradients of the individual octave frequency bands but nothing excessive, and thus no substantial differences were noted.

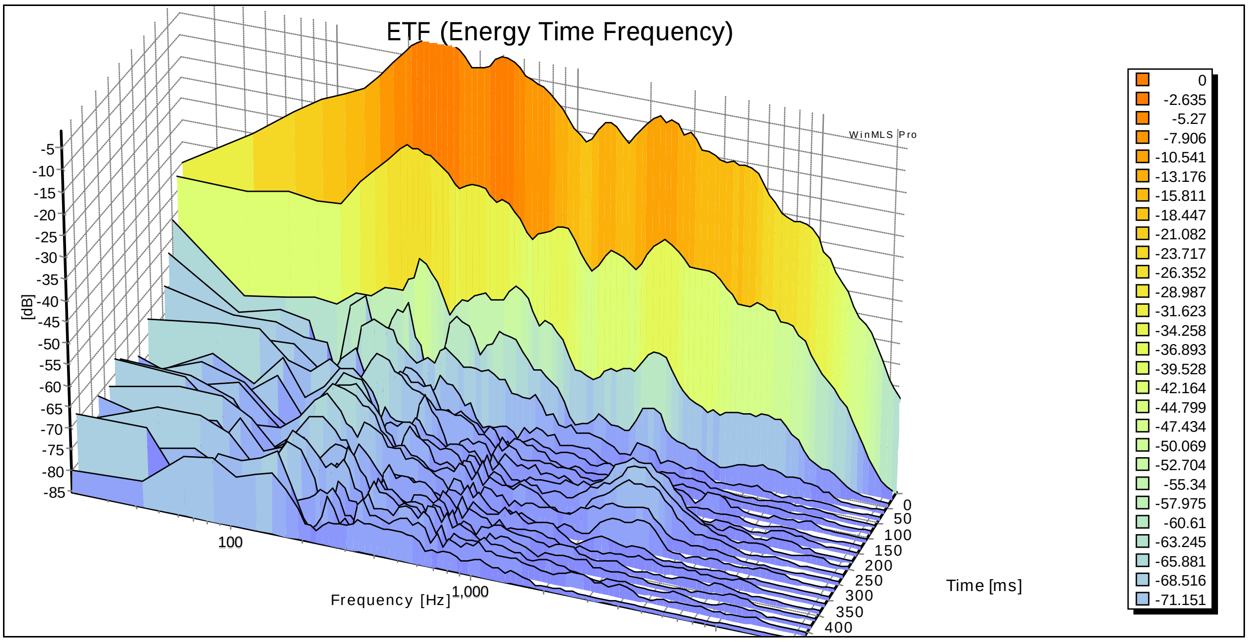
Figure 4.74: Octave specific Schroeder curve for impulse Ref-(2).



The dotted curved arrow labelled B shows the trend in RT increase with decrease in octave bands as would be expected. There is an inverse relationship between RT and frequency for this anechoic area measurement, which was what was expected.

The waterfall plot shows how the majority of sound energy originates from the low frequency bands.

Figure 4.75: Waterfall plot of a single impulse test.

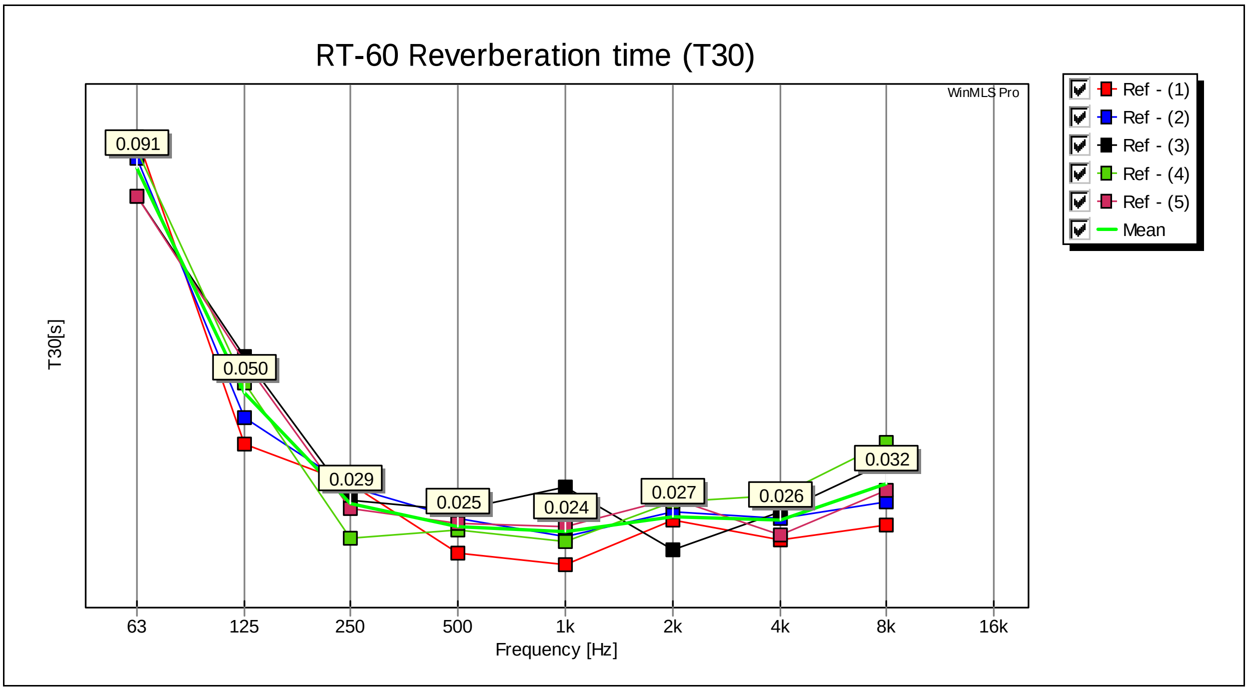


4.5.3.1.2 Analysing the Frequency Specific RT Characteristic

To obtain the mean RT and EDT for the anechoic area, the measures were plotted and a mean value was taken. The plots are shown in the next few figures. The numbered boxes that appear above each octave point represent the RT_{60} and are labelled for the mean value (green plot).

As there were five impulse tests conducted, there are five corresponding RT_{60} plots.

Figure 4.76: The RT_{60} for anechoic test with the mean shown.



The highest RT is found at 63Hz, which is true for all impulses in the above figure. Interestingly the highest octave-band does not exhibit the lowest RT.

Figure 4.77: The EDT for anechoic test with the mean shown.

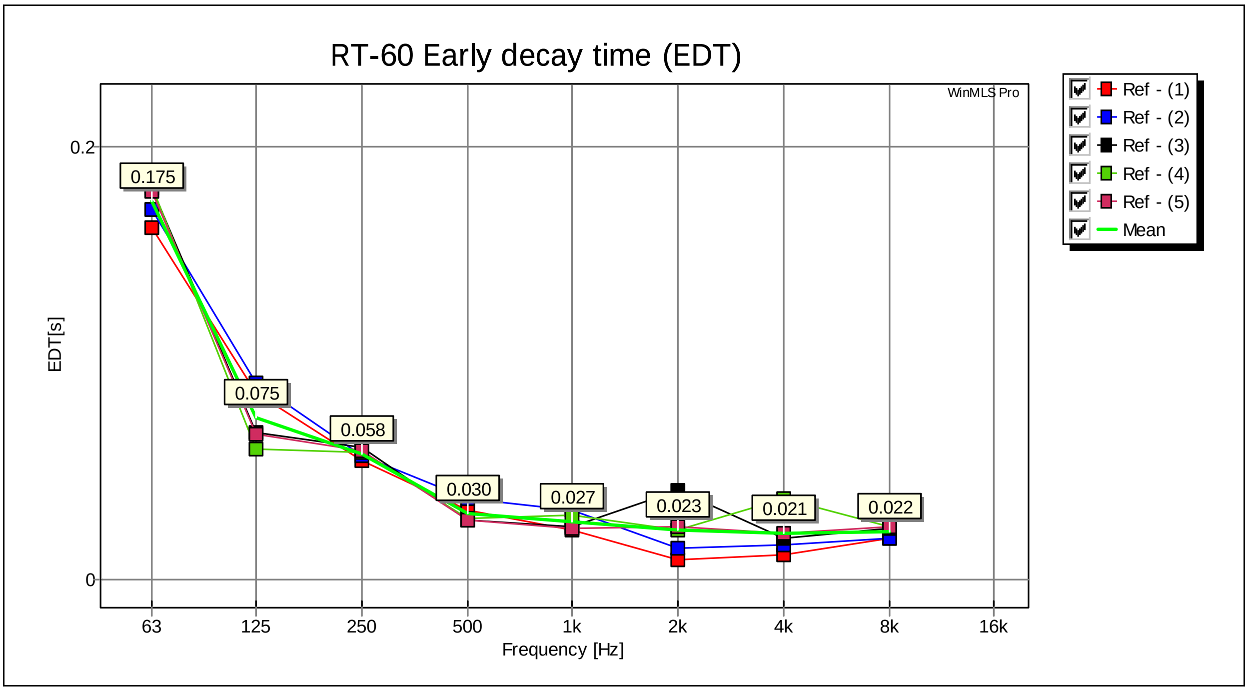


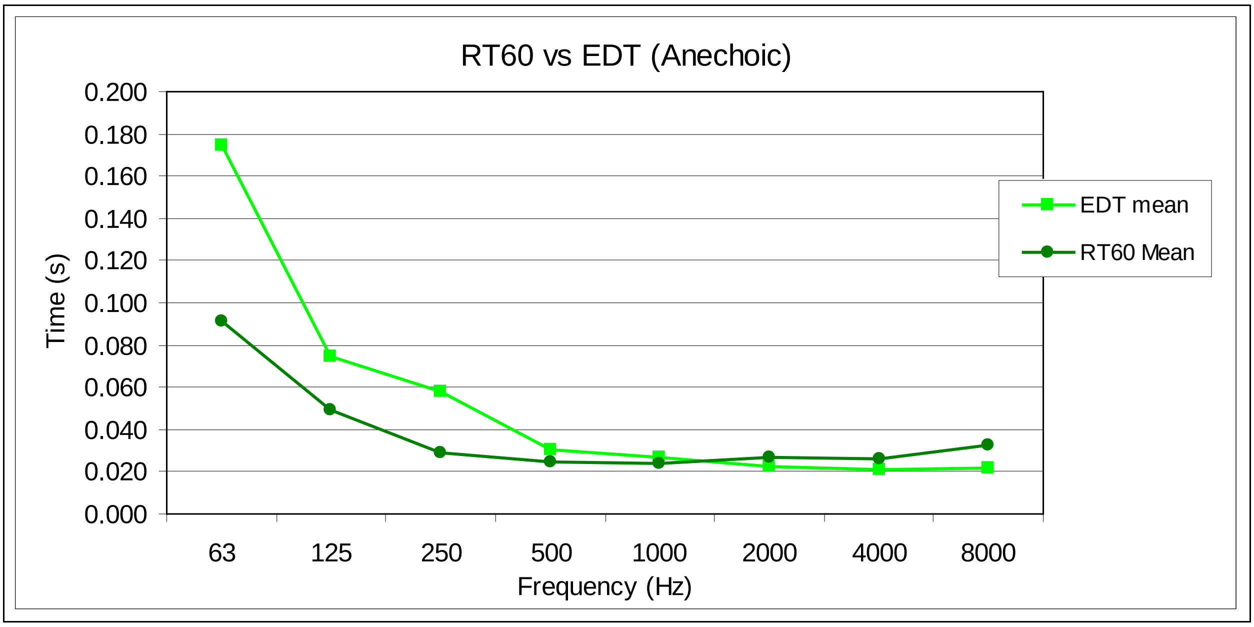
Table 4.19: Numerical summary of data for anechoic area.

RT60						
x-axis F [Hz]	Ref – (1) (s)	Ref – (2) (s)	Ref – (3) (s)	Ref – (4) (s)	Ref – (5) (s)	Mean (s)
63	0.097	0.093	0.086	0.094	0.086	0.091
125	0.040	0.045	0.056	0.051	0.055	0.050
250	0.033	0.032	0.030	0.023	0.028	0.029
500	0.020	0.026	0.028	0.024	0.025	0.025
1000	0.018	0.023	0.032	0.022	0.025	0.024
2000	0.026	0.027	0.020	0.029	0.030	0.027
4000	0.022	0.026	0.027	0.030	0.023	0.026
8000	0.025	0.029	0.037	0.040	0.031	0.032

EDT						
x-axis F [Hz]	Ref - (1) (s)	Ref - (2) (s)	Ref - (3) (s)	Ref - (4) (s)	Ref - (5) (s)	Mean (s)
63	0.162	0.171	0.181	0.181	0.179	0.175
125	0.086	0.091	0.068	0.060	0.067	0.075
250	0.055	0.058	0.061	0.059	0.059	0.058
500	0.032	0.038	0.027	0.028	0.027	0.030
1000	0.023	0.032	0.025	0.030	0.023	0.027
2000	0.010	0.015	0.041	0.023	0.024	0.023
4000	0.011	0.016	0.019	0.038	0.022	0.021
8000	0.019	0.019	0.023	0.024	0.024	0.022

Comparison of the RT_{60} and the EDT for the anechoic area is shown in the next graph.

Figure 4.78: Comparison of RT_{60} and EDT for anechoic area.



4.5.4 Discussion

The results obtained from the *Rooikraal Landgoed* farm can be used as the “dead” environmental response as the RT and EDT are considerably lower than the small room, the lecture hall and the arts theatre. A comparison has not been undertaken in this section as the anechoic results were compared thoroughly in the following chapter. The results for the echoic locations were not corrected for audience occupancy and thus should be considered as such.

4.6 Conclusion

This chapter provided a practical implementation of the ISO 3382 standard that was reviewed in Chapter 3. Four different acoustic environments were used to obtain their reverberation characteristics. Many graphs were presented in this section. This is not uncommon in acoustics where acoustical indexes are best presented in a graphical method with RT being high on that list.

This chapter sets the trend for the next chapter as it follows a similar sequence to this one. The results obtained in this chapter are used as the basis for the next chapter, which attempts to manipulate the anechoic impulse response to mimic that of each echoic response conducted on an individual basis.

Chapter

5

5 ARTIFICIAL REVERBERATION APPLICATION: OBJECTIVE TEST

5.1 Introduction

This chapter uses the results obtained from the analysis section as a base for an objective test in applying AR. It attempts to prove that a software-based method of reverberation control can be applied successfully in manipulating the reverberation characteristic of the anechoic environment to match that of the echoic room (from Chapter 4). This section forms part of the applications part of this dissertation and is partnered to the next chapter (subjective testing).

The objective test consists of matching the octave specific RT for the anechoic environment to each of the three echoic rooms, starting with the small room, then followed by the lecture hall and ending with the arts theatre. The goal is to match the RT of the artificially reverberated impulse sounds to be within a 10% deviation of their naturally reverberated counterparts between the frequency ranges of 125Hz to 4kHz. This frequency range was used as ISO 3382-2:2008 standard for the measurement of RT specifies this range for adequate RT measurements. As the chapter heading states, this is an application of method rather than an analysis of the method itself. The focus is on the results and whether the stated goals could be reached.

The term anechoic is used for the reference impulse that lacks any noticeable reverberation and thus in terms of this study is considered a “dead” sound. After the anechoic or dry impulse was artificially reverberated to match each of the livelier impulses from the echoic rooms, they were processed using octave specific BPFs and compared with each other.

5.1.1 Procedure Overview: Objective test

The effective impulse test response obtained in the dead area was manipulated using Cool Edit Pro acoustic software. The response was adjusted numerous times with various levels of signal processing applied in order to obtain a best fit to that of each room. This process was sequential in nature. There were three experiments in total, which are in the same order as the experiments in the previous chapter; that is, the small room was first to have its RT and EDT measured in Chapter 4 and is still the first room to be compared to the anechoic in this chapter. The only difference is that the anechoic is now taken as the reference and thus there are only three sequential experiments in this chapter, which are: small room and anechoic (*Rooikraal Landgoed* farm), lecture hall and anechoic, arts theatre and anechoic.

5.2 Experiment 1: Anechoic Adjusted to Exhibit the Small Room's Reverberation Time Response

The mean RT_{60} and EDT that were obtained from the previous section have been used as the goal for the anechoic RT response. As this is the first application test, more detail was included in this first test starting with a summary of the before condition. The anechoic RT_{60} is shown in the next figure followed by the small room's RT_{60} plot.

Figure 5.1: The mean octave band RT_{60} plot for the anechoic area[37].

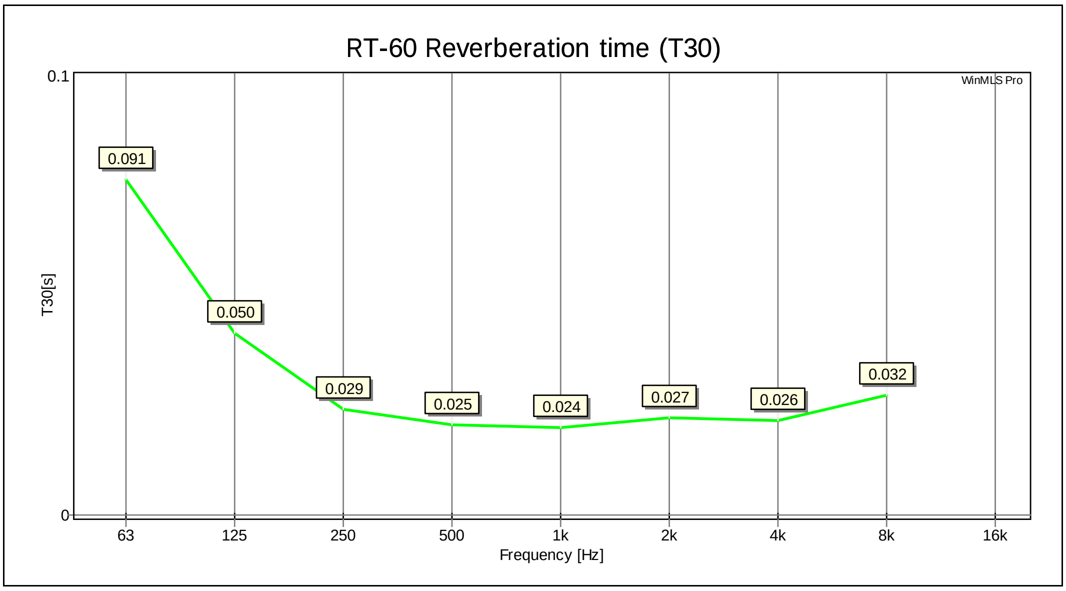
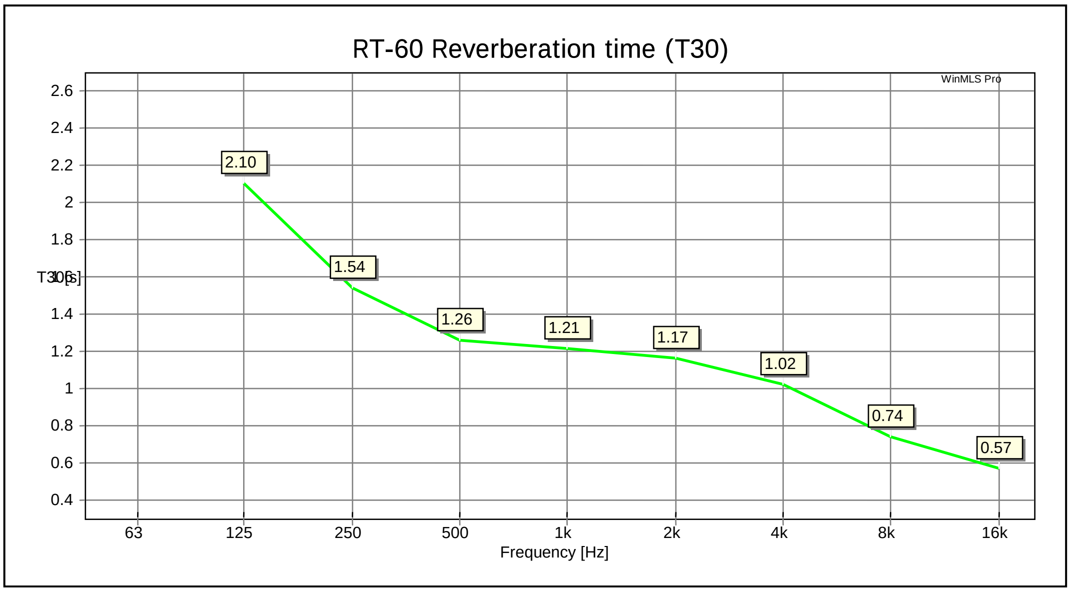


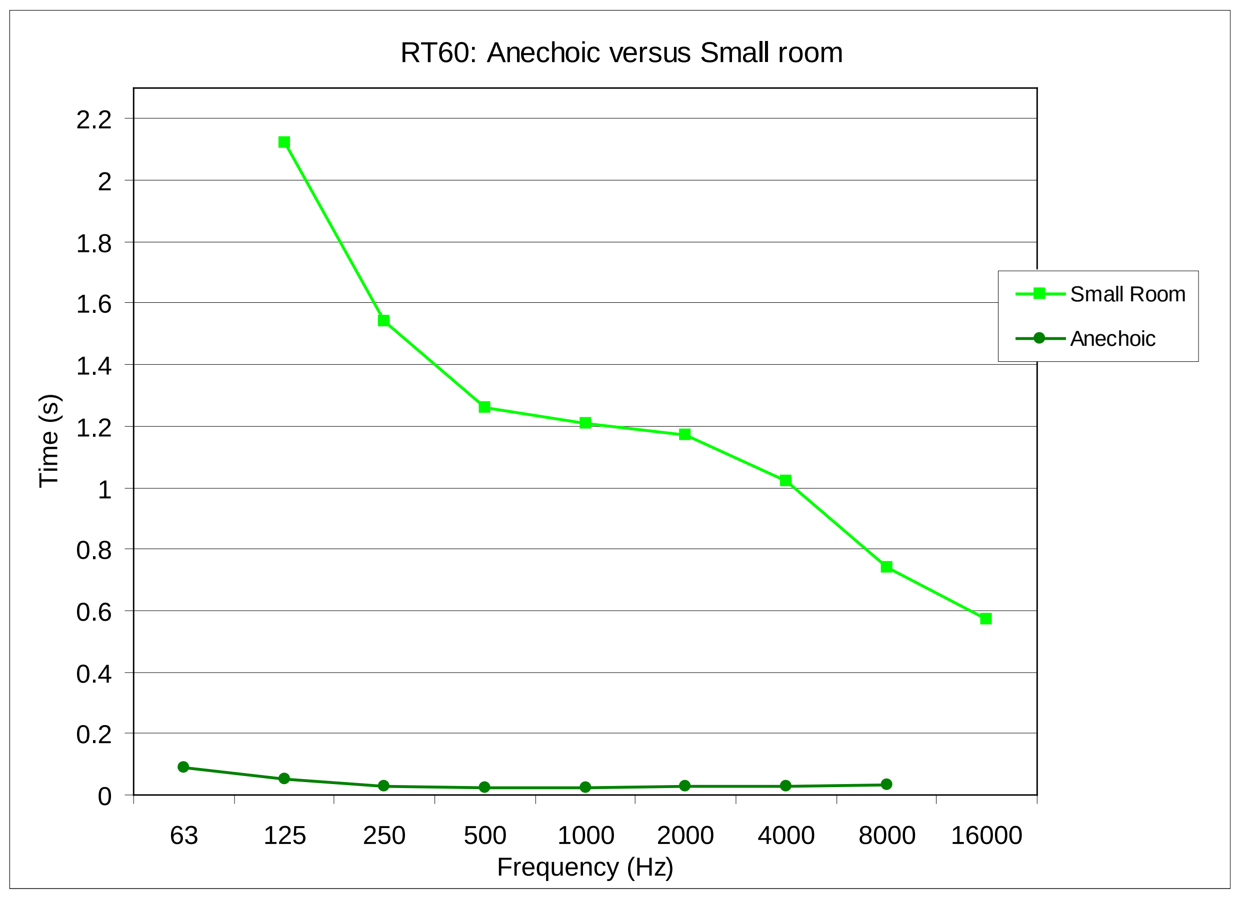
Figure 5.2: The mean octave band RT_{60} plot for the small room.



different RTs

[37] The actual raw wave file that was processed to match the small room's RT was waveform Ref-(5) as the mean shown in the graph is not a wave file but rather a mathematical plot. The closest curve to the mean was chosen and is discussed further under the "Method" section.

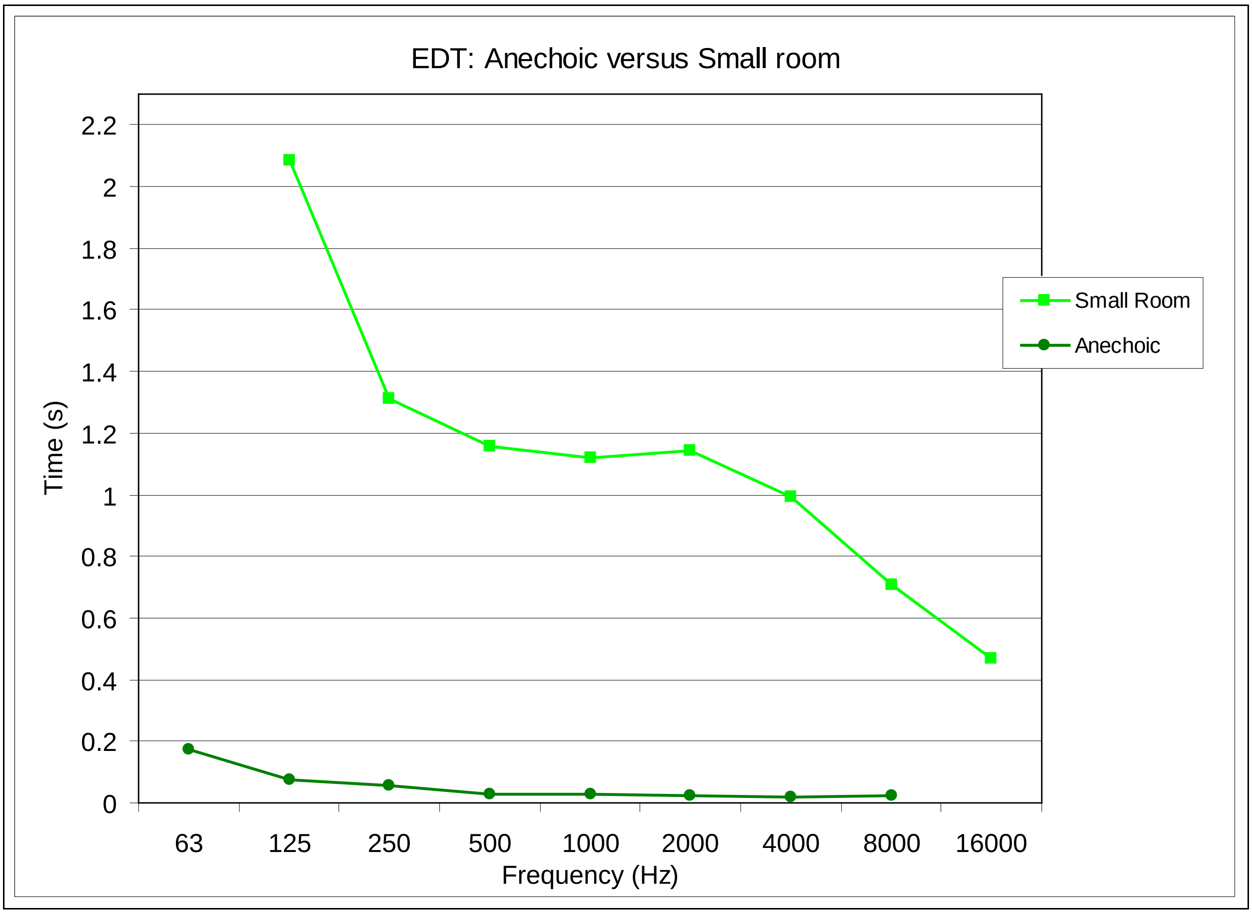
Figure 5.3: Relative comparison of anechoic and small room's RT_{60} shown on same axes.



Comparing the two RT curves on the same axes provides a clear relative reference as to how "dead" the reference measurement was in terms of its RT. The anechoic response is almost flat when compared to the small room, yet the anechoic octave specific plot is not actually flat. However, when plotted on a larger time base, it flattens out considerably.

The goal was to make the anechoic curve as close as possible to the small room's RT curve. Thus, two similar octave-specific RTs signalled a positive result. The next figure shows the EDTs compared on a single set of axes.

Figure 5.4: Relative comparison of anechoic and small room's EDT shown on same axes.



The difference in RT between the anechoic and the small room is large. Table 5.1 shows how many times larger the RT is for each octave band. By browsing through the "Times larger" column, it can be concluded that this application presents an extreme example of AR. Some of the octaves show an excess of 50x more reverberation than the anechoic impulse.

Table 5.1: Comparing the anechoic and the small room's RT_{60} and EDT.

x-axis	Anechoic RT60 Mean (s)	Mean RT60 - Small Room (s)	Times larger		x-axis	Anechoic EDT Mean (s)	Mean EDT - Small Room (s)	Times larger
125	0.05	2.1	**42.0**		125	0.075	2.08	**27.7**
250	0.029	1.54	**53.1**		250	0.058	1.31	**22.6**
500	0.025	1.26	**50.4**		500	0.03	1.16	**38.7**
1000	0.024	1.21	**50.4**		1000	0.027	1.12	**41.5**
2000	0.027	1.17	**43.3**		2000	0.023	1.14	**49.6**
4000	0.026	1.02	**39.2**		4000	0.021	0.99	**47.1**
8000	0.032	0.74	**23.1**		8000	0.022	0.71	**32.3**

5.2.1 Method

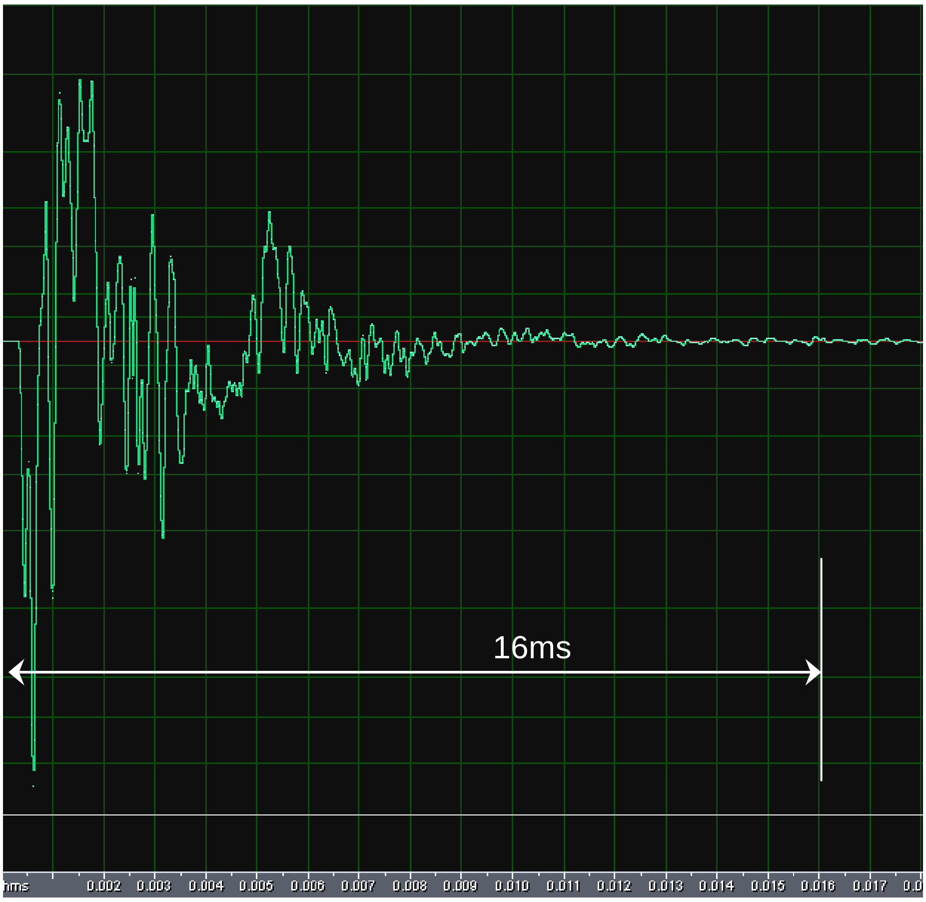
The anechoic wave file was imported back into Cool Edit Pro. Importing the wave file does not have any effect on the contents of the file. All signal processing was conducted using Cool Edit Pro software. The mean anechoic curve is not a wave file but a mathematical mean of the total of all the reference impulses; thus, one of the reference impulses needed to be used as the wave file that was to be processed and used for the application of AR. Referring back to the anechoic RT analysis in Chapter 4, it was found that the reference plot number five [Ref–(5)] was the closest plot to the mean plot for the anechoic area. Thus, the Ref-(5) impulse sound wave was used as the reference wave file. The following table shows how similar the two RTs were between Ref-(5) and the mean of the anechoic area. The same applied for the EDT.

Table 5.2: Comparing the reference RT_{60} and EDT [Ref-(5)] with the mean.

x-axis F [Hz]	RT60 Ref - (5) (s)	RT60 Mean (s)		x-axis F [Hz]	EDT Ref - (5) (s)	EDT Mean (s)
63	0.086	0.091		63	0.179	0.175
125	0.055	0.050		125	0.067	0.075
250	0.028	0.029		250	0.059	0.058
500	0.025	0.025		500	0.027	0.030
1000	0.025	0.024		1000	0.023	0.027
2000	0.030	0.027		2000	0.024	0.023
4000	0.023	0.026		4000	0.022	0.021
8000	0.031	0.032		8000	0.024	0.022

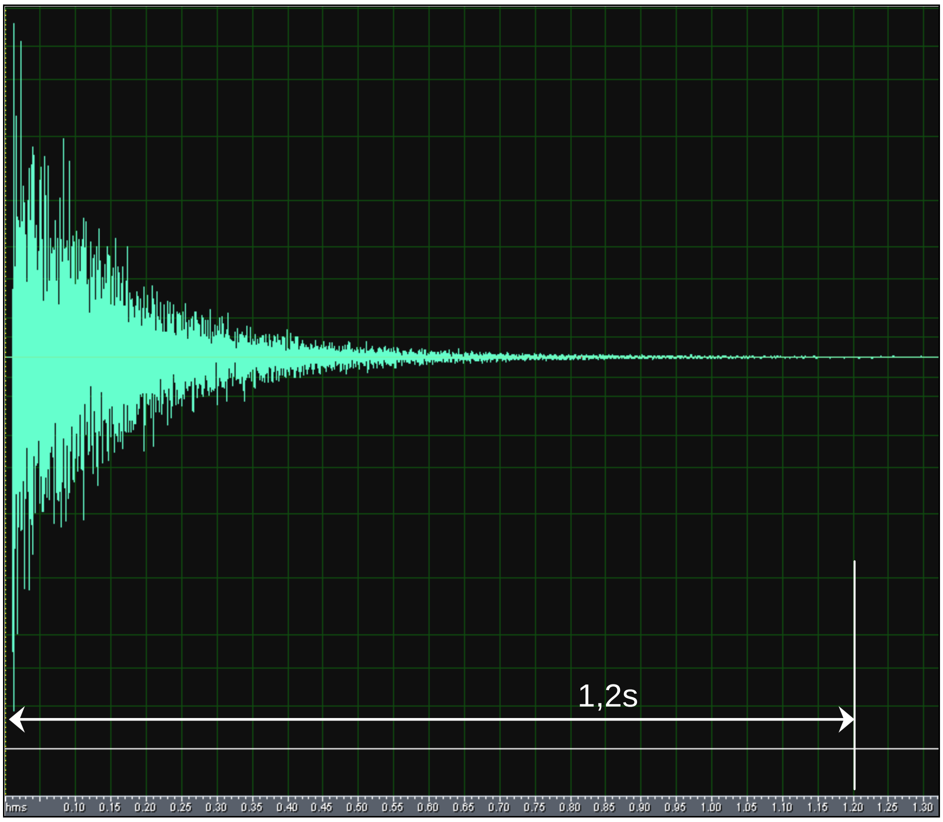
The anechoic impulse was short in time duration and was best viewed in the millisecond time base. A screen shot of the wave file that is to be processed follows next. The waveform shows SPL in the vertical plane and time on the horizontal axis.

Figure 5.5: Screen shot of the raw reference impulse before any signal processing applied.



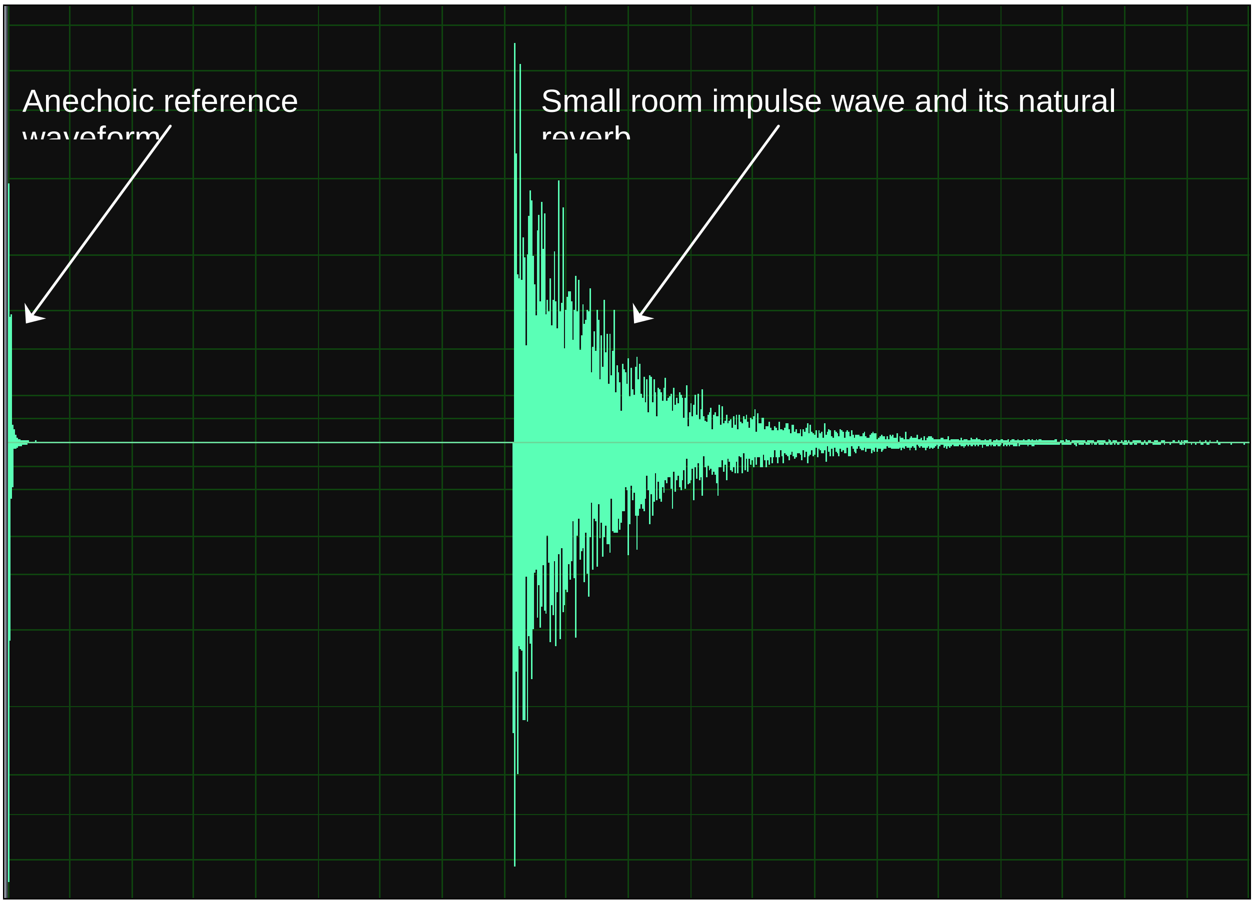
Comparing the anechoic impulse to the small room's impulse wave follows next.

Figure 5.6: Example of a small room's impulse decay response.



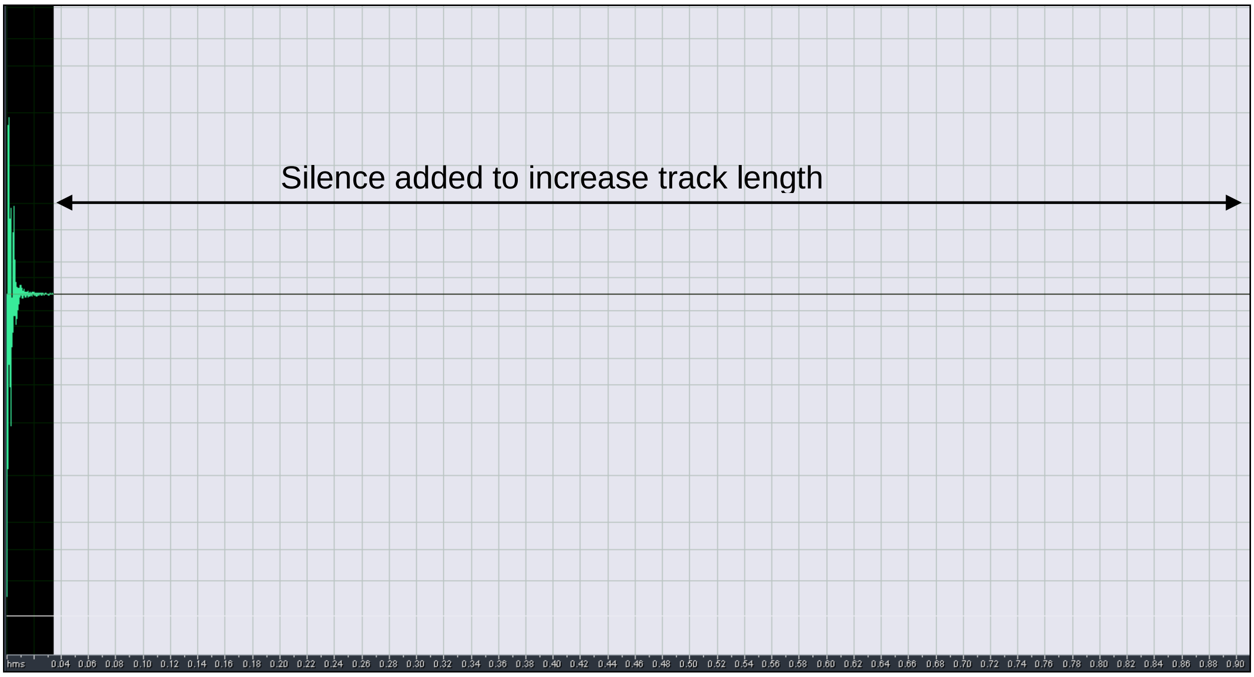
Viewing the two waveforms on the same time base highlights the clear RT difference between the two locations.

Figure 5.7: Comparing the anechoic reference wave with an example waveform from the small room sample.



Preparing the anechoic reference wave file starts with an adjustment to its track time. The wave file was extended in its time duration by adding additional time space to the plot. The reason for this was that the small room's maximum RT was 2,12s while the anechoic was 99ms. Thus, to make the anechoic similar to the small room, the time period available in the wave file needed to be at least 2,12s even though there is no useful impulse sound after 100ms. This was accomplished by the addition of a silence period after the anechoic impulse had decayed fully (Fig. 5.8).

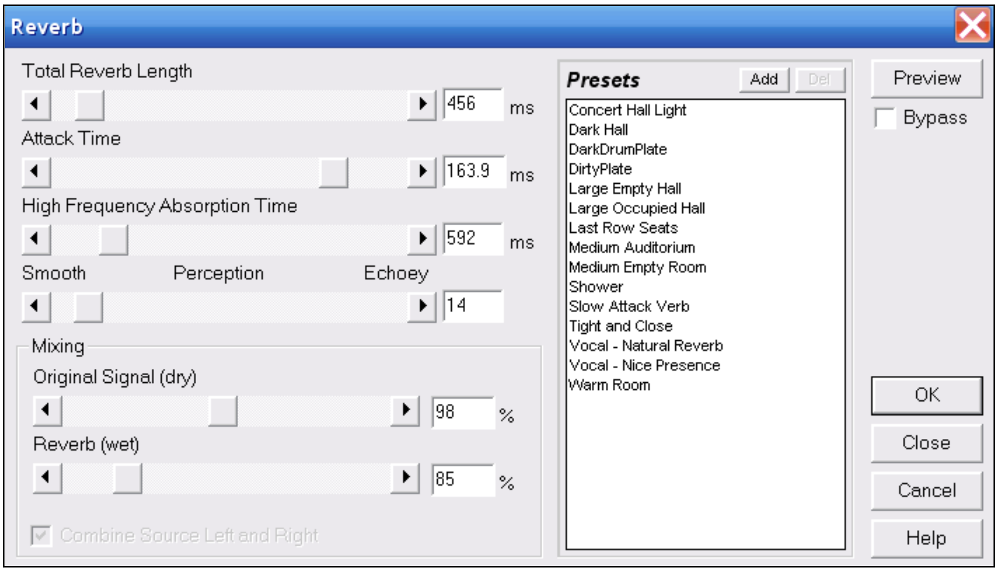
Figure 5.8: Anechoic waveform with a silence component added to extend the track time.



The anechoic impulse is now ready for signal processing. A method for adjusting the RT of one wave to match the RT of another wave using acoustic recording software has not been found in the literature. Thus, this started as a trial and error attempt. At least 125 different methods, with some having more than six signal processing steps were attempted before tangible and repeatable results were obtained. The following procedure is not a fixed method. There were many variations to this experimental method especially since this is an exaggerated example of applying AR. The method used for this first room differs from that of the other two, mainly as part of the learning process to allow alternative signal processing sequences to be carried out. It is important to note that this is not the way acoustic editing software would normally be used, as listening tests are fundamental to the process of audio editing. Secondly, the edited files would not be assessed using MLS software. A question may arise as to why this method was undertaken in the first place. The reason is that listening tests are subjective and may have a large area of uncertainty when interpreting the results as it is based on personal preference. Using MLS software provides a mathematical evaluation of the edited files instead.

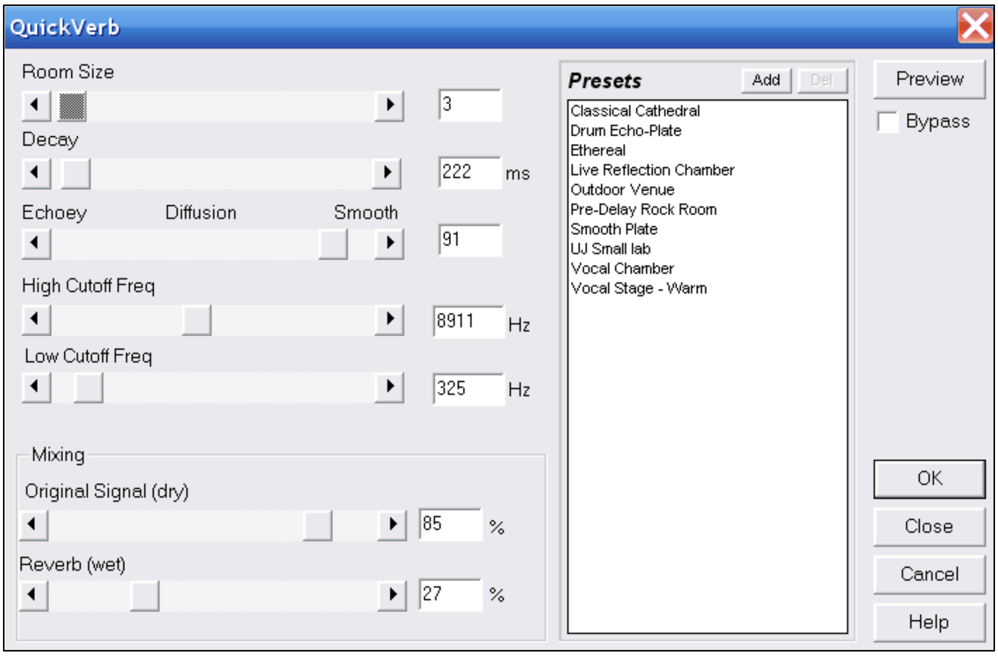
Cool Edit Pro has three reverberation effects to choose from. They differ in their content. The screen shots of the three options are shown in the next figures.

Figure 5.9: Option 1: Standard Reverb.



The standard reverb field allows for RT length adjustments (Fig. 5.9). There are also adjustments for attack and absorption time amongst other settings. A column of presets is available for all the reverberation effects. The software allows the user to choose pre-defined sound types. For example, one of the presets is "shower". Once chosen, the software attempts to edit the current track to sound like a shower's reverberation style. The user can also define his/her own presets. The second option follows in Figure 5.10 and is termed "Quickverb".

Figure 5.10: Option 2: *QuickVerb*.

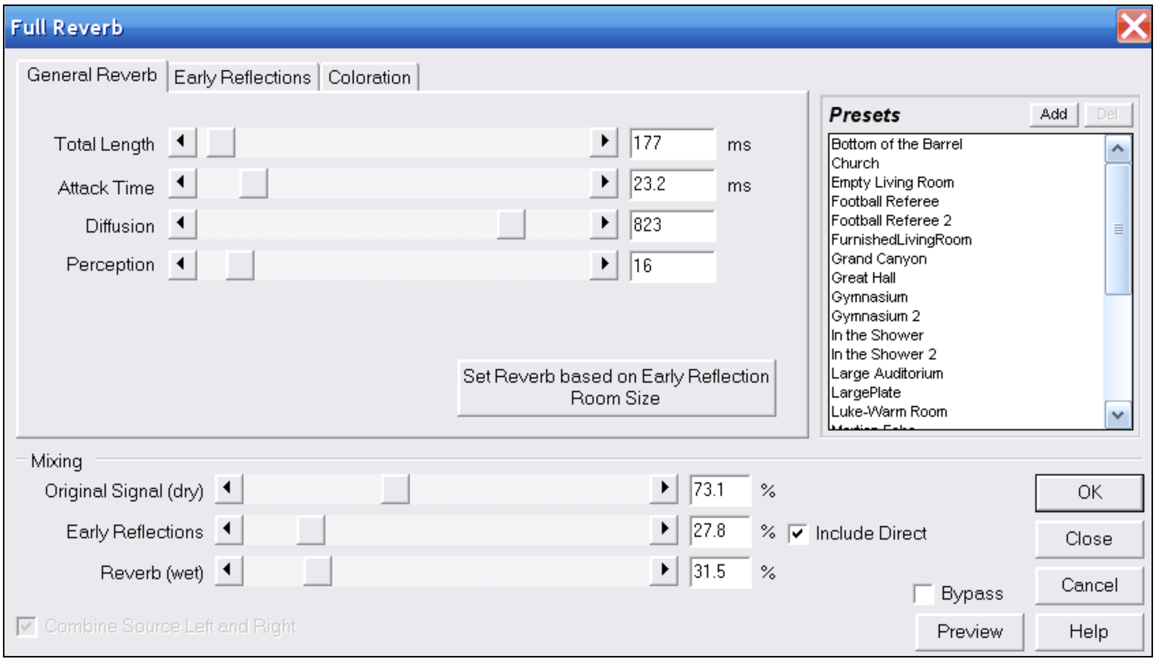


The *QuickVerb* has a feature called "Room size". This is a useful feature and together with the decay setting, the user can adjust their choice of reverberation type. Both the standard reverb and the *QuickVerb* were <u>inefficient</u> in attempting to match the anechoic impulse to the small room impulse. In summary, when adjusting the settings in these two reverb options, the outcomes were unpredictable and did not allow for accurate manipulation of the anechoic impulse in terms of octave-specific RT. However, it should be highlighted again that this application is an extreme example of applying AR.

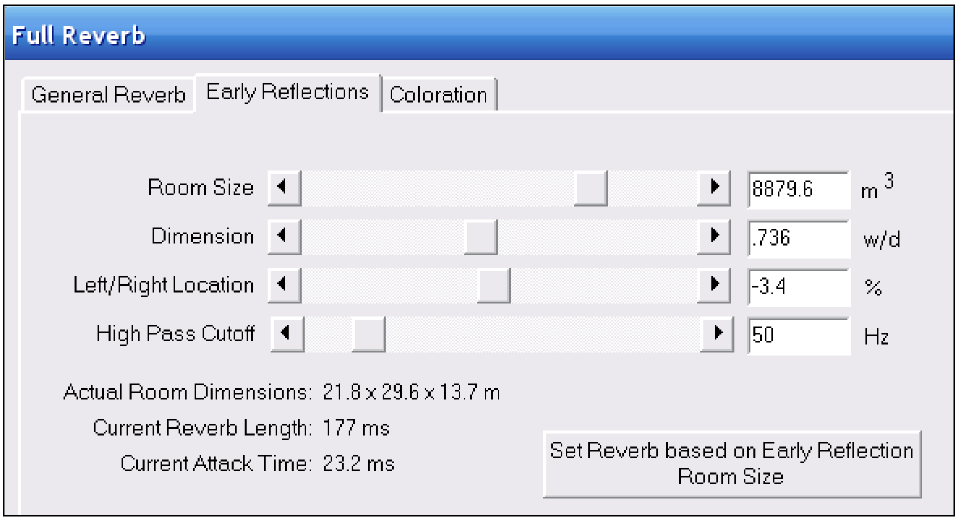
The third option shown in Figure 5.11 provided a better output that was easier to predict and quantify. This option has three parts to it, namely general reverb, early reflections and colorations. The screen shots are shown next. Note that the colorations field has a graph with a quality factor and period option (Fig. 5.11c).

Figure 5.11 a-c: Option 3: Full Reverb.

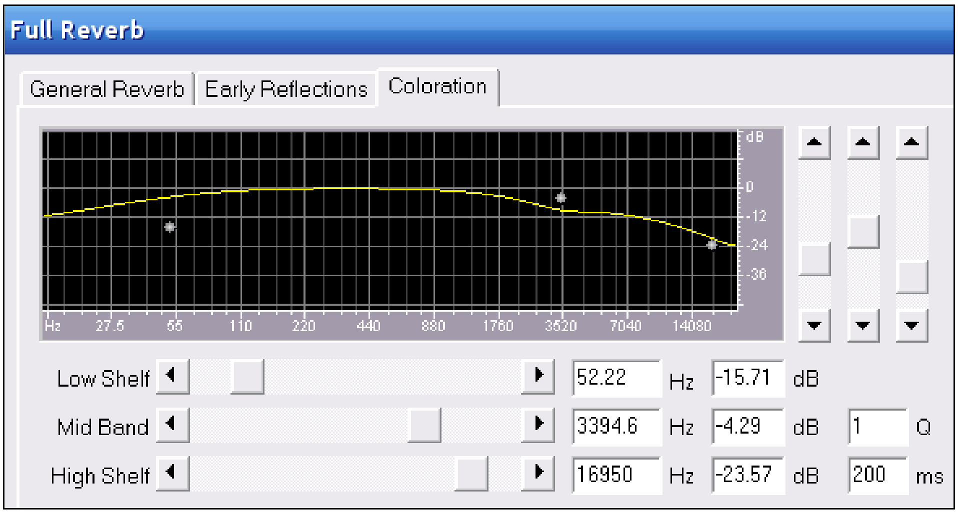
a) General Reverb Tab.



b) Early reflections Tab.



c) Colorations Tab.



After much trial and error[38], the colorations field was found to be the best option in making large RT_{60} adjustments. This may be an unorthodox approach as it may evoke an unplanned methodology. This was not the case; incremental steps were taken based on the findings of each previous step, the knowledge of acoustics and the software's response. This was constantly analysed by the WinMLS software for the viability of the result as compared to that of the desired response. One challenge in acoustics is that the editing of sounds needs to be handled case by case and thus it is difficult to blanket all sounds with the same procedure. If a procedure does need to be provided it would either be highly specific for that case, or if applied to many cases, it would be a general guideline[39].

The colourations menu had the ability to focus certain frequencies. The vertical axis shows a dB range while the horizontal axis is in Hz. There are three adjustment bands, low shelf, mid band and high shelf that all have frequency and dB labels. An example of how the colorations field can be adjusted is shown in the next two figures. The low and high shelves were reduced while the mid band was increased. The selectivity was

[38] A question may arise as to why trial and error attempts are part of an engineering study? In essence, this chapter is a test to see if the goal was in fact possible: the goal of using readily available software to perform AR on a dead impulse sound. The experiment does state a clear starting point and a defined end. The method is variable which is not uncommon when a new application is required. Before automation of a method is available, some trial and error work occurs. This is true across many fields and often creativity is favoured over conformance to classical engineering methodology. Once the proof of the result takes place, then a repeatable generalisable method can be developed that can be automated. It is unlikely that a completely automated method for all impulse sounds would be attainable but there is room for some automation. In the case of this chapter, it was decided that adding the comprehensive procedure would add tens of pages of case specific information, which would shift the weight of this chapter from proof to methodology. This would be unnecessary as adjustments to RT are conducted and evaluated by listening tests. Listening tests are used to make decisions when performing AR. However, to be accepted as an engineering proof, this chapter was compiled. This chapter therefore stands in merely as an objective measure to add further weight to the subjective chapter's result and improve the balance between pure engineering and art. Having said that, this chapter still has some ad hoc steps, as it is an original experiment with very little supporting work found in the literature. This issue is discussed further at the end of this chapter under the heading "Engineering Methodology". One of the aims of this dissertation was to determine if one could use already available means to perform AR to such a degree that it could compensate for rooms that have little natural reverberation yet provide the desired sound that the musician would like.

[39] This is true of many fields that have a human perceptual component as a measure of goodness.

affected by the *Q* setting. The graph attempts to depict the focussing effect but is merely a guideline.

Figure 5.12: Coloration settings with a high mid band and low selectivity (Q).

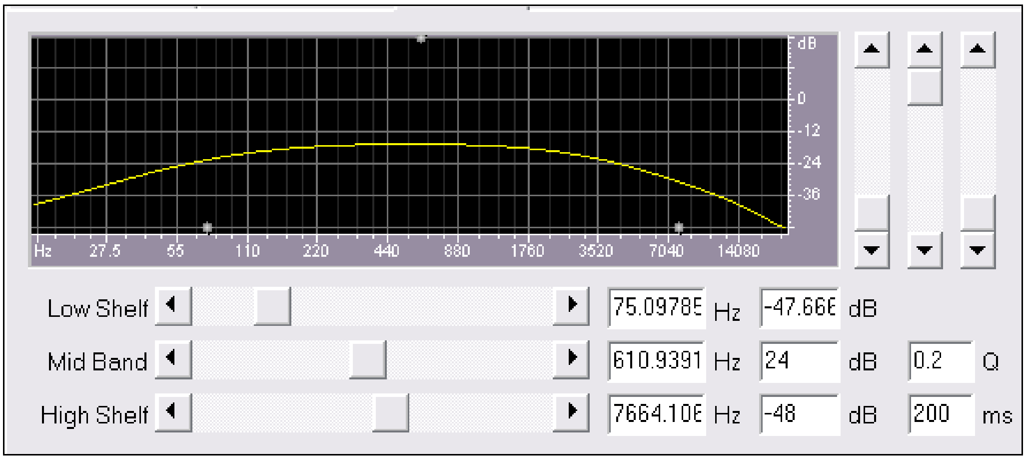
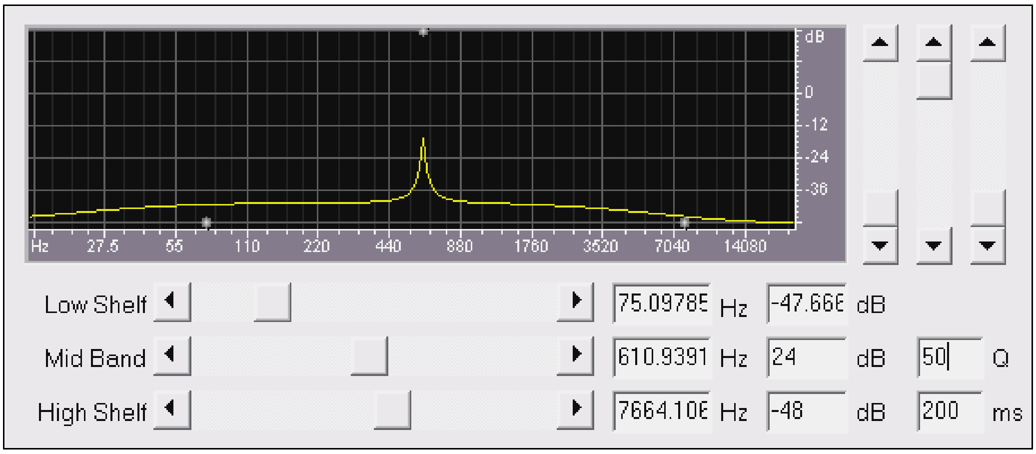


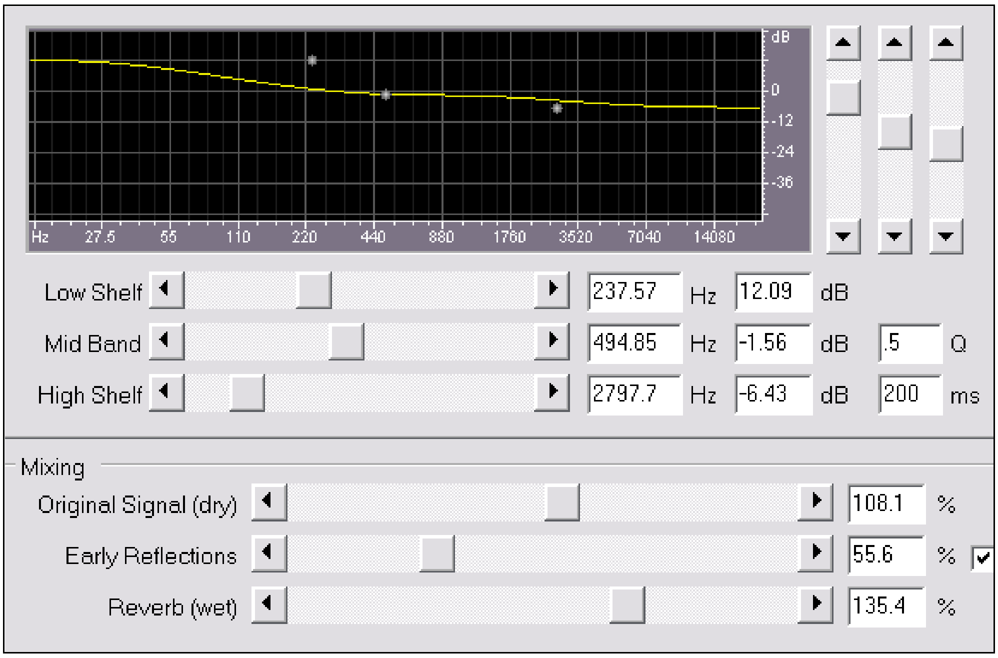
Figure 5.13: Coloration settings with a high mid band and high selectivity (Q).



The frequency range was verified using the WinMLS software and was found to be both consistent and fairly accurate. For example, when making an adjustment to the 1kHz frequency with a low *Q* value chosen and then with high *Q* chosen, the result was shown clearly in WinMLS's octave RT plot. The numerical data was found to be more accurate than the graph though. It is important to have confidence in the software that one uses, thus the software should be verified to provide the user with what it states it is doing. The features used in Cool Edit Pro, such as amplification, graphic equalisation and filtering have been checked to confirm that they provide the user with their stated specification. Thus, from this point forward it is assumed that the parameters that are adjustable in Cool Edit Pro are correct and do provide the user with outputs that are acceptable. This point provides an interesting question to the recording engineer. How accurate are the different software based editing suites available and can they be correlated with each other? This topic is presented again in the future work section of this dissertation (Chapter 7).

The anechoic track was adjusted using the *full reverb* option. The colorations adjustment is shown in the next figure. The noted settings include the low shelf amplified by 12,09dB and the high shelf attenuated by 6,43dB with a period of 200ms. The figure shows the actual preliminary settings that were found to provide the closest RT_{60} response to that of the small room. Further adjustments were made to accommodate the EDT correctly but the settings shown were the first close approximation to the small room. The *Q* was set to 0,5. Under the mixing section, the early reflections were chosen to be 55,6%. The settings not shown include: total length of RT; attack time; diffusion; perception; room size; dimension; left/right location and high-pass cut-off, all of which were adjusted until predictable results were obtained. Unfortunately this was a delicate task as there were more than 10 AR settings with each having a large range of adjustment, thus this process had more than 10 continuous type variables.

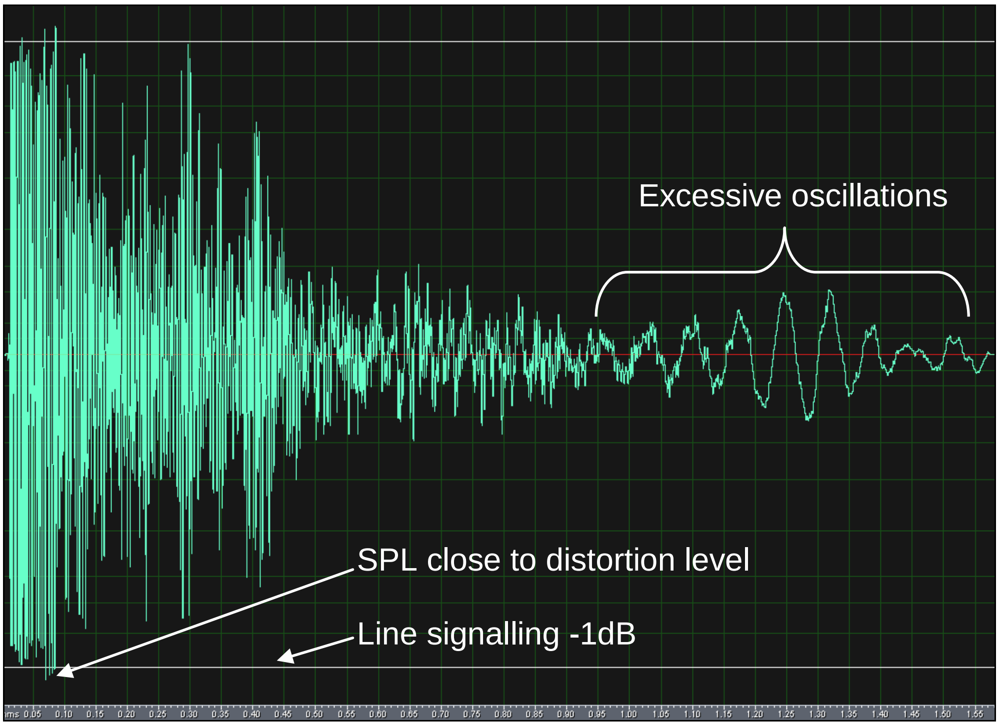
Figure 5.14: Preliminary settings for the full reverb delay function.



When applying large amounts of amplification to a reverberant sound using the colorations option, oscillation was sometimes found to occur. This was a problem that needed to be solved. An example of these oscillations are shown in Figure 5.15. To counteract this problem, the following process was followed:

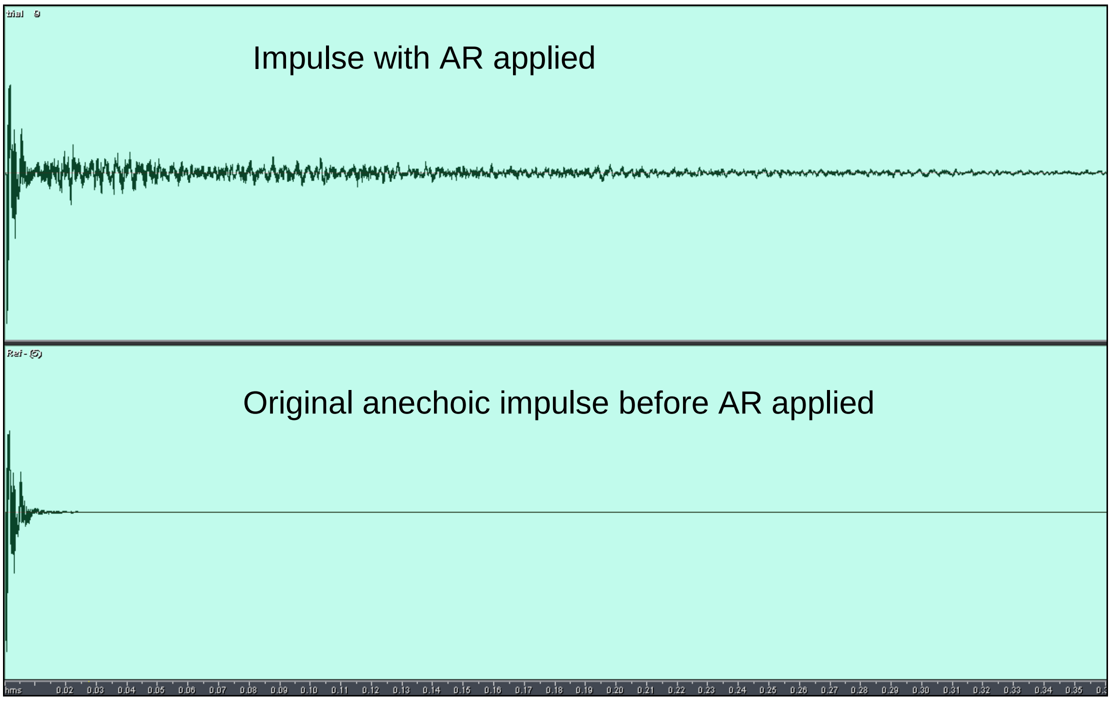
- Reverb applied but with half the amplification set.
- The entire track was then attenuated between 3-8dB.
- Reverb was then reapplied with the other half of the amplification requirement.
- Awareness of the peak sound limit is needed to stop any amplitude distortion (shown graphically as a sinusoidal wave that has taken a square wave appearance at its peaks and troughs).
- Amplification and attenuation settings chosen based on the headroom available handled case by case.

Figure 5.15: Oscillations that may occur when too much coloration is set[40].



The anechoic impulse after the first successful reverberation applied is shown next on one set of axes and the original anechoic impulse is shown below it on another set of axes. The time scale is common to both waves. Note how the top decay continues well after the bottom decay waveform. Also, note that there is similarity in the early part of both waveforms.

Figure 5.16: Comparing the reverberated impulse with its original.



[40] If excessive oscillations were found by visual examination [or heard in the sound test-chapter 6], those outputs were discarded. Not only does the oscillations add to the noise component, they are also not part of the original sound and thus cannot be accepted.

5.2.1.1 Analyses of the Artificially Reverberated Wave

It was noted from the previous figure (Fig. 5.16 showing the comparison between the before and after impulses), that the after impulse seemed to have an inconsistent decay roll-off. The waveform was analysed further and is shown in the next two figures. The first figure shows the impulse response, which depicts the typical roll-off shape (Fig. 5.17). The second figure exhibits the inconsistent decay slope with the labels showing the uneven decay points (Fig. 5.18). An inconsistent decay slope does not necessarily mean there was a problem; however, it was an area of interest, as the small room's decay slope did not exhibit such large inconsistencies. The inconsistencies were that the SPLs of the roll-off waves did not seem to match the small room's SPLs. While there were similar parts of the decay slope between the artificially reverberated impulse wave and the small room's wave, there were also parts that did not match each other in terms of SPL. This raises an important point: manipulating a wave in terms of reverberation is only one of many acoustic parameters that need to be addressed. What is clear here is that the decay SPLs also need to be examined further.

Figure 5.17: View of reverberated wave.

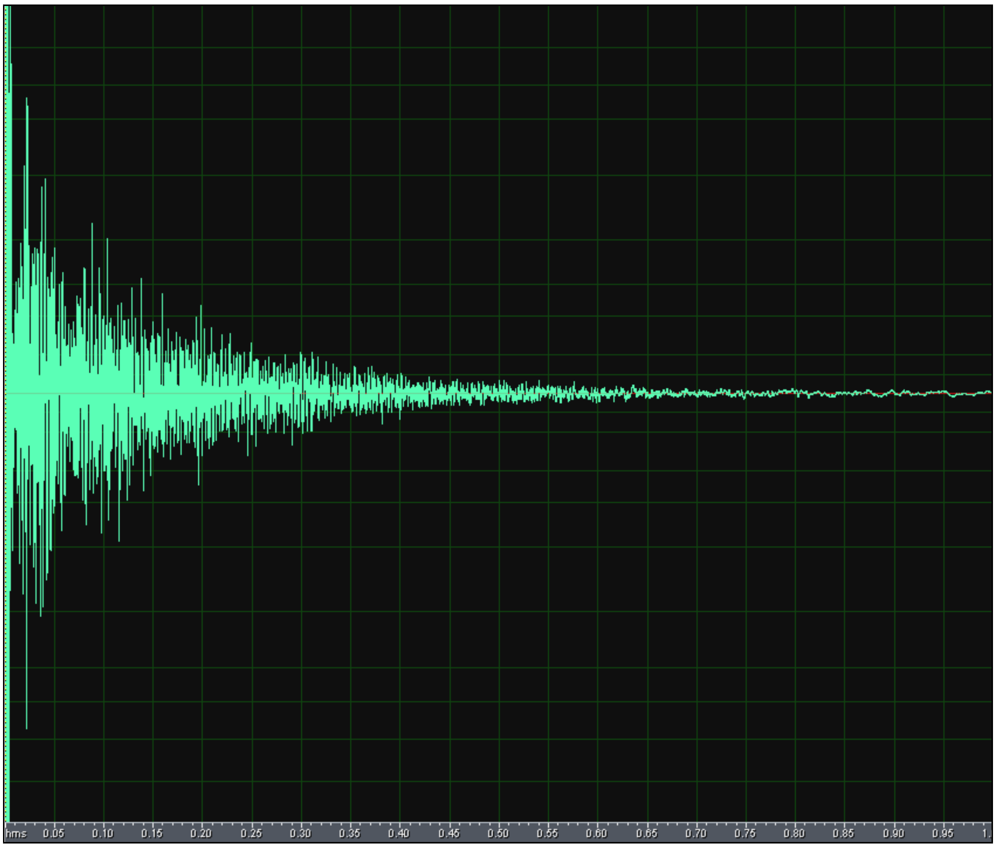


Figure 5.18: Zoomed view of the reverberated wave showing uneven decay.

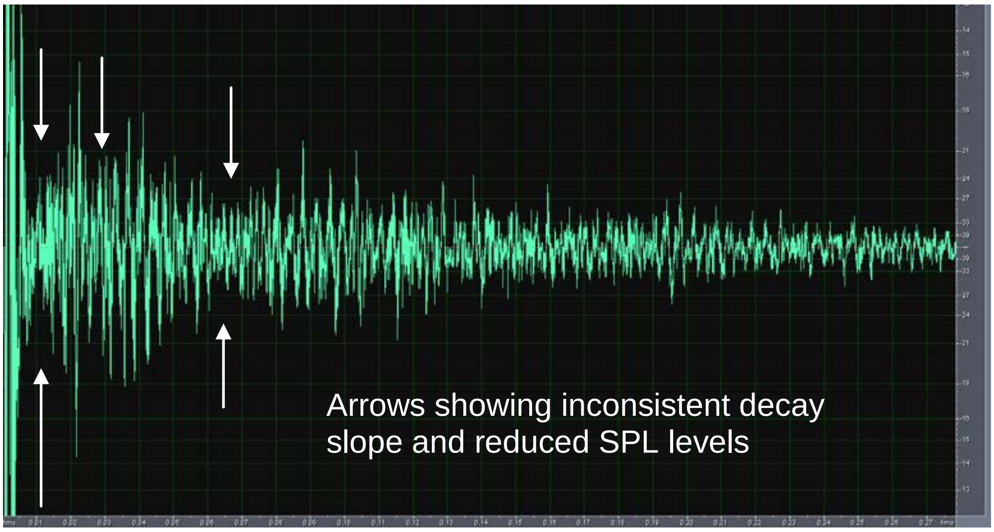
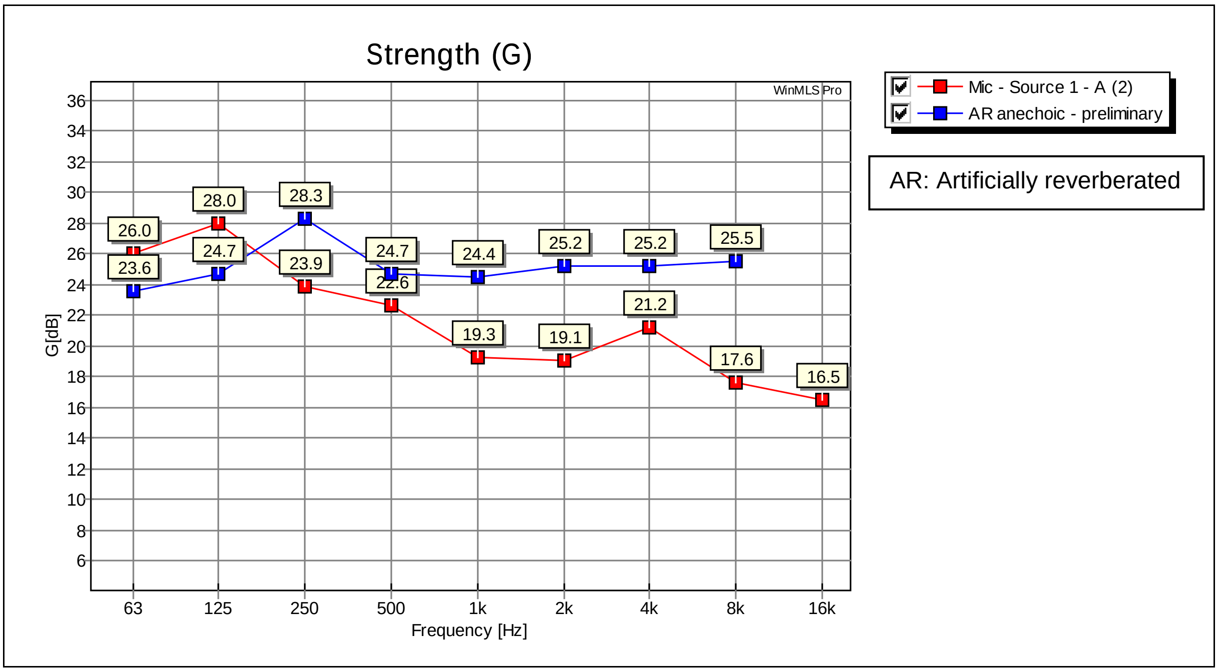


Figure 5.19 shows the sound strength of the artificially reverberated wave and the small room's wave. Both these waves have different sound strengths at different octave bands. This topic is visited again in the next chapter.

Figure 5.19: Sound strength of two impulses that have similar RT yet differ in their resultant sound strength for each octave band.



Continuing with the application of AR, the anechoic impulse, which was artificially adjusted was then imported into WinMLS for octave RT checking until a close match was found. This was initially a trial and error approach. As the parameters available from Cool Edit Pro were adjusted and analysed in WinMLS, an understanding as to what the effect was of each setting became apparent. The process became quicker as the delay effects became more predictable. The 380-page user manual of Cool Edit Pro does explain each feature of its program but does not expand on the technical process underlying the DSP steps nor provide the specific frequency values that each effect would reflect. For example, to follow is an extract describing the function of the Q setting which was shown in Figure 5.11c (Syntrillium Software Corporation, 2002:138):

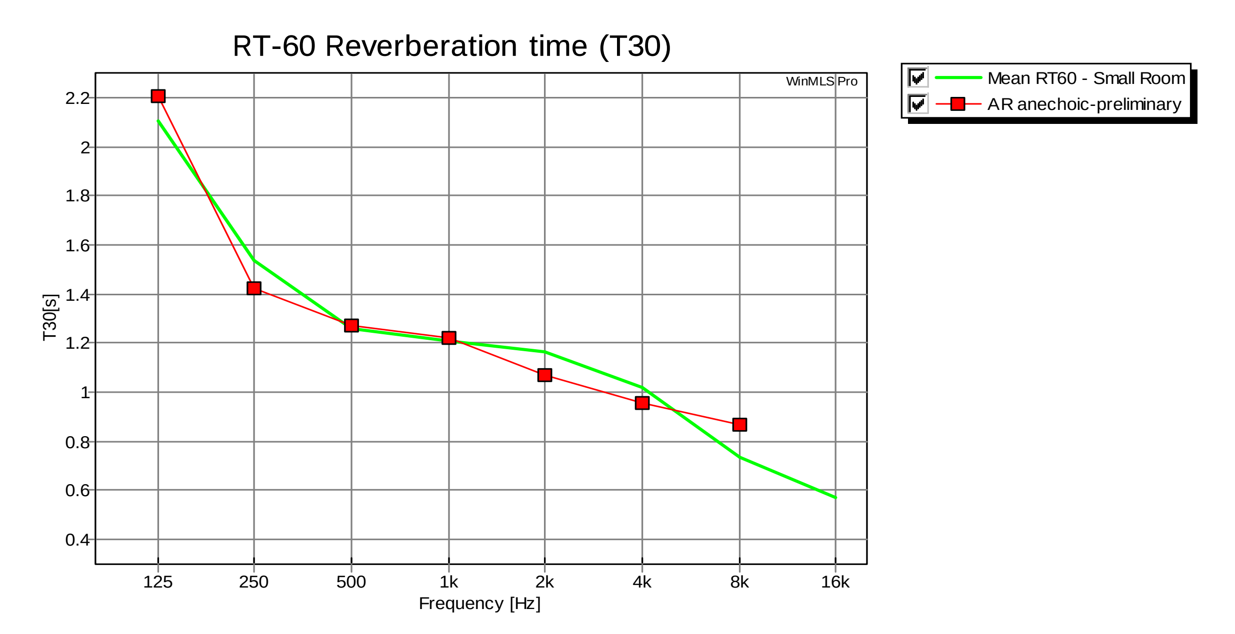
> Q Box:
> This value describes how wide the mid band's affective area is frequency-wise. Lower values affect a wider range of frequencies (hence higher values will narrow the effective range). For distinct resonance, use values like 10 or higher. For just general boosting or cutting a wide range of frequencies, use lower values like 2 or 3.
> Sometimes it may be desirable to enhance some quality of the audio being processed, like to bring out a singer's voice. Try boosting the frequencies just around the natural frequency of their voice to enhance resonance in that range – say, in the 200Hz to 800Hz range.

The manual explains in detail how to use the program, which is as expected. Competing software suites in this regard also have extensive manuals but also fall short in uncovering specific values for each DSP step. This is not surprising as firstly they are profit driven companies who would not readily explain all the programming that makes the software do what it does, as that is the very thing they are selling. Secondly, the target market of recording suites are people who generally do not have extensive knowledge in the inner workings of electronics and the software programming skills that go with it. Thirdly, even if the consumer is a specialist in this regard, they would probably not be interested in the behind the scenes aspects as but rather the outputted sounds that the software generates, which is what the majority of musicians use as a measure of fit.

There are various methods that can be used to perform software based AR. The software packages perform them using algorithms, which were originally based on the replication of the hardwired electronic methods. The application of post-processing AR is essentially the convolving of an input signal with the specific impulse response of an acoustic space. The multiplying of these two responses often relies on FFT for post-processing options. Real-time RT adjustments would not readily use FFT as the entire input stimulus would need to be defined prior to processing. For real-time AR a convolution technique such as direct linear finite impulse response would be used where the processing occurs in the time domain for each input sample (Browne, 2001). There are numerous ways to model the sound space including geometric models, waveguide reverberation, multiple-stream reverberation (Roads, 1996). In terms of this chapter, the focus was on the efficacy of one particular software method in performing AR. This AR process is post-processing and thus does not occur in real-time.

The next figure shows the preliminary result of the application of AR for the small room.

Figure 5.20: Preliminary result for artificially reverberated anechoic impulse for small room (RT_{60}).



The red plot (AR anechoic-preliminary) is the modified wave derived from the anechoic wave. The overall similarity is close; however, the linearity after 1kHz differs from the curvature of the small room's mean RT_{60}. The tabulated results and their percentage deviations are shown in the following table.

Table 5.3: Comparison of preliminary artificially reverberated anechoic versus small room's mean RT_{60}.

x-axis F [Hz]	Mean RT60 - Small Room (s)	AR (artificially reverberated) anechoic – preliminary (s)	Percentage deviation (%)
125	2.10	2.21	5.24
250	1.54	1.42	-7.79
500	1.26	1.27	0.79
1000	1.21	1.22	0.83
2000	1.17	1.07	-8.55
4000	1.02	0.95	-6.86
8000	0.74	0.87	17.57
16000	0.57	---	---

The EDT of the impulse was manipulated to be as similar to the small room's EDT as possible. Adjusting the EDT is a sensitive process. Minor setting changes were found to have large affects in terms of the EDT. In general, more parameters affected the EDT than the RT_{60} and thus adjusting the EDT was easier, however it took longer as the increments required were very small. The results are plotted in the next figure. The artificially reverberated anechoic impulse follows the EDT of the small room closely until 5kHz where it starts to deviate.

Figure 5.21: Preliminary result for artificially reverberated anechoic versus small room (EDT).

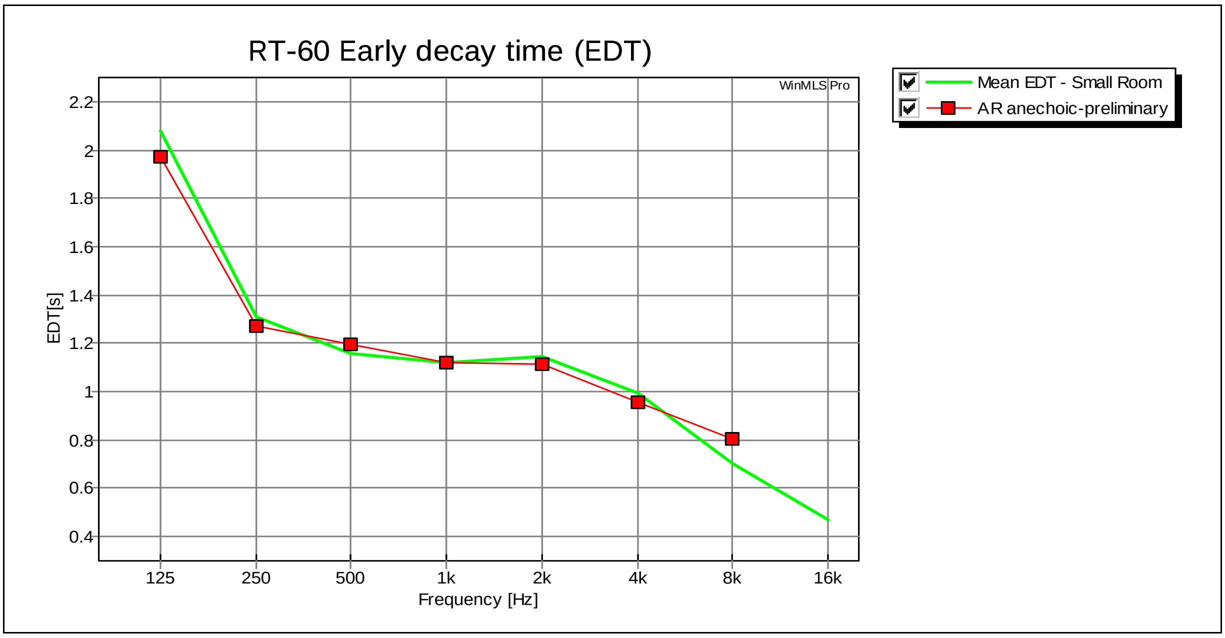


Table 5.4: Comparison of preliminary artificially reverberated anechoic's EDT versus small room's mean EDT.

x-axis F [Hz]	Mean EDT - Small Room (s)	AR anechoic – preliminary (s)	Percentage deviation (%)
125	2.08	1.97	-5.29
250	1.31	1.27	-3.05
500	1.16	1.19	2.59
1000	1.12	1.12	0.00
2000	1.14	1.11	-2.63
4000	0.99	0.95	-4.04
8000	0.71	0.81	14.08
16000	0.47	---	---

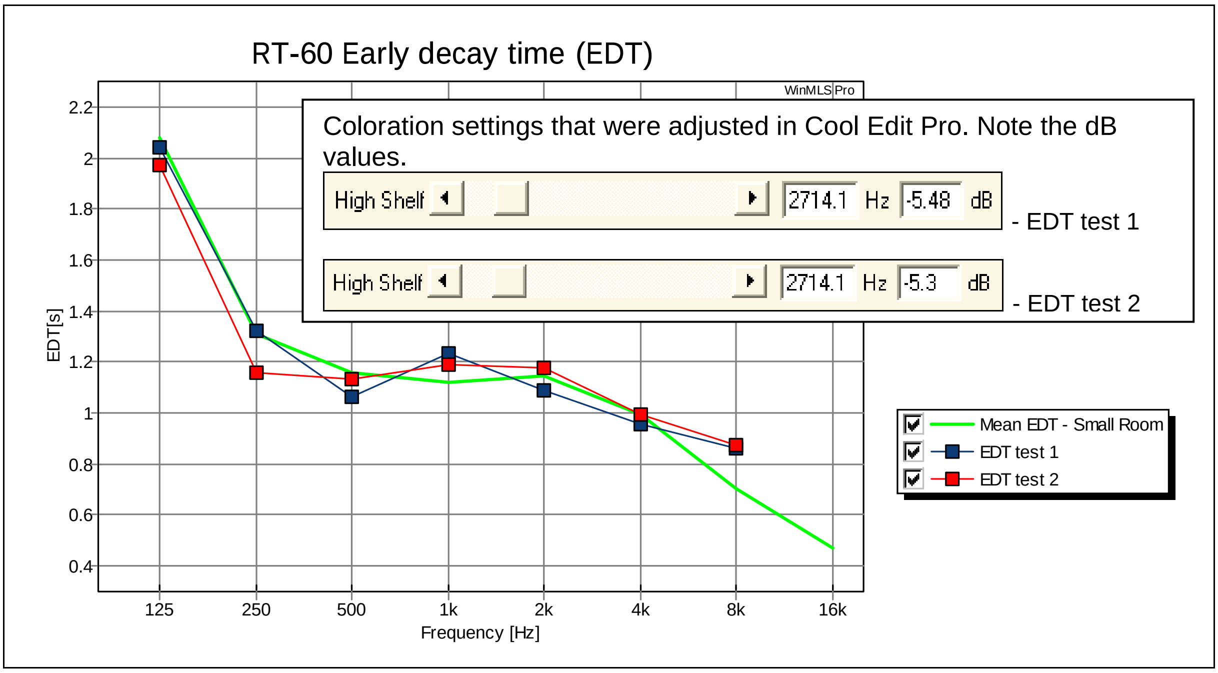
Overall, the EDT of the artificially reverberated anechoic impulse compared to that of the small room was a good match. The EDT and RT_{60} are characteristic of the single anechoic impulse that was artificially reverberated. While it would be much simpler to have one edited wave file of the anechoic to match the RT_{60} and then another separate wave file edited to match the EDT of the small room, but that would be technically incorrect. The point is that the anechoic impulse is supposed to have the same RT characteristic as the small room, therefore only a single wave file is used for both, and once finalised it must provide a similar response for both RT_{60} and EDT without any further adjustments. Unfortunately, making changes to the RT_{60} affects the early decay response considerably. Thus, it was found that first the RT_{60} was matched and then the EDT was matched while keeping the RT_{60} as unaffected as possible. The procedure is summarised as follows:

- Using AR software such as Cool Edit Pro, adjust the anechoic wave's RT to as close to the desired RT as possible. This starts as trial and error. When the AR is applied, the wave file needs to be assessed by octave specific RT software in this case WinMLS was used.

- Obtain a close match for RT_{60} and save the wave file with comments and signal processing settings. (As there are many attempts and settings, comments make it easier to understand the software's response and allow for rolling back to previous settings).
- Load the preliminary wave file into the octave band EDT software and assess.
- Adjust the EDT in the AR software until it is close to the desired EDT by checking the response in the RT software.
- Keep checking that the adjustments to EDT do not affect the RT_{60} response already obtained. This is a catch twenty-two situation where some parameters improve EDT but change the RT_{60} and thus compromise is required. Some settings may need to be omitted and compensated with other settings.
- When EDT is obtained and the RT_{60} is still within the desired range of less than 10% deviation, the goal is achieved in terms of a preliminary RT match.

It is much easier to match either EDT *or* RT_{60} individually. However, the challenge was to match them both for the same wave file. The EDT settings are extremely sensitive, for example, a change of only 0,18dB to one parameter has a large impact on the octave EDT as shown in Figure 5.22. The desired wave is in green and is the mean EDT for the small room. The two test waves are labelled "EDT test 1" and "EDT test 2". The only difference between the two was an adjustment to the high frequency (high shelf) coloration dB setting from -5,48dB (EDT test 1) to -5,3dB (EDT test 2). At first inspection, it may seem that the result is unpredictable as only the high shelf was adjusted, yet the two test waves now differ in their low bands as well. The reason was the Q-factor setting which changed. For these two tests a very low *Q* setting was chosen (0,6). Note the change in magnitude and curvature for only a slight dB adjustment.

Figure 5.22: Example of high sensitivity for EDT from small changes in AR settings.



5.2.2 Results

The results for the RT_{60} are shown in the next figure. The original anechoic (before condition) as well as the artificially reverberated response (after condition) are shown with the mean RT_{60} of the small room plotted in green. Only the before and after responses have their RT labels shown in the figure. The full tabular values are shown after the figure. Only the results that could be compared were included, for example, the small room had no 63Hz octave measurement while the anechoic had no 16kHz, thus those were excluded.

Figure 5.23: Before and after AR was applied compared to the small room's mean RT_{60}.

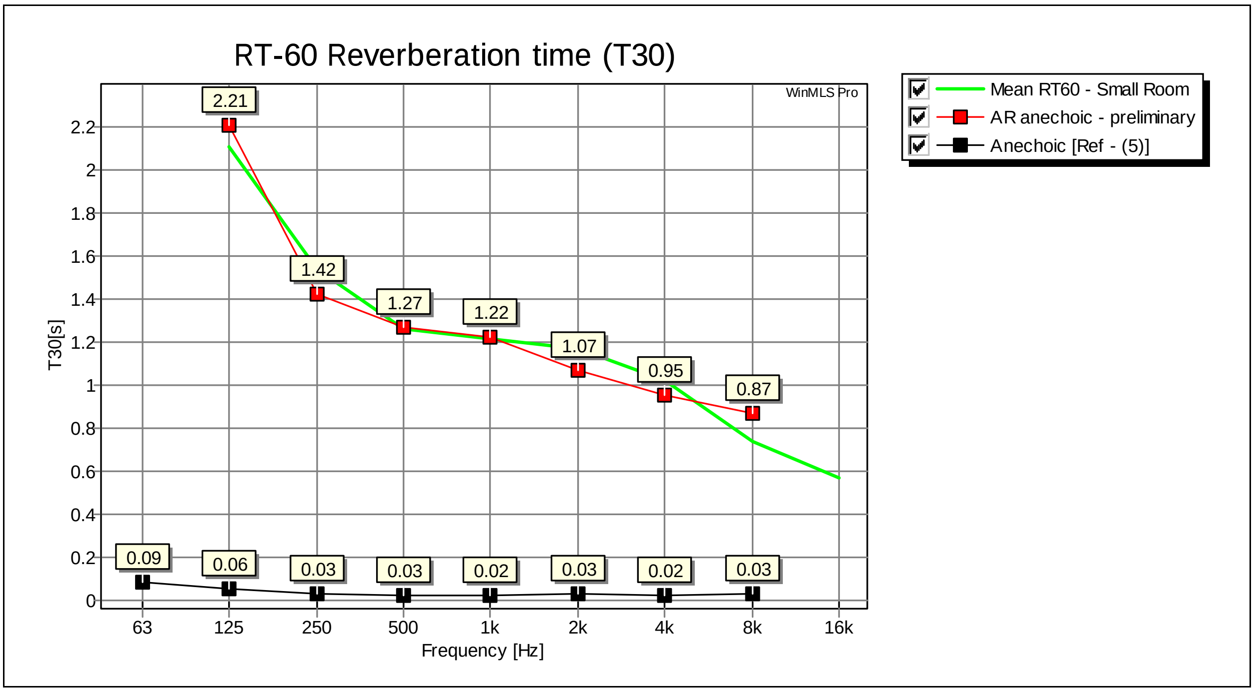


Table 5.5: Summary of results for RT_{60}: Anechoic converted to small room's response.

x-axis F [Hz]	**BEFORE** Anechoic (Original) [Ref - (5):RT_{60}] (s)	**AFTER** AR anechoic – preliminary (s)	**GOAL** (Mean RT60 - Small Room) (s)
63	0.086	---	---
125	0.055	2.205	2.10
250	0.028	1.421	1.54
500	0.025	1.271	1.26
1000	0.025	1.222	1.21
2000	0.030	1.070	1.17
4000	0.023	0.955	1.02
8000	0.031	0.866	0.74
16000	---	---	0.57

The results obtained for the EDT are summarised next.

Figure 5.24: Before and after AR applied compared to the small room's mean EDT.

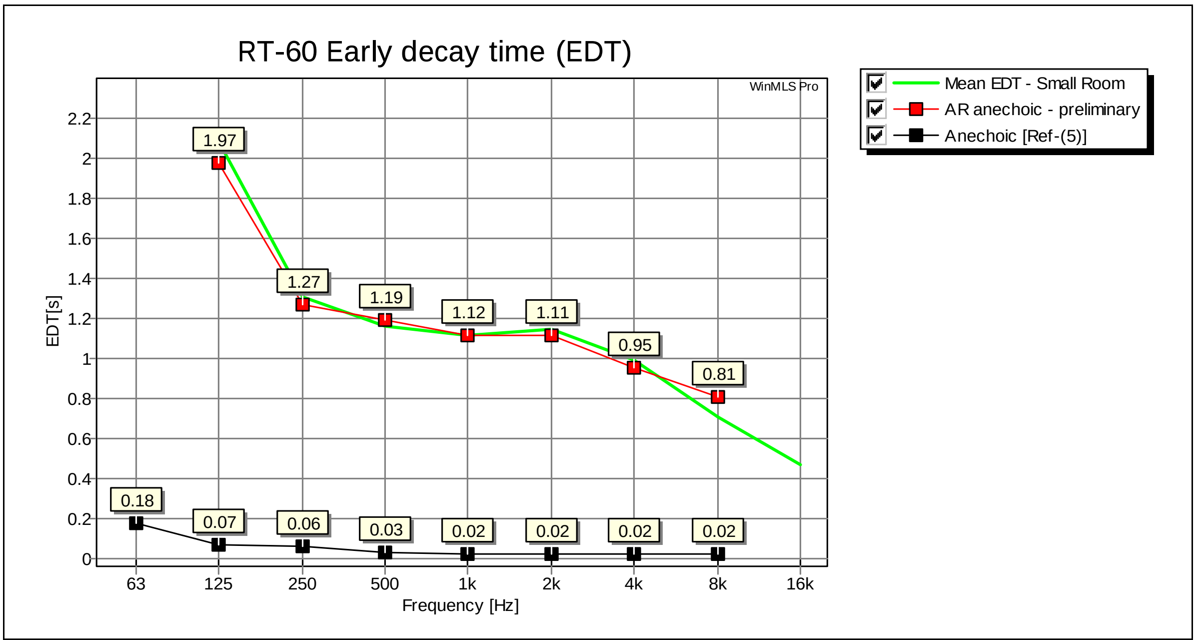


Table 5.6: Summary of results for EDT: Anechoic converted to small room's response.

x-axis F [Hz]	**BEFORE** Anechoic (Original) [Ref - (5):EDT] (s)	**AFTER** AR anechoic – preliminary (s)	**GOAL** (Mean EDT - Small Room) (s)
63	0.179	---	---
125	0.067	1.97	2.08
250	0.059	1.27	1.31
500	0.027	1.19	1.16
1000	0.023	1.12	1.12
2000	0.024	1.11	1.14
4000	0.022	0.95	0.99
8000	0.024	0.81	0.71
16000	---	---	0.47

Both the RT_{60} and EDT matched each other well through the use of AR. The EDT was a better result as it had a lower percentage deviation for each of the octave bands of interest.

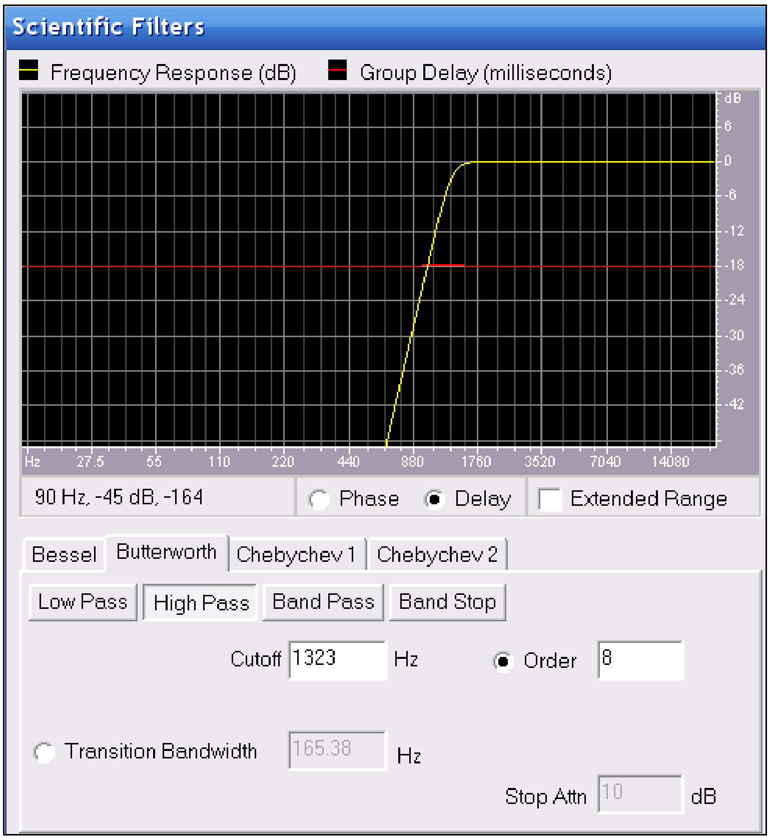
5.2.2.1 Further Artificial Reverberation Applied

Further signal processing was applied to the impulse wave to obtain a better RT_{60} for two reasons, firstly to illustrate another method, secondly to show the degree of control that is available from software based AR.

As the full reverb parameters were unable to provide a closer match to the small room's mean RT_{60}, it was decided to manually generate reverberation in the octave bands that needed additional reverb; in particular, the octave bands above 1kHz were further manipulated. The method used differed from the previous and will be discussed next.

The preliminary impulse was amplified and then filtered using the "scientific filters" option in Cool Edit Pro. This option allows for the various filter types and their respective roll-off types. A high-pass filter (HPF) with a Butterworth response was chosen that had an eighth order roll-off (48dB/octave). This filter type was chosen for its flat passband yet steep roll-off characteristic. The settings are shown in the next figure.

Figure 5.25: Impulse filtered with HPF with cut-off of 1323Hz.

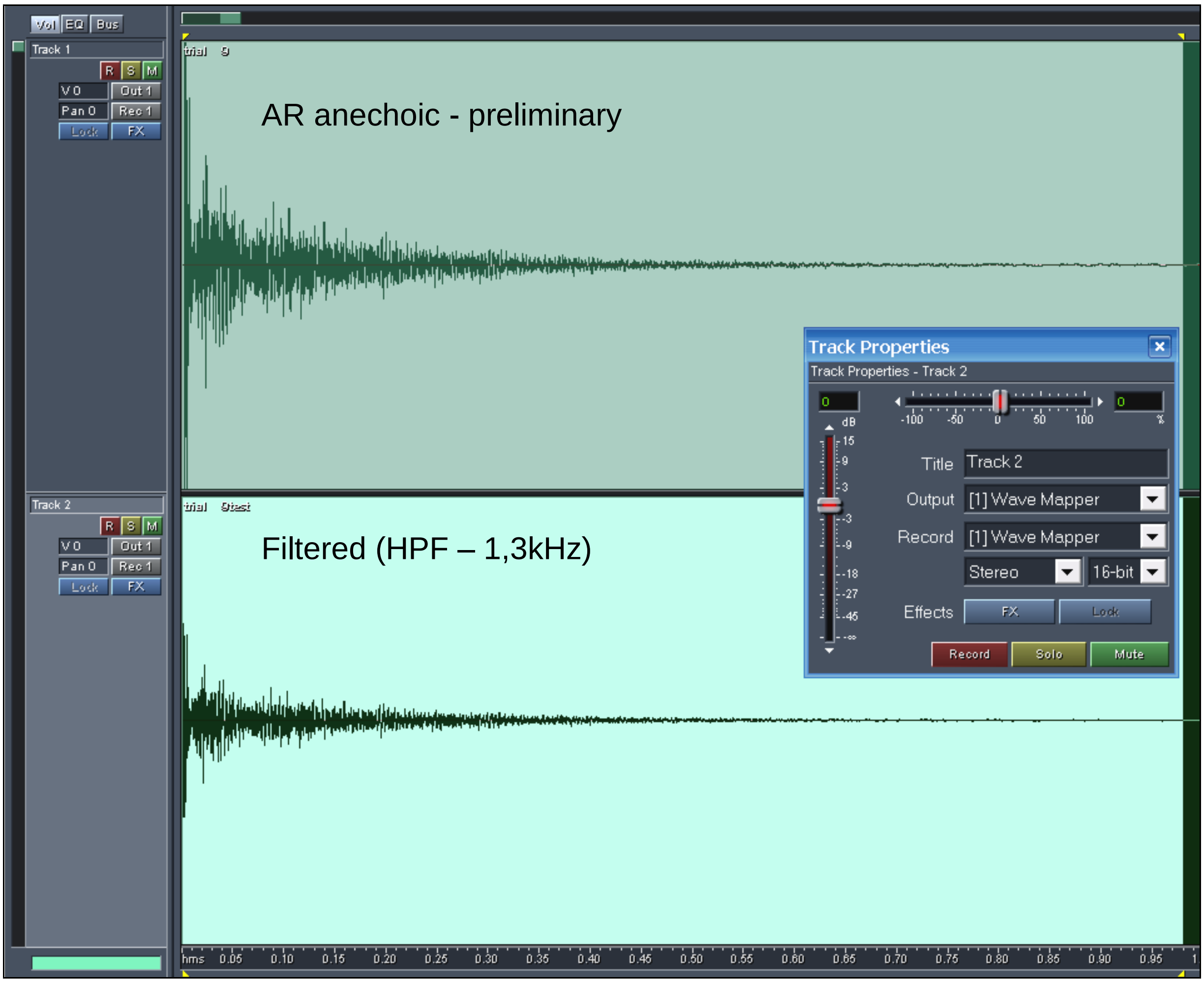


The filtered track was saved as another file; thus, there were two tracks, the artificially reverberated anechoic preliminary and the newly filtered artificially reverberated anechoic preliminary. The filtered track was then subjected to further AR in a similar manner as was done earlier. The filtered track was then inserted into a multitrack[41] view with the original artificially reverberated anechoic preliminary impulse. They were then volume adjusted as when two similar waves are consolidated into one, there is usually an increase in amplitude (unless phase cancelling occurred). The two tracks were mixed together into a single track, which was then checked in terms of octave-specific RT. A slight change in starting time allowed the one track to start slightly out of phase with the other, which needed to be undertaken 26 times before the correct amount of delay was realised. The result was that further reverberation was added in the higher octave bands by manually delaying the one track from that of the other which is essentially what reverberation is. The octave bands below the cut-off point of the Butterworth filter were not affected as much as the octaves within the filter's passband which was as expected. Thus, when the track containing the frequencies above 1.3kHz were delayed and mixed into the original track, it was as though an increase in reverberation for frequencies above 1.3kHz had occurred and thus the octave specific plot shows that the RT of the octave above 1kHz had an increased RT (Figure 5.27). After 4kHz the RT started to reduce.

The settings shown resemble those of a successful attempt. The mixdown view is shown next.

[41] This is a feature of recording software suites that allows more than one track to be placed into an editable viewing area for merging into a cohesive whole or what is commonly termed a *mixdown*. It is not necessarily a new feature as it was available in the 1970's, however those were hard wired electronically based and not software based.

Figure 5.26: Mixdown view of filtered and artificially reverberated track in one window.



The new RT_{60} curve is shown in Figure 5.27 as well as a zoomed view of the area of interest in Figure 5.28. The revised table of results follows in Table 5.7. Note the improvement in the curvature of the decay-slope as well as in the 8kHz octave band.

Figure 5.27: Improvement to the artificially reverberated preliminary wave.

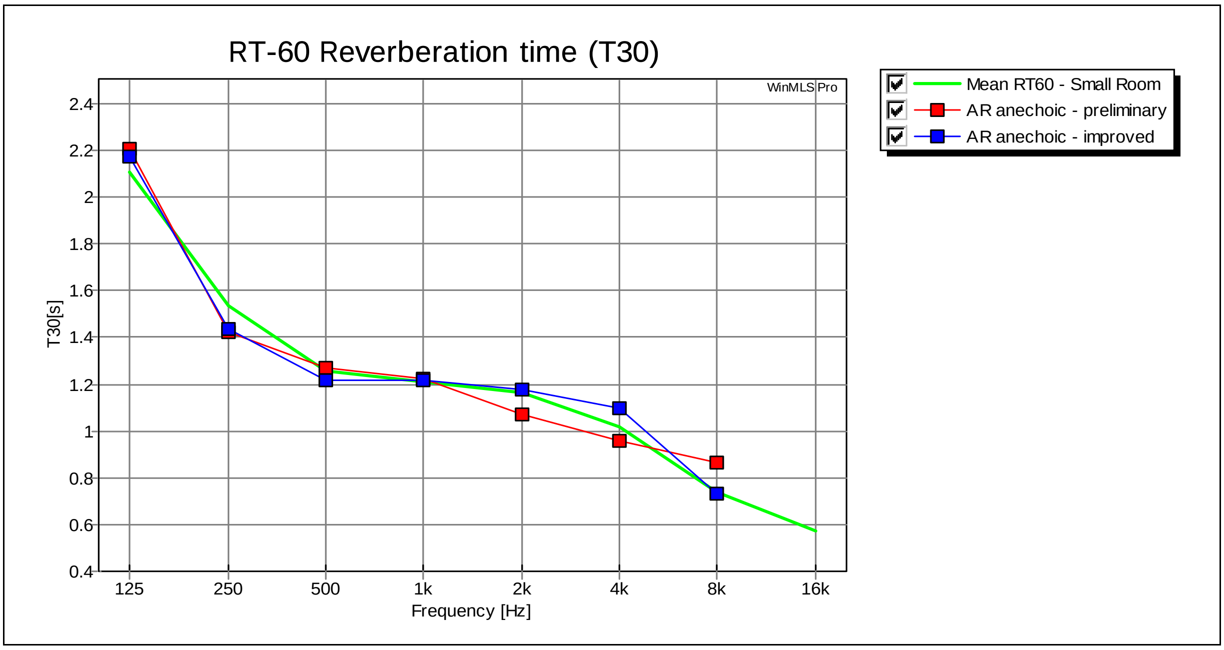


Figure 5.28: Zoomed view of area of interest.

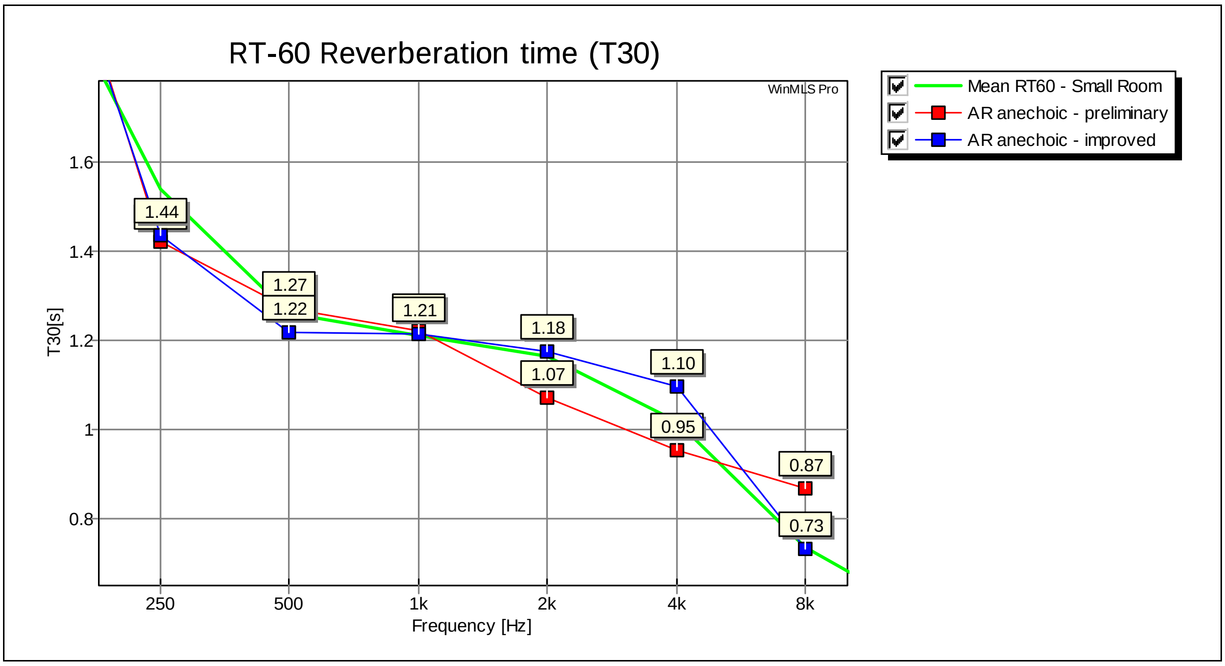


Table 5.7: Summary of results before and after improvement.

x-axis **F [Hz]**	**BEFORE IMPROVEMENT** AR anechoic – preliminary (s)	**AFTER IMPROVEMENT** AR anechoic – improved (s)	**GOAL** (Mean RT_{60} - Small Room) (s)	**Percentage deviation Before (%)**	**Percentage deviation After (%)**
63	---	---	---	---	---
125	2.205	2.170	2.10	5.00	*3.33*
250	1.421	1.435	1.54	-7.73	*-6.82*
500	1.271	1.217	1.26	0.87	-3.41
1000	1.222	1.215	1.21	0.99	*0.41*
2000	1.070	1.175	1.17	-8.55	*0.43*
4000	0.955	1.097	1.02	-6.37	7.55

8000	0.866	0.733	0.74	17.03	*-0.95*
16000	---	---	0.57	---	---

The improved results are italicised in the last column. The main enhancement was that the 2kHz and 8kHz octave bands were considerably improved in their ability to map the small room's response. This came at the expense of slight further deviation in the 500Hz and 4kHz bands. The point of this improvement was to highlight the ability to control the reverberation response. In terms of sound quality, it is unlikely that one would intentionally reduce the higher octave bands as they are usually attenuated naturally and one may rather try to increase them slightly. The additional AR applied provided a closer approximation of the small room's octave specific RT and thus the 8kHz band was now also within the desired specification. The EDT unfortunately would now need to be re-edited using the procedure stated earlier, as changes to the RT_{60} effects the EDT. Changes to the EDT do not affect the RT_{60} that much and sometimes not at all.

5.2.2.2 Sound Test

The DSP applied was verified using WinMLS software to meet the criteria of being an "objective" test. Thus, no comments have been made about the sound of the before and after artificially reverberated anechoic impulse. Out of interest, the impulse was compared to the desired impulse of the small room and it was a positive experience to hear that they sounded quite similar in terms of the reverberation. There was a difference in the loudness of certain frequency bands, for example, the artificially reverberated impulse sounds tinny (lack of bass) compared to the small room's impulse, which sounds fuller in the bass to mid-bass frequency band. It was highlighted earlier that the two impulses did not accurately match each other in their SPL over time plot. Adjustments to the graphic equaliser settings can be used to further match the sound. This can also be conducted in an objective manner by obtaining the SPL of the octave bands for the before and after impulses and adjustment of the graphic equaliser to amplify or attenuate the chosen frequencies. It was found that simply amplifying a single octave band had minimal effect on the RT_{60} as amplification has no effect on the delay or the position in time of the individual frequencies.

5.3 Experiment 2: Anechoic Adjusted to Exhibit the Lecture Halls' Reverberation Time Response

The original anechoic impulse was now manipulated to closely match the response of the lecture hall. The method was similar to the previous experiment and thus only the before condition and results are shown for this experiment. The lecture hall's RT_{60} is shown with the unmodified anechoic impulse response in the next figure and thereafter the EDT is plotted in Figure 5.30.

Figure 5.29: Relative comparison of anechoic and lecture hall's RT_{60} shown on same axes.

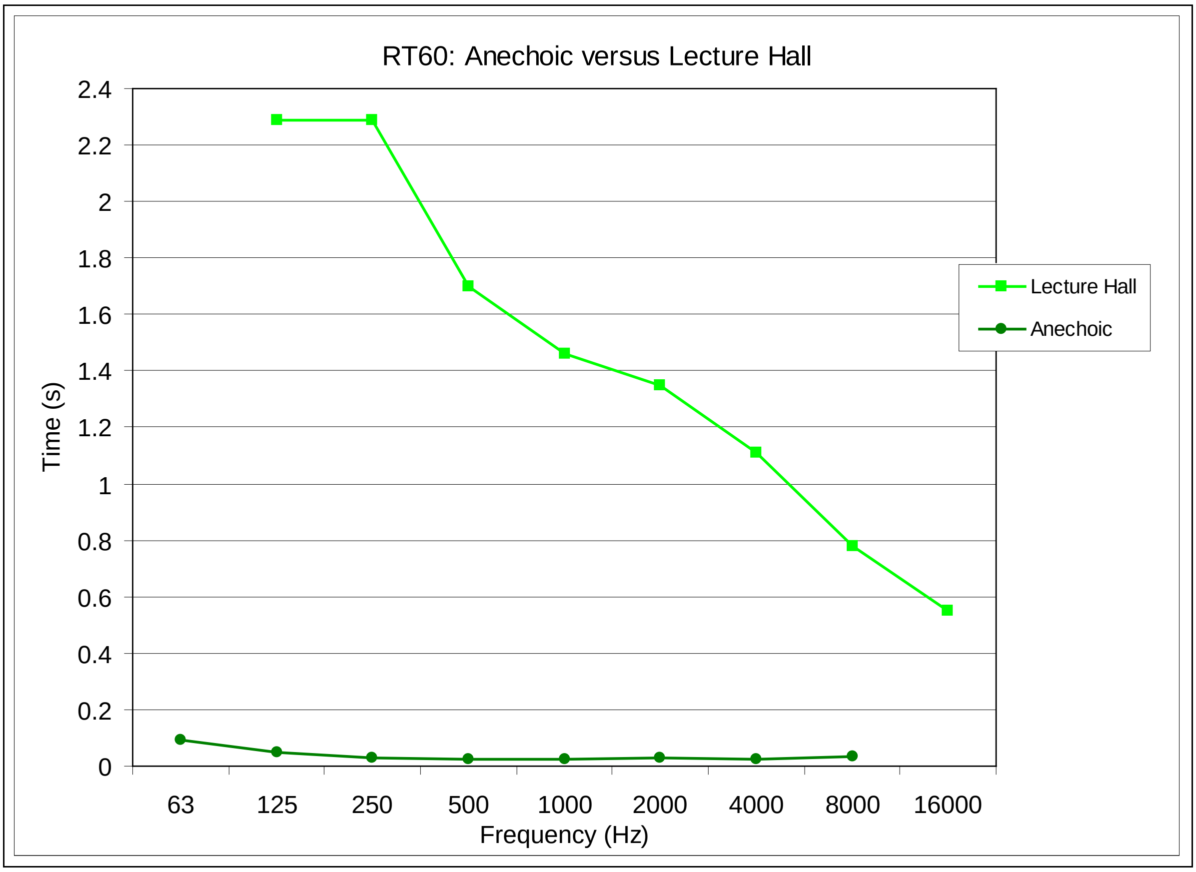
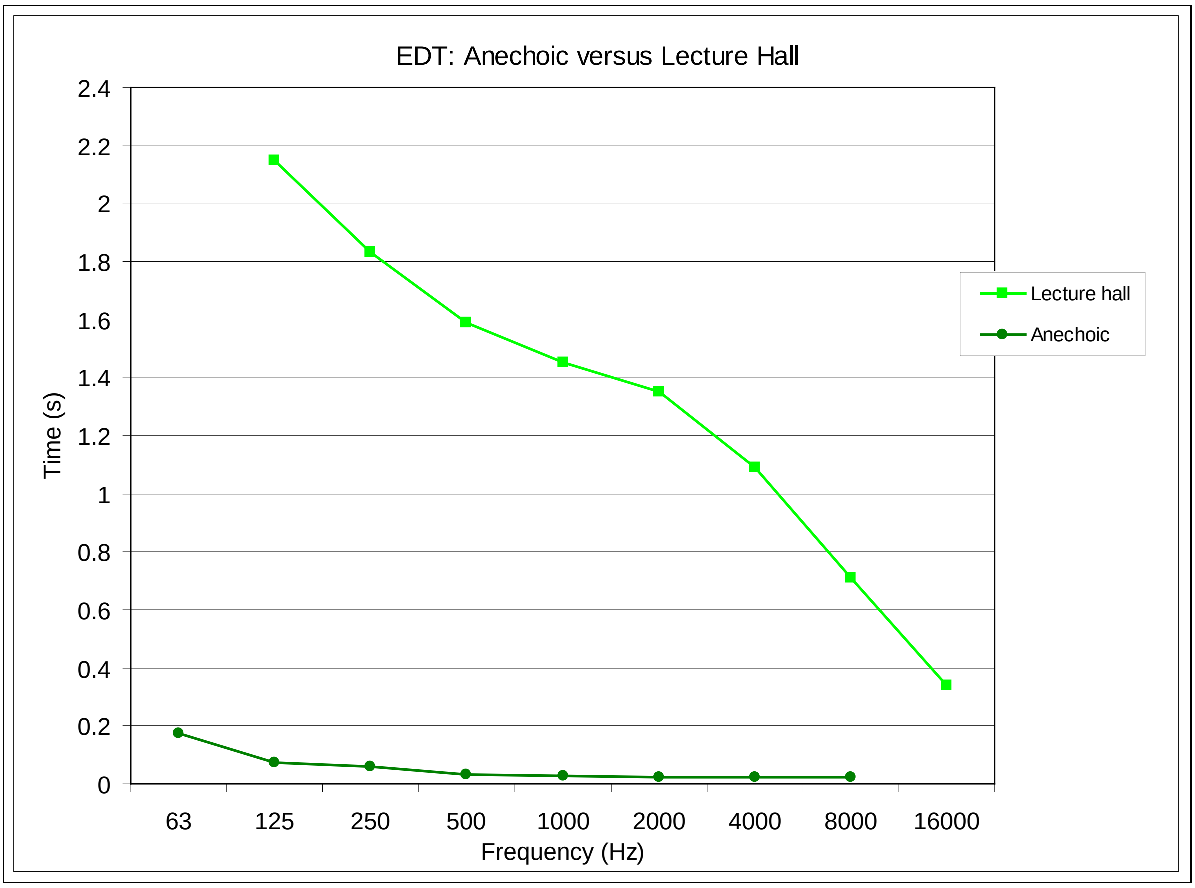


Figure 5.30: Relative comparison of anechoic and small room's EDT shown on same axes.



After applying artificial reverberation, the following results were obtained for the RT_{60} as shown in the next figure and associated table.

Figure 5.31: Before and after AR applied compared to the lecture hall's mean RT_{60}.

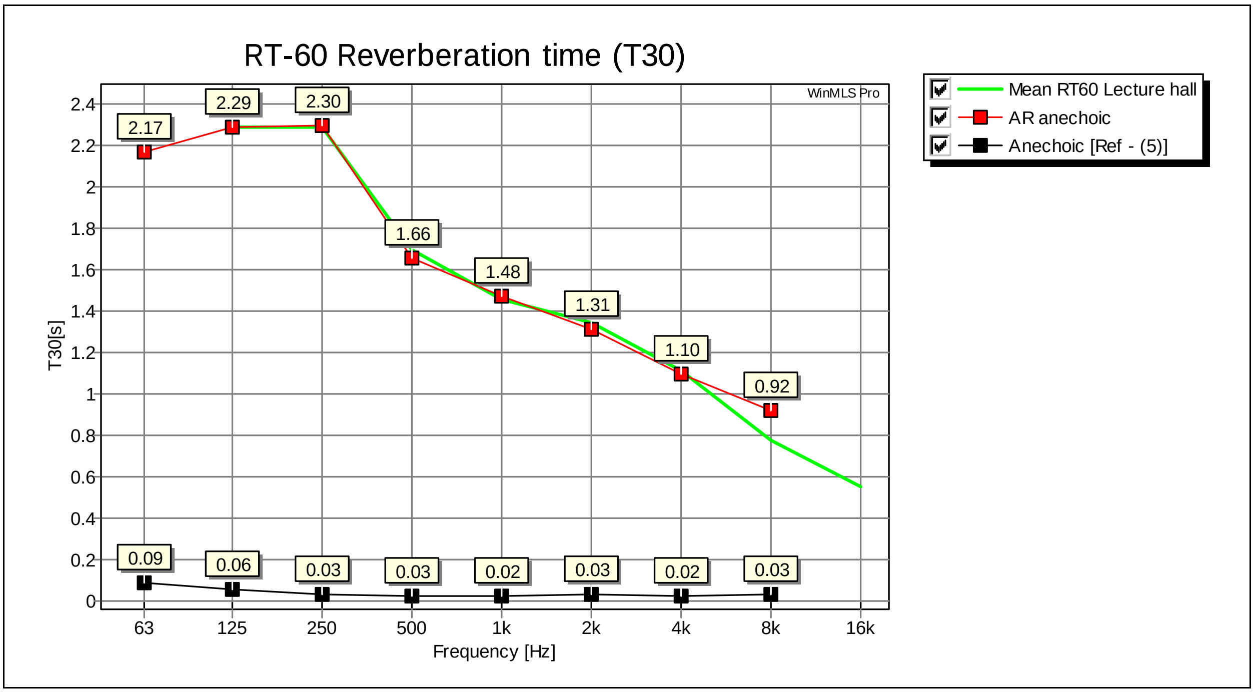


Table 5.8: Summary of results for RT_{60}: Anechoic converted to lecture hall's response.

x-axis F [Hz]	**BEFORE AR** Anechoic (Original) [Ref - (5):RT_{60}] (s)	**AFTER AR** AR anechoic (s)	**GOAL** (Mean RT_{60} – Lecture Hall) (s)	**Percentage deviation (After AR vs. Goal) (%)**
63	0.086	2.17	---	---
125	0.055	2.29	2.29	0.00
250	0.028	2.30	2.29	0.44
500	0.025	1.66	1.70	-2.35
1000	0.025	1.48	1.46	1.37
2000	0.030	1.31	1.35	-2.96
4000	0.023	1.10	1.11	-0.90
8000	0.031	0.92	0.78	17.95
16000	---	---	0.55	---

The results obtained for the EDT are summarised next.

Figure 5.32: Before and after AR applied compared to the lecture hall's mean EDT.

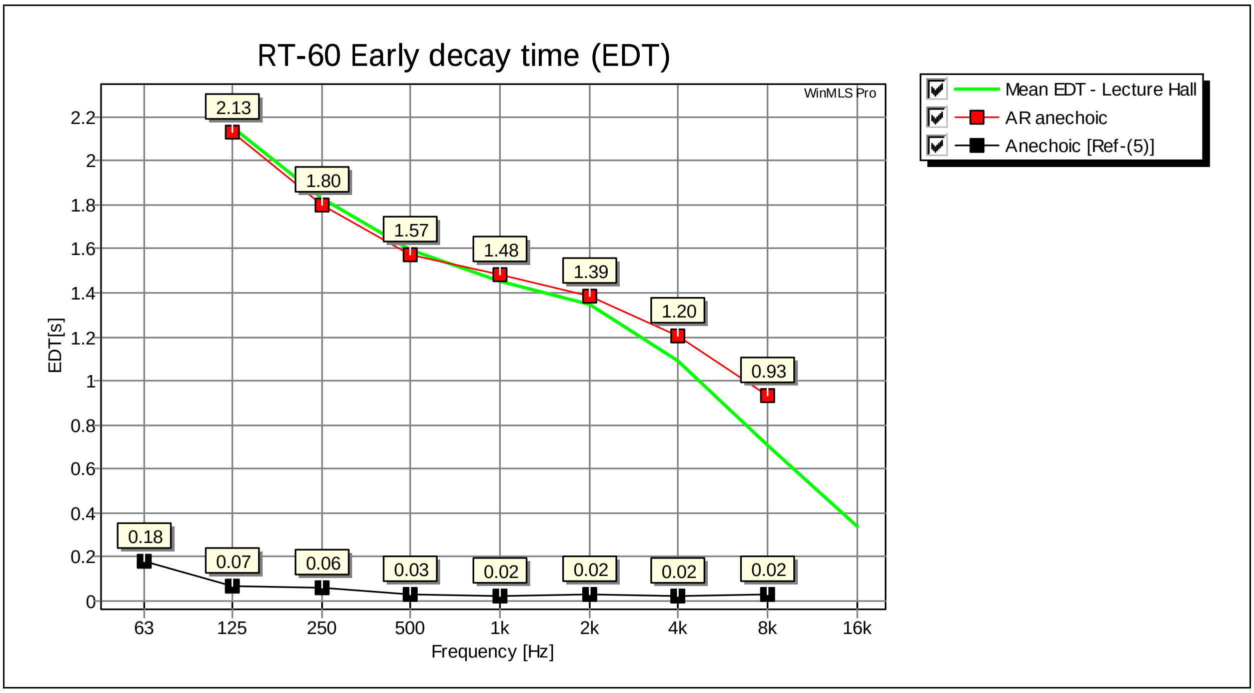


Table 5.9: Summary of results for EDT: Anechoic converted to lecture hall's response.

x-axis F [Hz]	**BEFORE AR** Anechoic (Original) [Ref - (5):EDT] (s)	**AFTER AR** AR anechoic (s)	**GOAL** (Mean EDT – Lecture Hall) (s)	**Percentage deviation (After AR vs. Goal) (%)**
63	0.179	---	---	---
125	0.067	2.13	2.15	-0.93
250	0.059	1.80	1.83	-1.64
500	0.027	1.57	1.59	-1.26
1000	0.023	1.48	1.45	2.07
2000	0.024	1.39	1.35	2.96
4000	0.022	1.20	1.09	10.09

8000	0.024	0.93	0.71	30.99
16000	---	---	0.34	---

As stated earlier, after the EDT has been artificially adjusted, some change may occur in the *RT_{60}* response. This was the case with this experiment. The *RT_{60}* was re-assessed by loading the wave file that was just modified in terms of EDT and plotted in octave specific RT. The result is shown in Figure 5.32. Note the two lowest octaves have deviated from the desired goal. All the other octave bands were either unaffected or improved slightly. An overall positive result was obtained for both *RT_{60}* and EDT and the desired goal of achieving a match with a percentage deviation of less than 10% was achieved for the octave bands 125Hz to 4kHz.

Figure 5.33: RT_{60} after EDT artificially reverberated.

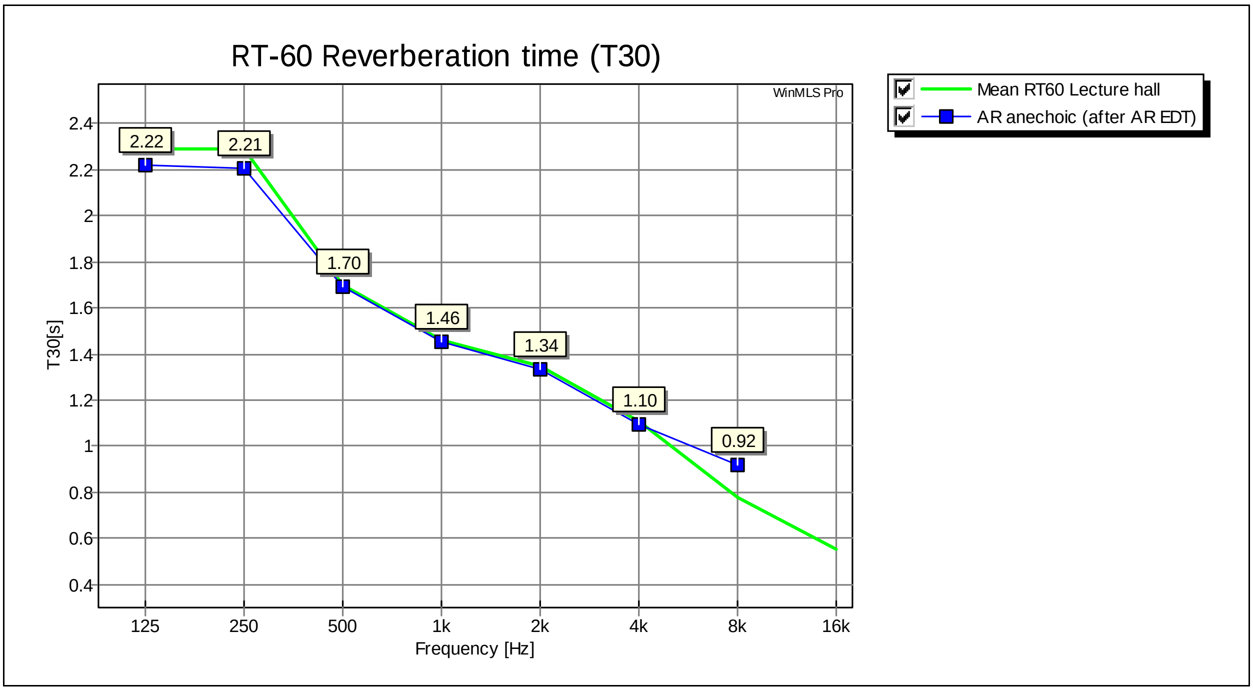


Table 5.10: Summary of results for RT_{60} after EDT was artificially reverberated.

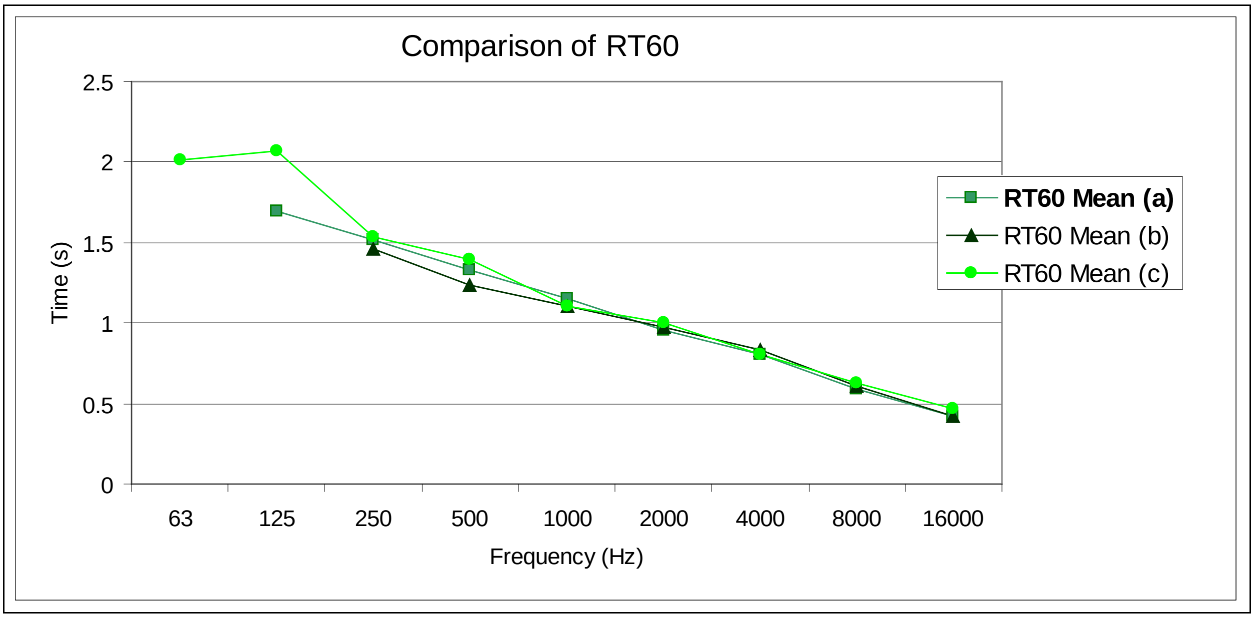
x-axis F [Hz]	**AFTER AR EDT** AR anechoic (after AR EDT)	**GOAL** (Mean RT_{60} – Lecture Hall) (s)	**Percentage deviation (After AR EDT vs. Goal) (%)**
125	2.22	2.29	-3.06
250	2.21	2.29	-3.49
500	1.70	1.70	0.00
1000	1.46	1.46	0.00
2000	1.34	1.35	-0.74
4000	1.10	1.11	-0.90
8000	0.92	0.78	17.95
16000		0.55	---

5.4 Experiment 3: Anechoic Adjusted to Exhibit the Art Theatre's Reverberation Time Response

5.4.1 Method and Results

In the RT analysis section, the arts theatre was split into three sections. The different plots are shown in Figure 5.34 as a reminder and motivation for the decision to choose the *RT_{60}* mean (a) plot for this experiment (shown below). The reason for this choice was that plot "c" was too similar to the lecture hall's RT and would be repetitive. The plot "a" was the first almost completely linear slope, which would be different from the previous two experiments (small room and lecture hall). Plot "b" did not have a value for the 125Hz band and thus too would not be adequate for this test.

Figure 5.34: Comparison of RT_{60} for a, b, and c for the arts theatre.



Using plot "a" as the goal for this experiment and plotting it against the anechoic RT to obtain a relative measure of the two plots follows next.

Figure 5.35: Relative comparison of anechoic and arts theatre's RT_{60} shown on same axes.

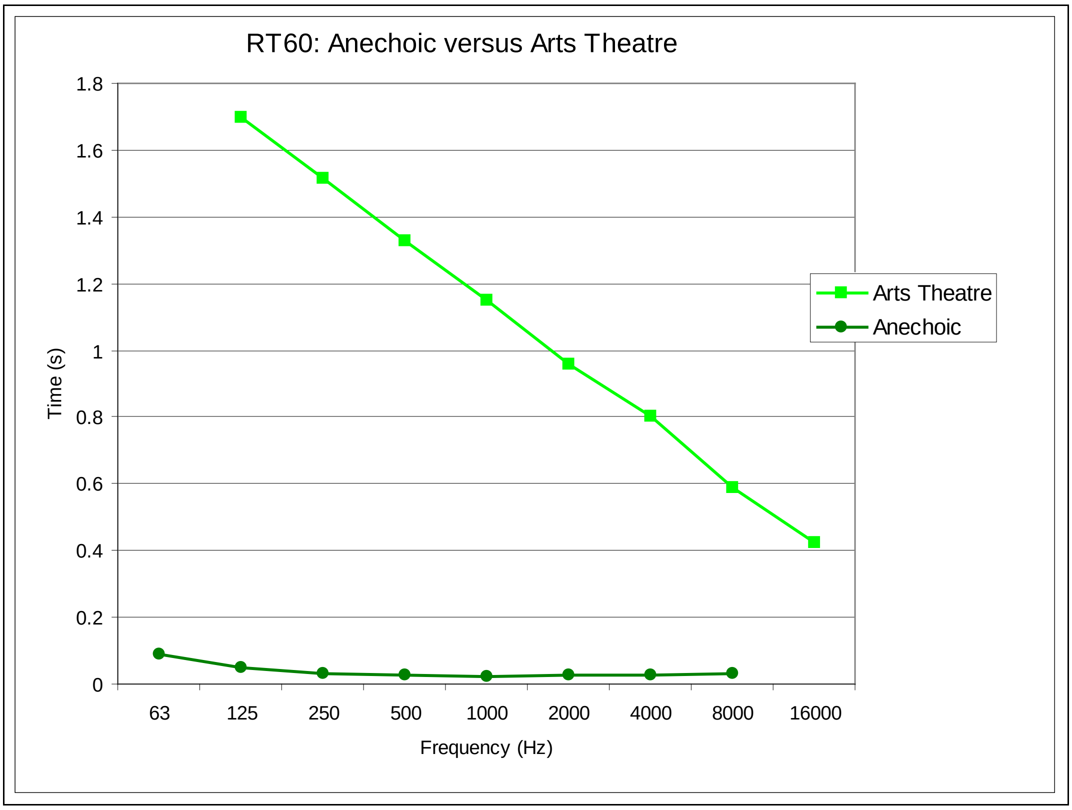
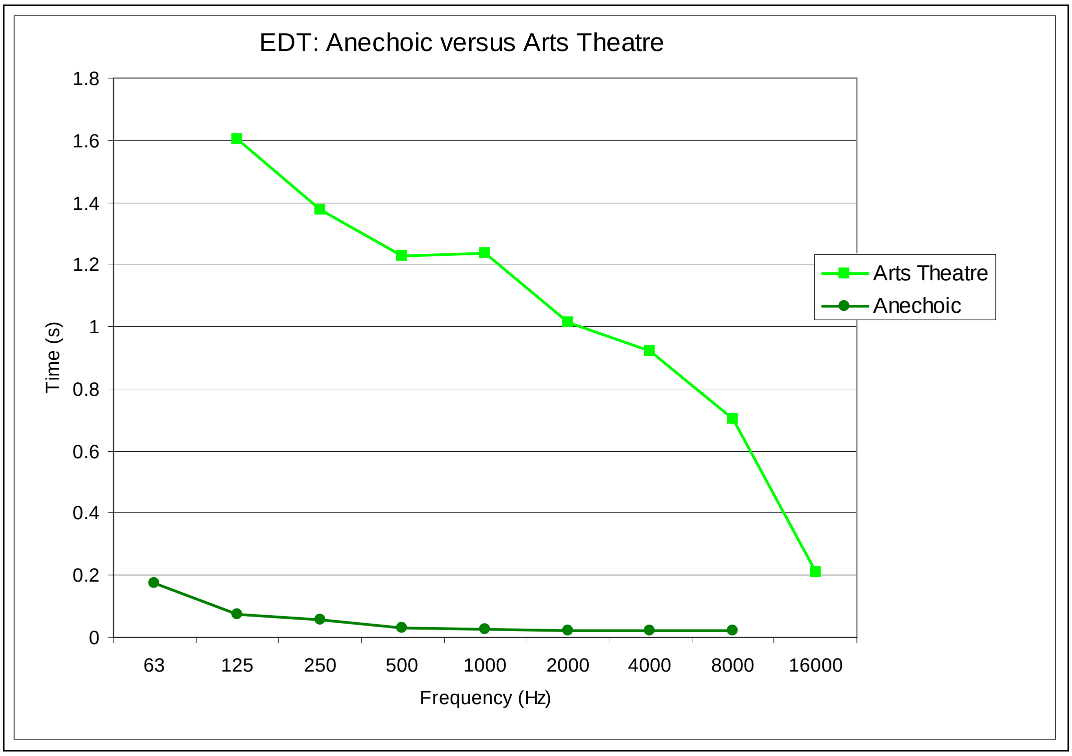
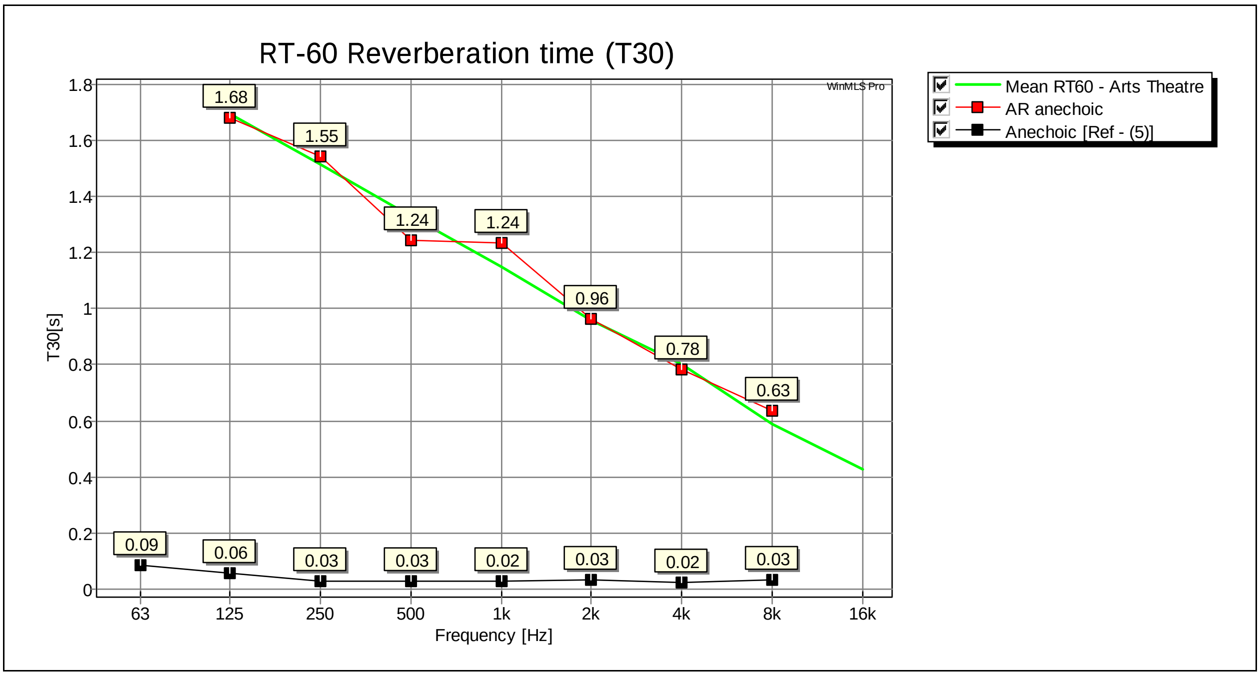


Figure 5.36: Relative comparison of anechoic and art theatre's EDT shown on same axes.



After applying artificial reverberation, the following results were obtained as shown in the next figure and associated table.

Figure 5.37: Before and after AR applied, compared to the art theatre's mean RT_{60}.



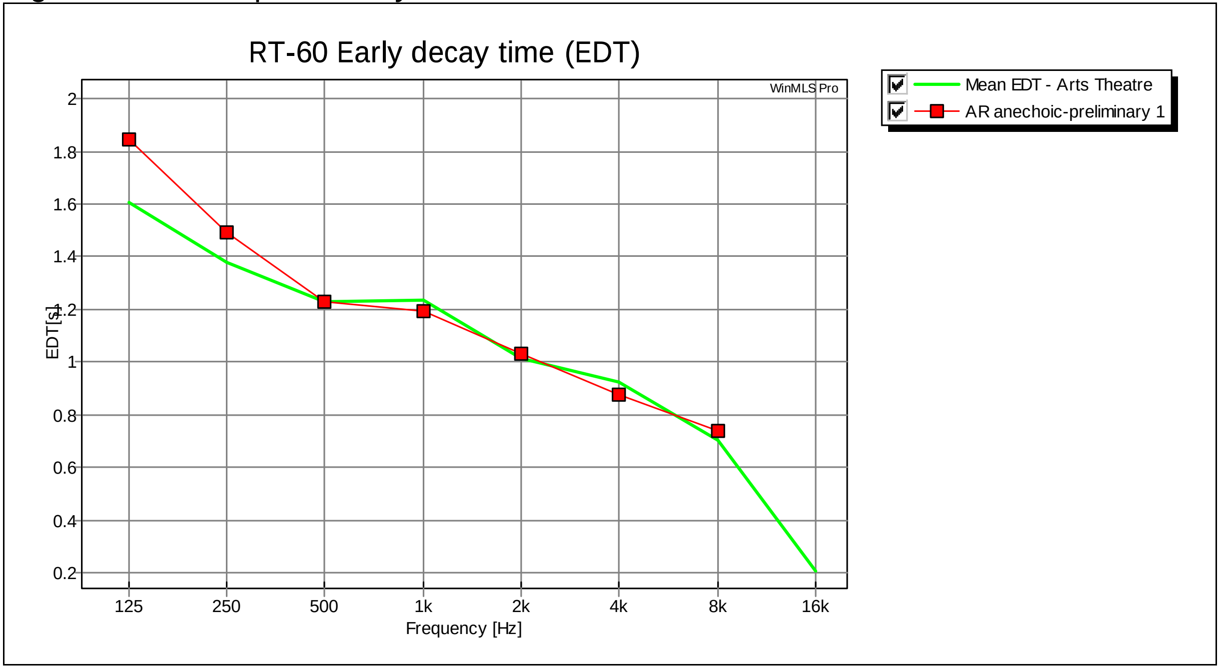
The artificially reverberated RT_{60} shows a larger deviation at points 500Hz and 1kHz from the desired plot.

Table 5.11: Summary of results for RT_{60}: Anechoic converted to art theatre's response.

x-axis F [Hz]	**BEFORE AR** Anechoic (Original) [Ref - (5):RT_{60}] (s)	**AFTER AR** AR anechoic (s)	**GOAL** (Mean RT_{60} – Arts theatre) (s)	**Percentage deviation (After AR vs. Goal) (%)**
63	0.086			
125	0.055	1.68	1.70	-1.18
250	0.028	1.55	1.52	1.97
500	0.025	1.24	1.33	-6.77
1000	0.025	1.24	1.15	7.83
2000	0.030	0.96	0.96	0.00
4000	0.023	0.78	0.80	-2.50
8000	0.031	0.63	0.59	6.78
16000			0.42	

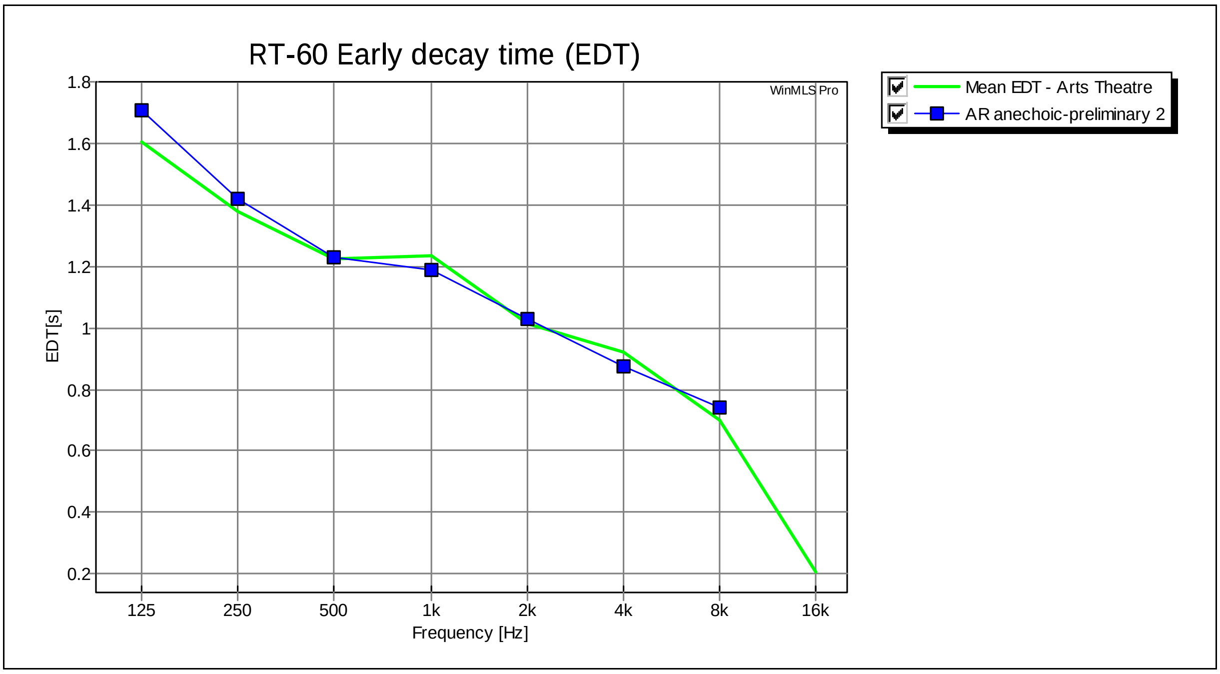
Owing to the fact that the EDT and RT_{60} were derived from the same audio sample, it was challenging to obtain the correct AR for each. Obtaining an acceptable RT_{60} did not mean that the EDT was correct. Thus, this process of applying AR relied on accessing both RT_{60} and EDT for each DSP step applied to the audio sample. The initial result for EDT is shown in Figure 5.38. The lowest octave bands did not meet the goal and thus further work was needed.

Figure 5.38: First preliminary result for arts theatre.



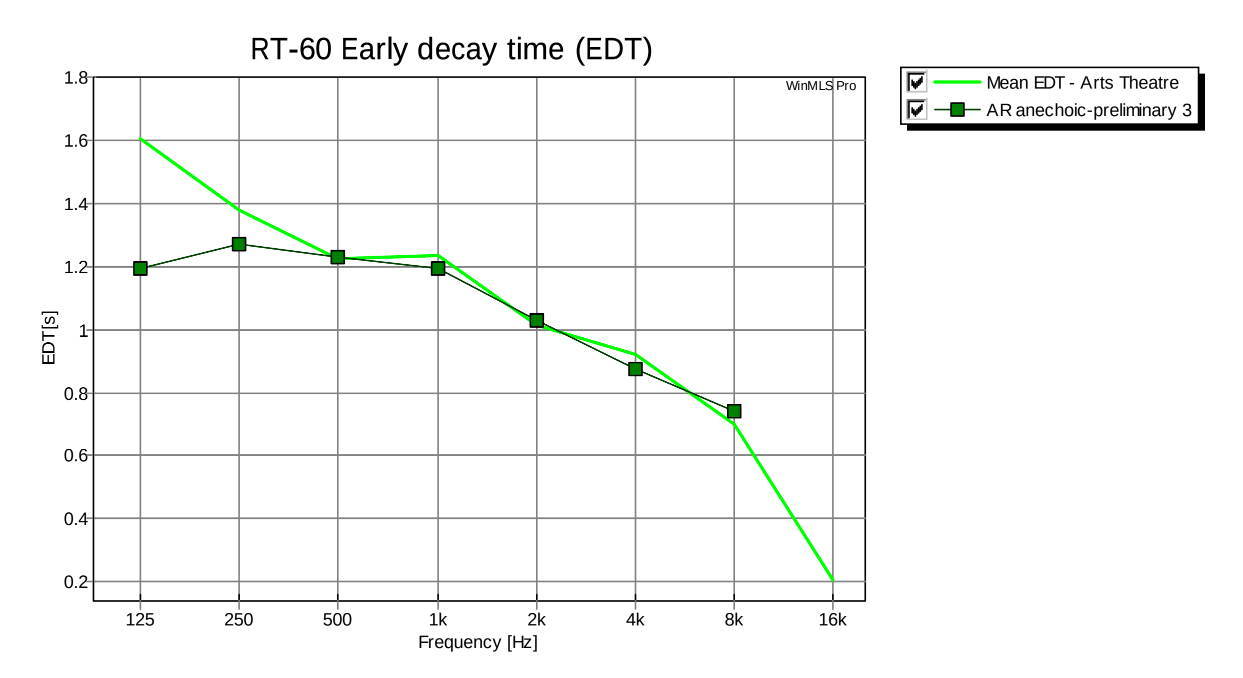
The octave bands of 250Hz and 125Hz have a large deviation (more than 10%). The RTs in the lowest bands were too high and a method of reducing them was required. It was not possible to obtain a close match for the higher octaves and still have the lower octaves matched after applying the AR. Thus, it was decided to accept a deviation in the low octave range and use filtering to reduce the RT for those two bands. A high-pass Bessel 4th order filter having a roll-off slope of ±24db/octave with cut-off frequency of 450Hz was applied. This filter was chosen as it has a smooth roll-off response and a gentle transition into the passband. The cut-off value of 450Hz was selected, as octaves above 500Hz were acceptable and did not need to be changed. The following result was obtained as shown as *AR anechoic-preliminary 2* in Figure 5.39.

Figure 5.39: Second preliminary result for arts theatre.



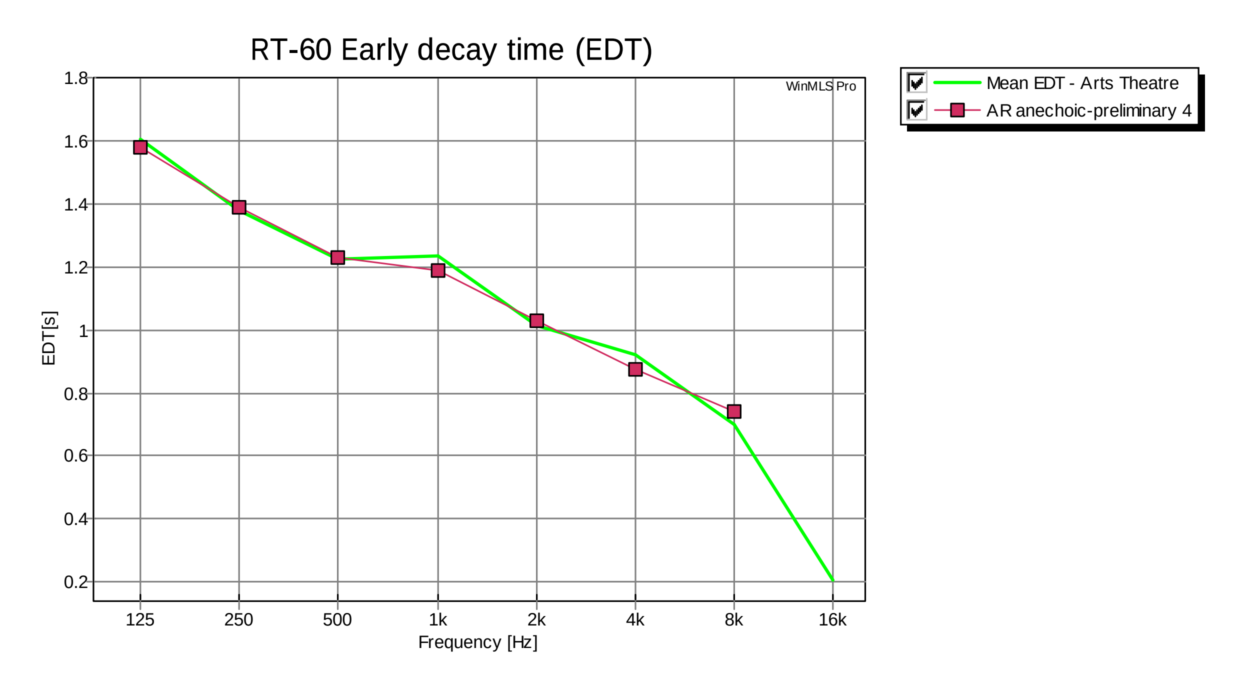
The filter managed to reduce the RT for the two lowest octave bands but the overall response was still shy of being a close match to the desired response. The Bessel filter was reversed (the wave file was put back into the condition prior to the application of the Bessel filter) and a Butterworth filter was applied instead [also high-pass]. The cut-off frequency remained the same but the order was changed to 10^{th} order. This was a drastic change and was chosen to evaluate whether the increase in filter order would have adequate range in its ability to change the EDT. Thus, 10^{th} order was chosen. The finding was that using filters such as Butterworth and Bessel, a change in the EDT was available. In this case with a 10^{th} order Butterworth response, the decay slope was set too steep and from Figure 5.40, it can be seen how the octave bands in question now have an RT for 125Hz and 250Hz bands that is far too low.

Figure 5.40: Third preliminary result for arts theatre.



Applying a 3rd order Butterworth instead resulted in the goal being achieved (Figure 5.41).

Figure 5.41: Fourth preliminary result for arts theatre.



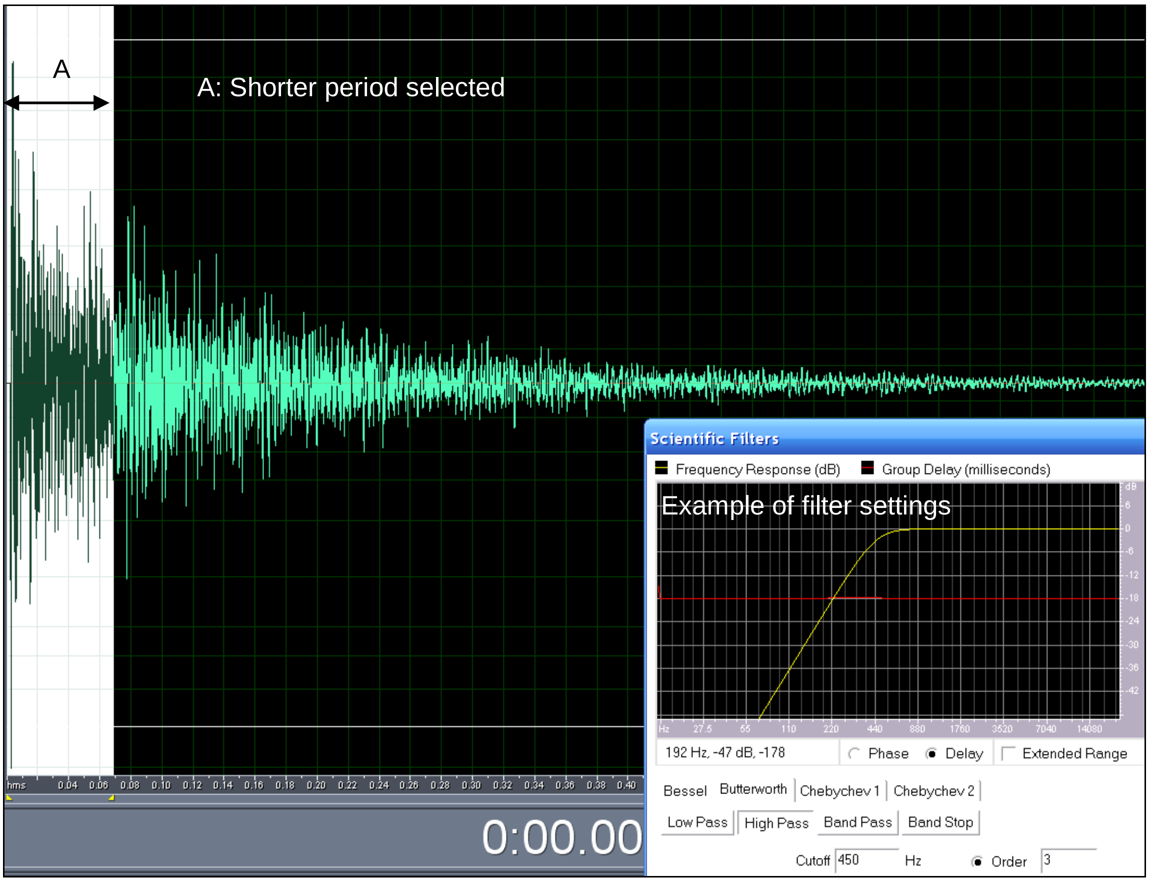
The following table summarises the numerical data for the preliminary measurements.

Table 5.12: Summary of results for all preliminary attempts for EDT.

x-axis	AR anechoic-preliminary 1	AR anechoic-preliminary 2	AR anechoic-preliminary 3	AR anechoic-preliminary 4	**GOAL** (Mean EDT – Arts Theatre) (s)	Percentage deviation (Preliminary 4 vs. Goal)
63	1.78	1.60	1.22	1.40	---	--
125	1.84	1.71	1.20	1.58	1.61	-1.86
250	1.49	1.42	1.27	1.39	1.38	0.72
500	1.23	1.23	1.23	1.23	1.23	0.00
1000	1.19	1.19	1.19	1.19	1.24	-4.03
2000	1.03	1.03	1.03	1.03	1.01	1.98
4000	0.87	0.87	0.87	0.87	0.92	-5.43
8000	0.74	0.74	0.74	0.74	0.70	5.71
16000	---	---	---	---	0.21	---

After obtaining the EDT, the RT_{60} was re-checked to make sure that it was not affected by the filtering. The RT_{60} was affected which highlights the shortcoming of this technique. To counteract this problem, manipulating the amount of the audio sample that is to be filtered reduces the impact on the RT_{60} yet still provides adequate improvement for the EDT. The reason for why this is successful is that the EDT is only a small portion of the full audio sample. Applying filtering to the entire waveform was unnecessary as the EDT only accounts for the audio up to the first 10dB drop in amplitude. Thus, if the first portion of the waveform has the filtering applied, the EDT is improved without much change in the RT_{60} characteristic of the entire waveform. The next figure shows an example of this method.

Figure 5.42: Selecting a shorter period for applying filtering.



5.5 Problems Encountered

As this application relied on software and computer techniques, the problems were mainly pertaining to that subject area.

a) <u>Log Files</u>

Using Cool Edit Pro would have been easier if there was a log file that tracked all the file changes during the AR process. The software does have the ability to undo changes but if all changes could be exported into a text file for reference and event viewing, this would be an ideal situation. This would have resulted in a massive time reduction as noting each signal-processing step was tedious.

b) <u>The Sensitivity of the EDT</u>

As stated earlier, minor adjustments to the available RT settings within Cool Edit Pro resulted in massive changes to the octave specific EDT.

c) <u>Digital Storage Space</u>

Owing to the fact that there were hundreds of wave files generated during this project, storage space became problematic. This is a common problem in recording studios where a large array of hard drives is required to store all the music files. As uncompressed pulse code modulation requires a large storage space, extra disk space was needed, for example, even the memory stick used was 8GB and had a read/write speed of 30MByte/s and 8MByte/s respectively. During the AR process, many files remained open in the Cool Edit Pro program, which resulted in extensive RAM (Random Access Memory) requirements. The laptop had 1GB of RAM on board which was sufficient however the laptop's responsiveness was reduced. The laptop's specification was not great and it was four years old. Thus, it is maintained that the hardware requirements for applying AR were not state-of the art, but one would need sufficient storage space for wave files.

d) *<u>A Matched RT does not Mean the Sounds are the Same</u>*

After manipulating the anechoic impulse to mimic the echoic impulses, they were checked to see if they matched each other as shown in the octave RT graphs that were shown in this chapter. The results were promising. However, on listening to the artificially reverberated impulse and then listening to the desired echoic impulse, they did not actually sound the same. While they sounded quite similar, and some more than others, they were not exact copies. This issue did not pose a problem in this chapter but was a focal point in the subjective testing section (Chapter 6). The obvious finding was that just because the RTs were similar, that did not mean the impulses sounded the same. Other factors such as sound strength, sound balance and IACC amongst other parameters would also need to be adjusted. This problem further motivates the joint objective and subjective co-dependency in the field of acoustics.

5.6 Engineering Methodology

One may argue that the methodology used was not in keeping with strict engineering principles and could not be readily automated for the application of AR. While this chapter did rely on mathematical methods in obtaining AR and evaluating the results, it is unlikely that there will be a fixed one-size fits all method available in the near future for a variety of reasons. Firstly, each audio sample is unique. Secondly, the results while verified mathematically would still need to be verified through listening tests as a final measure of acceptability. Thirdly, the amount of reverberation required is based on personal preference, which too is difficult to quantify. Fourthly, the signal processing requirement would require an agreement on the definitions of the acoustic parameters which are still in debate (Miller, 2007). Fifthly, the designers of the acoustic suites are corporate companies operating in a capitalistic economy, which makes it unlikely for a collaborative effort. Even if audiophiles and engineers did agree on the definitions and technical parameters, there would still be dispute on whether the sound that the device outputs is an example of the settings of the program. This last point may be solved by using MLS software for evaluating the waveforms and proving that the RT is correct as was conducted in this chapter; however, the finding that a matched RT does not mean the sound is the same could be applied and this becomes a circular argument. This is the reason that this dissertation used both a mathematical measure and a listening test as the proof of attaining the goals of the study.

Altogether, this chapter's experiment provided the goal AR six times and was successful for each one. Thus, while not automated, the process was repeatable to some extent.

5.7 Conclusion

An anechoic impulse that differed considerably from three echoic samples was successfully adjusted using AR to match the three echoic impulses in terms of RT_{60} and EDT within a 10% deviation specification. The results showed that it was possible to match a sound that had an RT of less than 100ms, to that of a naturally reverberated RT of at least 1,5s. The major challenge was obtaining a match for both RT_{60} and EDT from a single wave file. Overall, it was possible to achieve the desired goal in all accounts, verified through octave-specific RT graphs.

This chapter relied on mathematical methods in evaluating the artificially reverberated impulses with octave filters and graphical plots provided by WinMLS acoustic analysing software. This has been termed an objective test, as no listening tests were required. However, in practice recording engineers and musicians would use listening tests as a measure of whether they are getting the sound they want. This method is subjective and arguable and thus by incorporating a mathematical evaluation of the end condition as performed in this chapter, the validity of this experiment is improved in terms of the engineering methodology.

The results obtained from this experiment are important to the recording engineer as the recording engineer and musician rely on their environment to obtain their desired level of RT. With the use of AR, the recording engineer is given further range and freedom in his/her task of getting the reverberation correct. The reliance on scarce large-scale lively rooms can be substituted by using AR and moderately echoic rooms. In this study, an almost anechoic sample was adjusted as an extreme example. In practice, a moderately echoic room may be used and AR applied. The process is quicker as there is already some reverberation present and only further reverberation would be required. It was also shown that excess RT was removed in the low frequency bands; however, that was not the focus of the study.

The improvement to the field is in the reduced time required for the recording engineer to get the recording environment correct as this can be compensated largely by use of software methods and AR. Further, the cost and complexity required for AR need not be excessive. The results obtained from this chapter show that even modest recording equipment could provide the desired goals[42]. The use of readily available software too improves the accessibility of this method.

[42] The acoustic analysing software bears the highest cost and is only required for the proof of this experiment. The musician/recording engineer would not need the WinMLS software as part of their recording setup.

The next chapter introduces the subjective side of the application of this dissertation, which seeks to prove that two groups of people would not be able to identify the artificially reverberated sounds from the echoic sounds that had natural reverberation.

Chapter

6

6 ARTIFICIAL REVERBERATION APPLICATION: SUBJECTIVE TEST

6.1 Introduction

This section introduces a listening component as part of the testing. As hearing is a subjective experience, it was decided to deal with it in its own chapter as it differs considerably from the previous section. In this chapter, a vocal passage recorded in the anechoic environment was subjected to AR and then compared to the vocal passage recorded in each of the three echoic rooms (small room, lecture hall and arts theatre). The vocals were evaluated by two groups through a questionnaire and the results were then statistically processed.

6.2 Procedure Overview: Subjective Test

The same environments were used as in the objective tests; however, instead of an impulse sound, a vocal passage was used. A female vocalist[43] sung the same verse in each of the four locations numerous times. The microphone used was the *Shure* SM58 vocal microphone[44]. The cables were all balanced and of XLR type. The recordings were digitally converted using the same soundcard laptop configuration as in the impulse testing in the previous chapter. Thus, the recordings were conducted at a minimum of 48kHz [some were recorded at 96kHz] with a bit depth of 16. The distance between microphone and singer were controlled as this also affects the RT of the recording (Nisbett, 1995). The SPLs of the recordings were monitored and a few different SPLs were used for each recording. As this was a subjective test, the majority of decisions were made after hearing the output rather than by mathematical methods. For example, in the first recording location a loud SPL was obtained easily as the room was small. To obtain the same SPL in the arts theatre the singer would need to project her voice, which has the affect of possible pitch changes as well as sounding different in terms of the flow of the vocals. Thus, some amplification settings were adjusted for these recordings. No other signal processing was applied to the normally reverberated

[43] A professional singer sang the vocal passages. The details of the singer and the other purposively sampled participants used in this test are included in Appendix D.

[44] There are recording microphones that are better suited for studio recording; however, the SM58 is an accepted performance microphone and was in the economic range of this project.

vocal samples and if any amplification was applied, it was applied for the entire frequency range.

The song chosen for this test was Evanescence’s (2003) *Bring me to life* from the album *Fallen*. This track was chosen as it has verses that require the singer to sing loudly followed by softer sections and then loudly again. This allows the listener to hear both the EDT as well as the longer RT decay. The loud section requires the singer to project her voice and remain in the same key for a few seconds.

The tracks were filed and labelled according to their location of where they were recorded. Each location had a few recordings completed. It is highlighted that each successive recording was completed in the same position in the rooms and with the same room layout, that is, the successive recordings were duplicated as closely as possible. This is a critical point as recording the singer in a different position within the same room could translate to a different reverberant sound. The volume of the recording was calibrated and thereafter a few recordings were conducted with the best two recordings chosen as the track samples for each echoic area. The anechoic recordings were recorded in the same way and labelled accordingly.

The lecture hall had an additional recording type: the singer was asked to hold the microphone 40cm away from her mouth in order to create a larger RT. A few recordings were also conducted for this distance, again with attention paid to the reproducibility of the position, volume, pitch and projection. The two most similar samples were used as an example of excessive reverberation for this room. Thus the lecture hall was used twice in the subjective listening test, firstly for a standard recording and then secondly for a highly reverberated recording example.

After all the recordings had been completed and filed accordingly, the anechoic recordings were analysed and narrowed down to the best four samples. A process of artificially reverberating the anechoic samples followed. This relied on various reverberation parameter adjustments; including changes in delay, filtering of select frequency groups, amplification and attenuation adjustments, and graphic equalisation. Careful observation of the headroom levels were conducted by use of the dB level indicators and the full digital dB scale. This was a lengthy process and relied on repetitive listening tests, which is similar to that used in the mastering process of recorded music prior to the completion of the final copy. In the case of this study, the editing process was extensive owing to the nature of the original sound being almost anechoic and the requirement of it matching a pre-defined reverberant sound. This would not be the case in the recording studio whereby the live sound would have at least some reverberation present and would not need to be matched, but rather made to sound lively, but not necessarily a copy of a reverberant sound that had already been recorded. Thus, to manipulate an anechoic sample several times to sound similar to each individual echoic sample required a comparative sound analysis. The anechoic sound was constantly compared to the echoic sample for each room until a close match was obtained in both reverberation and colouration. Room dimensions and layout not only adjust the singer’s reverberation but the complete graphical mix of the sound. This is discussed further under the limitations heading.

The anechoic sample was manipulated to sound like the small room, the lecture hall, the arts theatre and then the lecture hall's excessive reverberation sample. Thus, in total four separate artificially reverberated anechoic samples were created. These samples were then placed into a track that contained two of the live recordings for each echoic location, for example, one track contained the anechoic sample that was artificially reverberated to sound like the small room as well as two sample recordings from the small room. The next track contained the artificially reverberated anechoic sample that was manipulated to sound like the lecture hall as well as the two best samples that were recorded live in the lecture hall and so forth. Thus, one complete track contained three vocal samples with one of them being the artificially reverberated sample. An audio CD was compiled as well as an accompanying questionnaire that was used in the listening test. The results were statistically processed. A copy of the testing CD is included and is enclosed on the last page in a CD pocket.

6.3 The Reverberation Listening Test: Questionnaire and Audio CD

6.3.1 Aim of questionnaire

To determine the extent to which two population samples could differentiate between an artificially reverberated vocal sample from that of its two live vocal counterparts that had no AR applied to them. This questionnaire provides a method for subjective testing of the effectiveness of software based reverberation techniques applied for the use of AR.

6.3.1.1 Hypotheses

6.3.1.1.1 <u>Hypothesis 1</u>

The mean score obtained for the first reverberation test by the general group (group 1) will be lower than the "musically skilled" group (group 2).

H_o: $\mu_1 = \mu_2$

H_1: $\mu_1 < \mu_2$

μ_1 = the mean of the Reverberation Test 1 score for population group 1.

μ_2 = the mean of the Reverberation Test 1 score for population group 2.

(α=0.01)

6.3.1.1.2 <u>Hypothesis 2</u>

The participants score in their first reverberation test would not be matched in their second reverberation test irrespective of population and qualifying characteristic.

H_o: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

μ_1 = the mean of the Reverberation Test 1 score.

μ_2 = the mean of the Reverberation Test 2 score.

(α=0.01)

6.3.1.1.3 Hypothesis 3

The participants love for music; time spent listening to music; ability to play and appreciate music as well as their critical listening aptitude all would not show a correlation to their ability to distinguish the artificially reverberated vocal samples from the normally reverberated samples. Thus, the participants score in the second reverberation test would not exhibit a correlation to the above stated independent variables.

H_o: $\rho=1$

H_1: $\rho<1$

($\alpha=0.01$)

There were four individual correlations conducted.

ρ_1 = the total population correlation between their love for music and their score in the second reverberation test.

ρ_2 = the total population correlation between their time spent listening to music and their score in the second reverberation test.

ρ_3 = the total population correlation between their ability to play and appreciate music and their score in the second reverberation test.

ρ_4 = the total population correlation between their critical listening aptitude and their score in the second reverberation test.

6.3.2 Methodology

6.3.2.1 Participants

Two sample groups were chosen. The first group was a convenience sample of 65 second and fourth year electrical engineering students with an average age of 24, and was called the "general" group. The second group was purposively sampled and consisted of six participants. The purposively sampled group of participants were chosen based on each participant's abilities and knowledge of acoustics and music and has been termed the "musically skilled" group. A brief description of each of the six musically skilled participants can be found in Appendix D. Both sample groups were given the same test. The test was explained to the participants and then handed in afterwards.

6.3.2.2 Informed Consent and Confidentiality

Participants were verbally briefed as to the goals of the research experiment. It was made clear that participation was voluntary. A preface to the questionnaire also stated the voluntary nature of participation. Once the form was completed and handed in, they were handled and stored by the researcher alone.

6.3.2.3 Tools

6.3.2.3.1 Acoustic Listening Skill Test on Reverberation

There was no available test that could be used for this subjective evaluation and thus the researcher compiled his own test. The test consisted of 12 questions. The questionnaire was divided into two parts. Part 1 consisted of four questions that were self-describing forced choice statements that the participants chose as a best match to describe themselves. These questions were ordered in a manner that would highlight blind guessing. Part 2 contained two reverberation listening tests. The first test consisted of four questions while the second test consisted of eight questions.

The first reverberation test required the participant to listen to two consecutive sample vocal tracks and choose the most reverberant track of the two. The tracks were extracted from the accompanying audio CD from the book *Critical Listening Skills for Audio Professionals* (Everest, 2007). Everest recorded a spoken verse that was repeated with varying amounts of reverberation in his tutorial on RT. The range varied from 0,5s to 2,5s of reverberation in steps of 0,5s. The first four questions asked the participants to choose the most reverberant sample out of the two sample tracks played. These four questions were qualifying questions and served as a pre-evaluation of whether the participants could differentiate between reverberant and less reverberant sounds. The second reverberation test followed with eight questions, which asked the participant to listen to three successive vocal samples (± 20 seconds each) and choose the one that had a different reverberant sound. The samples were the recordings of the female vocalist who sang in all the experimental locations studied in Chapter 6 (small room, lecture hall and arts theatre). The point of the questionnaire was to determine if the participant could hear the difference between the tracks recorded in the echoic rooms from that of the sample recorded in the dead environment, which was artificially reverberated. Thus, the first question [from the second reverberation test] consisted of an accompanying audio track that contained two recordings from the small room and one from the anechoic area that was artificially reverberated. The next question consisted of two recordings from the lecture hall and one from the anechoic area that was artificially reverberated and the same for the arts theatre. Each location was tested in its own question with its own associated audio track; however, the participants needed to answer it twice before it was marked correct to account for chance. For example, question five and six were of the same audio samples but re-ordered to reduce the effect of blind guessing. The same logic applied for question seven and eight, nine and ten as well as eleven and twelve. To sum up, there were two reverberation tests. The first was a qualifying tests and the second test contained the artificially reverberated samples and their associated live counterparts.

The questionnaire took an average of 20 minutes to complete including the explanation from the researcher [and the additional explanation audio track]. The participants were allowed to listen to the samples twice. The questionnaire can be found in Appendix D.

6.3.2.3.2 Test Validity

The test comprised of different vocal samples. The first reverberation test RT samples were taken from Everest (2007) and were verified for their stated specification. As reverberation in terms of this test related to vocals, the testing variable was vocally based with only the amount of reverberation being the varying quantity. The language used in the test was of a low level and any acoustic terms were explained and demonstrated to the participants. Thus, the first reverberation test is assumed to have construct validity. The conveniently sampled group were all students at a university. This is not believed to be a threat to the validity as the social spread of the group was wide in terms of life experience, age, family size, race, religion and job description (those who were working and studied part-time). The internal validity was found to be high as the effects of confounding, selection bias, maturation, mortality, instrument change, and attrition were not related to this questionnaire. The external validity was assumed to be high as well. The testing locations were chosen for their low background noise and adequate acoustic response. The test was conducted using high-quality equipment. The conclusions drawn in terms of this test have not been generalised to other forms of reverberation and thus the conclusions are based on the same sound type, that is, reverberation applied to vocals.

6.3.2.4 Testing equipment and Environment

Equipment used for the listening test is shown in Table 6.1 below.

Table 6.1: Equipment used in testing.

Function	Make	Model	Comment
Amplifier	NAD	3240PE	
CD player	Sansui	CD-α607	
Speakers	Bose	302 (2001 model)	
Headphones	AKG	271	Group 2 participant request
Headphones	Sony	MDRV-700	Group 2 participant request
Cabling	None stated		Low resistance copper RCA and speaker cable

The listening tests for the first sample group were conducted in two classrooms. The rooms did exhibit some reverberation and attention was focussed to finding a balance between playing the test CD at a loud enough volume but not so loud that the room’s reverberation would dominate. When listening to a play back of a musical performance that was recorded with reverberation, more reverberation may be heard if the playback room too adds reverberation. This may be problematic especially if the RT of the playback room is long and the resulting sound heard may be confusing[45] (Berg & Stork, 1982:212). An attempt was made to reduce this effect and another less reverberant class was used when testing the remainder of the sample group, however the results were similar. Using headphones instead of monitors may improve this problem. Group 2’s participants were tested in their own locations and thus the rooms differed from participant to participant, however all six listened via headphones and thus the room effect was minimal.

[45] A practical example of this problem is as follows: A person in a room that has a long RT (or poor direct-to-reverberant level) announces using a paging system. The paging announcements are played in a room that also has a long RT and the result is reduced speech intelligibility and often sounds like the mid range frequencies are only present.

6.3.2.5 Data Analysis

All the data was manually entered into Microsoft Excel. The participants' responses to the four self-describing questions and their available answers were tabulated with their respective percentage choice values allocated to their relevant groups. A spread of the self-describing choices and their relative incidence was obtained. The results for the two reverberation tests were obtained and converted into a percentage. Individual percentage scores as well as group means were obtained for the two tests. A t-test that ascertained the difference between the means of two independent groups was applied to the two sample groups with respect to their first reverberation test scores. A second t-test was conducted to compare the first and second reverberation test scores for each group. This was a dependent groups test as the same participant had two tests and was matched for first and second reverberation test. The dependent t-test was used for all participants and then used for the "qualified" participants. Four Pearson's *r* correlation tests were used to examine the subjective evaluations that the participants chose in their self-describing questions versus the participants' second reverberation test scores.

6.3.2.6 Anticipated Problems

The singer sang each verse individually in each location, thus there were slight differences in the vocal passage. The participant may hear the slight difference in timing and the occasional key change and believe this to be a sign of a different recording venue. The researcher has explained to the participants that there were slight differences in the track samples and that they should focus on the reverberation rather. Unfortunately, listening is not driven by logic and thus unconscious factors are at play so this issue remained a problem.

There may be participants who had hearing problems. As the first group was tested in sample groups of at least 10 persons, it is possible that a participant had a hearing problem but did the test anyway without telling the researcher. The reason for this is that some people while not completely hard-of-hearing may have a percentage hearing loss yet continue to function without any hearing assistance. These people may not feel the need to disclose their hearing weakness. The researcher did not state any limitations with regard to this issue, as he did not want to isolate anyone especially since the sample was conveniently chosen.

6.3.3 Results

6.3.3.1 Self-Describing Question

There were four self-describing questions. The results are shown in percentage form in the next table. The groups were calculated individually and then combined. The percentages indicate the amount of people who chose their relevant option. There were 65 participants in the first group and six participants in the second group. The table shows how different the two groups were in their chosen answers. This was expected as the second group was purposively sampled and consisted of participants who are active in the music field. For example, more than 80% of the participants in the second group selected the option that they loved music and could not live without it compared to just under 35% for the first group. Group 2 also had a much higher daily time average for listening to music, which was 67% compared to 28% for the first group, as well as about 67% playing or have played an instrument compared with only 17% for the first group.

Table 6.2: Summary of results for the self-describing questions[46].

Self-describing question 1: Which of the following phrases best describes your view of music?		
Available answers[47]:	Group 1 (%)	Group 2 (%)
I love it and can't live without it.	34	83
I enjoy it.	62	17
I could live without it.	3	0
I am indifferent.	2	0
I don't like it.	0	0

Self-describing question 2: How much time on average do you spend listening to music per day?		
Available answers:	Group 1 (%)	Group 2 (%)
Less than 25minutes.	5	0
More than 25 minutes but less than 1 hour.	18	0
More than 1 hour but less than 1,5 hours.	28	67
More than 1,5 hours but less than 3 hours.	25	17
More than 3 hours.	25	17

Self-describing question 3: Which of the following phrases best describes your musical ability?		
Available answers:	Group 1 (%)	Group 2 (%)
I play or used to play a musical instrument.	11	17
I play a musical instrument/perform and I consider myself a professional musician/performer.	6	50
I don't play any instrument, but can appreciate it when someone plays one well.	72	33
I like music but am no good at evaluating whether someone is good at music.	8	0
I have no musical ability.	3	0

Self-describing question 4: Which of the following phrases describes your music listening ability? A good ability would mean that you are able to hear minor changes in sounds while a weak ability means that you are not aware of minor sound changes.		
Available answers:	Group 1 (%)	Group 2 (%)
I have been told that I have a good ear for music.	14	0
I have been told that I have a good ear for music and I have told myself that as well.	31	83
I am useless at listening.	5	0
I am average at listening.	51	17

[46] Owing to the fact that there cannot be fractional people, the results were rounded to whole numbers. This resulted in slight inaccuracy. This inaccuracy was accepted for the gain of improved readability.
[47] The options for all questions are not in a logical sequence to reduce blind guessing. For example, the 5th option of self-describing question 1 does not correlate with the 5th option of self-describing question 2 etc.

There were two reverberation tests. The first test was a qualifying test, while the second test consisted of the anechoic samples that were artificially reverberated to mimic the sound of the three respective echoic locations. The following tables show the results including a surprising feature.

Table 6.3: Results for Reverberation Test 1 (Qualifying test).

Reverberation Test 1: Group spread	**Participants spread**			
Score obtained in test 1 (%)	Group 1 (number of people)	Group 2 (number of people)	Combined (number of people)	Percentage of total participants
0	4	0	4	6
25	18	0	18	25
50	24	2	26	37
75	18	4	22	31
100	1	0	1	1
Total participants:	65	6	71	

The results for the first test show that the two groups differed considerably from each other. Group 1 had 18 out of 65 (28%) participants obtain 75% for the test while all but two of the participants for group 2 obtained the same result. This was expected as group 2 was supposed to exhibit better results as they were the “musically skilled” group. A few interesting observations were noted. Firstly, the amount of people in group 1 who scored less than 75% for the test was surprising. Twenty-four people obtained a 50% mark, which meant that they could not differentiate between reverberant and less reverberant vocal samples beyond reasonable doubt. Only one person was able to get all his/her answers correct and score 100%. The results for the first test were surprising and raise some interesting questions about people’s ability to hear reverberant from less-reverberant vocal sounds. This has been addressed further in the final chapter. Further testing in this regard can be used to improve the validity of the test questionnaire and method. Secondly, the “musically skilled” group had two participants who got 50%, while the rest obtained 75%. This was also unexpected as it was thought that they would all get at least 75%. None of the “musically skilled” participants obtained 100% for their pre-test.

To determine the statistical difference between the means of two independent groups, a t-test was conducted. The results are shown in the following table.

Table 6.4: T-test: Differences between the means of two independent groups for Reverberation Test 1.

Differences between the means of two independent groups.	***Group 1: Reverberation Test 1***	***Group 2: Reverberation Test 1***
Mean score (%)	47.69	66.67
Variance	541.47	166.67
Participants in each group	65	6
Results		

t-Stat	-3.158
P(T ≤ t) one-tail	0.0067

It was hypothesised that the “musically skilled” group would score higher in their pre-test (first reverberation test). The results show a statistically significant difference in the mean scores and thus the first hypothesis was confirmed.

6.3.3.2 Comparing the Mean Scores for the First and Second Reverberation Tests

As there were two reverberation tests, a paired two-sample t-test was conducted to evaluate the statistical relationship between the first test score with its partner second test score. This was conducted for groups 1 and 2. The results are as follows:

Table 6.5: t-Test: Paired Two-Sample for Means – All Participants.

Group 1 (all participants)	***Reverberation Test 1***	***Reverberation Test 2***
Mean	47.69	9.23
Variance	541.47	186.90
Observations	65.00	65.00
	Results	
t Stat	12.59	
P(T<=t) two-tail	0.0000	
DF	64.00	
Group 2 (all participants)	***Reverberation Test 1***	***Reverberation Test 2***
Mean	66.67	4.17
Variance	166.67	104.17
Observations	6.00	6.00
	Results	
t Stat	11.18	
P(T<=t) two-tail	0.0001	
DF (small sample size thus DF is low)	5.00	

The results indicate that the second hypothesis was realised and that the first reverberation test mean score differs from the mean of the second reverberation test score by a statistically significant amount.

6.3.3.2.1 <u>The Effect of Applying a Qualifying Criteria to the Participants</u>

The first reverberation test was now set to act as a qualifying test. Participants who obtained at least 75% for this first test had their second test mark counted. The results are summarised in Table 6.6. Group 1 had only 19 people score 75% and above in their qualifying test, while the second group had 4 out of the initial 6 participants obtain the same result. The total “qualified” participants amounted to 23 when added together. None of the “qualified” participants scored above 25% for their second reverberation test. This was a pleasing result for two reasons. Firstly, none of the participants were able to distinguish beyond reasonable doubt (score of 50% or more) artificially reverberated samples from their normally reverberated sample groups. Secondly, the skill level of the participant [in terms of their musical ability] did not improve the result, as the “musically skilled” group had no participants achieve a score higher than 25%. This was further highlighted in the correlation test, which follows after this section.

The highest score was 25%, which meant that nobody was able to distinguish the artificially reverberated vocal sample from their naturally reverberated group beyond reasonable doubt. This confirmed the second hypothesis fully as well as the main research goal for this chapter.

Table 6.6: Results for Reverberation Test 2 (Musically skilled participants).

Reverberation Test 2: Group spread (Only those who qualified from test 1 with mark ≥75%)	**Participant spread**			
Score obtained in test 2 (%)	Group 1 (number of people)	Group 2 (number of people)	Combined (number of people)	Percentage of total qualified participants
0	10	3	13	57
25	9	1	10	43
50 and above	0	0	0	0
Total participants:	19	4	23	100%

Evaluating the data using the t-test exhibits the following results:

Table 6.7: t-Test: Paired Two-Sample for Means – qualified participants.

Group 1 (qualified only)	***Reverberation Test 1***	***Reverberation Test 2***
Mean	76.32	11.84
Variance	32.89	164.47
Observations	19.00	19.00
	Results	
t Stat	18.52	
P(T<=t) two-tail	0.0000	
DF	18.00	
Group 2 (qualified only)	***Reverberation Test 1***	***Reverberation Test 2***
Mean	75.00	6.25
Variance	0.00	156.25
Observations	4.00	4.00
	Results	
t Stat	11.00	
P(T<=t) two-tail	0.0016	
DF (small sample size thus DF is low)	3.00	

The research hypotheses were confirmed for the qualified groups.

6.3.3.3 Correlation Analysis

Owing to this being a subjective test, it was decided to correlate the subjective opinions of the participants with their second reverberation test scores. The four correlations were based on the four self-describing questions, which were correlated to the scores on Reverberation Test 2.

The correlation tests are linked to the third hypothesis, and are linked to the self-describing questions in the first part of the questionnaire and include both groups. As the independent variables were not on an interval measure but an ordinal type, the scale was converted and represented by numerical units. For example, with the first correlation test, which was the correlation of love for music with second reverberation test score, the love for music was scale rated from one to five with five being absolute love for music, thus a score of 1 equated to dislike for music. The same applied for the other three independent variables. The dependent variable, which was the second reverberation test score was already an interval measure.

6.3.3.3.1 <u>First Correlation Test: Love for music versus second reverberation test score</u>

The following table provides the result for this first correlation.

Table 6.8: Correlation data for first self-evaluative question versus second reverberation test score.

Correlation data	Love for music	Mark for Reverberation Test 2
Love for music	1	
Mark for Reverberation Test 2	-0.1364 (*r* - value)	1

The *r*-value was divided by two as the test was directional, thus:

$r = 0.682$

To obtain the p-value a derived t-test statistic was used. The correlation coefficient *r* can be converted into a statistic that has *t* distribution and N-2 degrees of freedom (Coetzee, Janeke, Terre'Blanche, Kruger & Payze, 2001:134). The following equation was applied:

$$t_r = \frac{r\sqrt{N-2}}{1-r^2} \qquad [6.1]$$

Where:

- t_r is the test statistic used to obtain the probability value
- r is the correlation value of the sample
- N is the sample size

Thus, t_r =10.59 which provides a *p*-value of 0.0000 and hence the hypothesis was confirmed.

6.3.3.3.2 *Second Correlation Test: Time spent listening to music versus second reverberation test score*

Table 6.9: Correlation data for second self-evaluative question versus second reverberation test score.

Correlation data	Time spent listening to music	Mark for Reverberation Test 2
Time spent listening to music	1	
Mark for Reverberation Test 2	-0.1062	1

The *r*-value was again divided by two as the test was directional.

$r = 0.0531$

The statistical probability does not need to be calculated as the hypothesis was confirmed already as r =0,0.

6.3.3.3.3 *Third Correlation Test: Musical ability versus second reverberation test score*

Table 6.10: Correlation data for third self-evaluative question versus second reverberation test score.

Correlation data	Musical ability	Mark for Reverberation Test 2
Musical ability	1	
Mark for Reverberation Test 2	0.1767	1

The *r*-value was divided by two as the test was directional.

$r = 0.088$

The statistical probability does not need to be calculated as the hypothesis was confirmed already as r =0,0.

6.3.3.3.4 *Fourth Correlation Test: Critical listening ability versus second reverberation test score*

Table 6.11: Correlation data for fourth self-evaluative question versus second reverberation test score

Correlation data	Critical listening ability	Mark for Reverberation Test 2
Critical listening ability	1	
Mark for Reverberation Test 2	0.0346	1

The *r*-value was divided by two as the test was directional.

$r = 0.017$

The statistical probability does not need to be calculated as the hypothesis was confirmed already as r = 0,0.

6.3.4 Discussion

This subjective test highlighted that the two groups did differ from each other in terms of their ability to differentiate reverberant from less reverberant vocal sounds. The "musically skilled" group scored a significantly higher average mean score, which was as expected. The difference in the mean scores was evaluated using a t-test, which confirmed that the purposively selected group differed from the conveniently sampled group in terms of their pre-test result. However, these results were not transferred to the second reverberation test. Actually, the mean score of the "general" group was just above 9% while the "musically skilled" group was just over 4%. Although there was a difference, the difference was irrelevant as the average was too low for any positive inferences to be made in terms of the group's ability to handle the second reverberation test. The two groups scored miserably low on the second test, which was a positive result in terms of this study as it confirms that the artificially reverberated vocal samples could not be differentiated from their sample tracks. One may argue that the participants should be assessed before hand to ascertain their ability in critically listening to reverberation. This point is important and thus the researcher decided to initiate qualifying criteria to the sample groups in a further statistical test. Only participants that scored at least 75% had their second test mark counted. This would mean that they were able to critically assess the reverberant from the less reverberant vocals in the first reverberation test and thus should be able to do that in the second test. The range of reverberation that the participants were expected to differentiate were increments of 0.5s. Thus, the participants who passed this first test could differentiate a RT of 1.5s from that of 2s, or that of 0.5s to that of 1s or any combination between the range of 0.5s to 2.5s.

Applying the qualification criteria resulted in a similar result. The "general" group achieved an average score of just less than 12% while the "musically skilled" group scored just over 6%. While both groups had a slight increase in their mean test results, this was still a very low score.

It was interesting to look at the correlational data and see that there were no correlations in terms of the subjective evaluations of the participants and their respective scores in the second reverberation test.

The questionnaire focused on subjective aspects including attitudes and beliefs. The listening test forms the subjective component of this application.

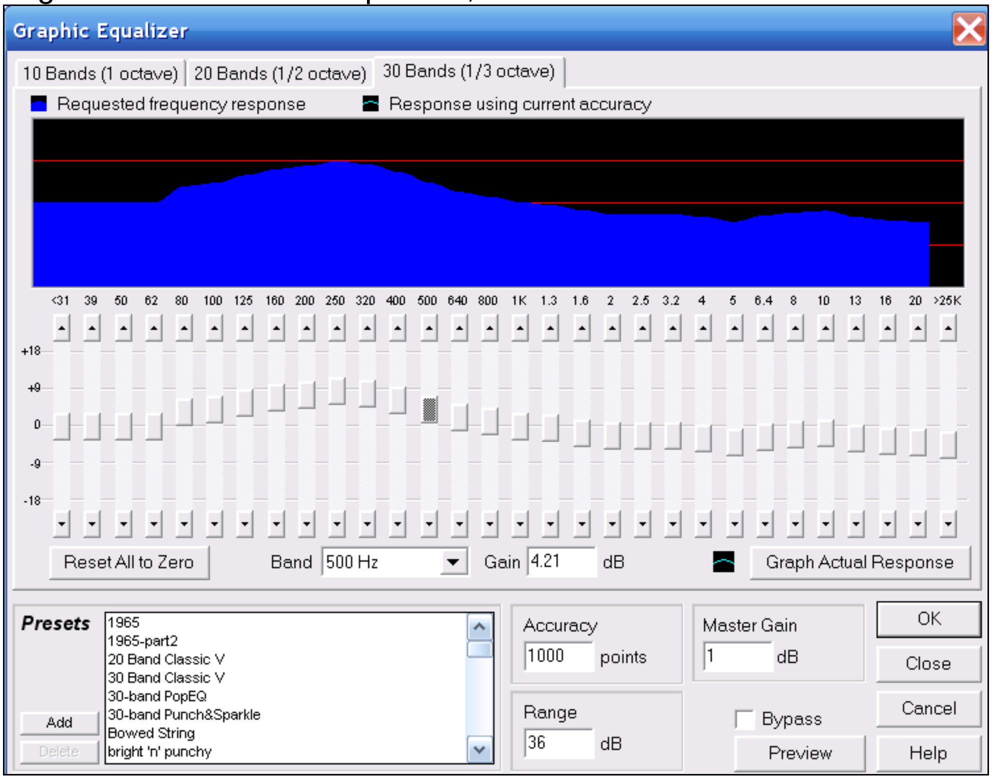
6.3.5 Limitations

The vocal passage was sung in each environment and thus there were slight differences in the recordings in terms of timing and pitch. As the nearby boundary layout of a room provides the musician with feedback, each environment would have an effect on the singer's attempt to reproduce the same vocal passage. Musicians often play/sing according to the feedback they get from the room. The singer was made aware of this issue and attempted to maintain her same sound irrespective of the room changes. This resulted in a few recordings being recorded in each room. It may be highlighted that the singer should have sung in the anechoic area and the recording played back through a high-fidelity speaker, which would then be re-recoded in each of the environments to enable perfect reproducibility of the vocal passage. That method was considered but it also has limitations, for example, a speaker would then need to be the sound source for each of the test locations instead of a human voice. The playing of a speaker in a room introduces new issues in terms of the coupling of the sound to the room (Eargle, & Foreman, 2002). Further, one of the points that this study attempts to highlight is that a moderately anechoic room can be used in place of a dedicated live recording room for recording of musicians in particular, vocals. Thus, the test needs to have a musician perform live in each location rather than a speaker replaying a vocal passage, as the recording studio does not record vocals from a speaker that contains already recorded vocals.

Matching the RT of two recorded sounds does not mean that they would sound the same. This is an interesting and important point. Two sounds may have similar RT_{60} and EDT yet have a very different sound characteristic. This can be shown by mapping the sound strength for two sounds as shown in Chapter 5. Thus, this chapter posed two challenges, firstly to match the reverberation characteristic of the desired sound as well as the graphical characteristic across the decaying wave. The matching of the graphics of the anechoic sound to that of the echoic vocals was conducted through listening tests of various frequency SPL changes. The task was essentially straightforward but it is noted that the researcher had some experience in this field. An example of the graphics tool used for adjustments is shown in the next figure.

One may argue that the original unedited vocal recordings should have also been included and compared in the listening test. They were not included as the anechoic vocals had a RT of less than 0.06s. The unedited vocals recorded in the three echoic rooms had a RT of well over 1s. Thus, if the participants could already differentiate between a 0.5s change in reverberation it is assumed that they would be able to perceive even larger amounts of reverberation change. Another variation would be to compare the original vocals recorded in the three lively rooms with each other. This was not conducted as the test was already over 10 minutes and this would be better suited for musicians and recording engineers as it relies on hearing the colourations of the sound, which is a difficult concept to explain without the participants having some acoustic knowledge.

Figure 6.1: Parametric equalizer, 30 bands.



6.4 Conclusion

In this chapter, an experiment was compiled to assess subjective components of acoustics. An experiment was conducted to assess the degree to which a group of participants were able to distinguish the artificially reverberated vocals from the normally reverberated vocals. A questionnaire design with an accompanying audio CD for listening was used to asses two independent groups. It was found through statistical analysis that the participants from both groups were unable to achieve a statistically significant result in identifying the artificially reverberated vocals from the normally reverberated vocals. Thus, this subjective test confirmed that artificially reverberated vocals could be successfully masked as naturally reverberated vocals.

There were a few self-describing questions included to gain insight into the sample groups personal beliefs and attitudes such as; love for music, experience in music and amount of time spent listening to music. It is important to keep in mind that acoustics is not a clear-cut objective field. It adds value to acoustic studies when personal experiential data are included. In terms of this study, interesting observations were identified through the subjective testing that was not seen in the objective testing. Together, both the subjective and the objective methods allow for a holistic approach to studying acoustical phenomenon.

Chapter

7

7 CONCLUSIONS AND RECOMMENDATIONS

7.1 Project Goals

The research undertaken in this study attempted to provide a scientific base for computer aided AR methods. It attempted to prove both objectively and subjectively that artificial methods can be used successfully in adding reverberation to sound impulses and vocals. Both these goals were achieved. This project also highlighted the relatively reduced requirements for recording and editing a vocal sample by the use of readily available equipment and software that is relatively cheap compared to traditional studio recording equipment.

7.2 The Bigger Picture

The music industry has been facing many technological challenges over the last few years. The introduction of MP3 coding methods with its reduced file size and ease of dissemination for example, has led many musical analysts to query the sustainability of the music industry (Liebowitz, 2004). Copy protection issues have been a controversial topic for some time now with DRM (digital rights management) presently taking a centre stage. Some major music recording companies have implemented a DRM-free approach that will open the door for music users to listen to their music on many mobile devices without the hassle of being confined to certain brands. They also hope that this move will promote people to buy more music (Sayer, 2008).

The widespread use of mobile devices that incorporate audio capabilities has further challenged the way music has traditionally been listened to. It has been noted that some artists have advised their recording engineers that their target market is mobile media device users and that that should be used as the reference sound goal.

The influence of the internet in lowering transmission costs for both authorised and unauthorised digital music copies has affected the market for copyrighted works considerably (Liebowitz & Watt, 2006). Music sales have shown a relative decline since 2001 (Zentner, 2005; Rob 2006). There has been and still is argument as to the cause of this decline (Oberholzer-Gee & Strumpf, 2007); however, it is clear that broadband internet, peer-to peer sharing, mobile media players and pirating have had some effect. Without entering that debate, one can highlight the many online websites that sell music legally which have increased in popularity over the years, not to mention the websites that while legal in their countries, infringe on copyrights of people who live by a different set of accepted copyright legislation in other countries.

Music purchased in its digital format without a hard copy (CD) has increased considerably with *Atlantic Records* stating that their digital sales have exceeded their CD sales (Arango, 2008). The format that the music is downloaded in is generally MP3 owing to its reduced file size as compared to its wave file equivalent. Most new music systems, whether dedicated hi-fi, home theatre or the like have the ability to play MP3 files as well as have interconnectability to flash drives and media players. As Jeff Goodell (1999) from *Rolling Stone* magazine stated before the new millennium, “the music industry is about to tumble head over heels into the Digital Age”. This has become a reality and has brought with it new opportunities as well as some disadvantages. In a video interview, Ken Hertz (2009) who is a lawyer representing top-rated music artists, provides an interesting perspective to the current state of the music recording industry. Hertz highlights two important points. Firstly, a new era in the contractual landscape between artist and record company has evolved. A break-away from 360-degree contracts[48] to multiple contracts with different companies all providing different services for the artist has now become commonplace with prominent musicians. However, this may not be the case for new musicians who are just starting their career.

Secondly and more relevant to this dissertation, there has been an increase in the available options for the aspiring [and established] artist particularly in the making and marketing of their music. This paragraph should not be misconstrued as an attack on the record companies as it is those companies who have developed and maintained the studio standard for recording of artists. It is just this point that highlights the dependence on the skill of the recording house to provide the aspiring artist with a high-quality musical recording. This comes at a considerable cost and it is understandable that record companies cannot offer every aspiring musician a recording contract especially when the artist is relatively unknown. Thus, many aspiring musicians find it difficult to obtain professional recording contracts. Many artists are also weary of the limitations imposed by some recording company’s exhaustive contracts on the new musician who is still somewhat of a risk to the recording company. The costs involved in obtaining the services of a professional studio recording company to provide a “once off” no royalty recording are also a stumbling block for new musicians. There is also the risk that the outcome of the recording process may not be what was desired and is generally left in the hands of the recording engineer to compile the “right” sound.

Many musicians have noticed that selling their music is quickly becoming the lesser goal of generating an income, but rather the stepping-stone for creating the demand for live performances, which can provide sustainable revenue.

[48] A contract whereby the musician agrees to one company (usually a record label) having the rights to multiple parts of the musician’s work including the sale of CDs, digital media, merchandise, concerts, endorsements and possibly even ringtones. The company obtains profit share from all these areas.

The use of computer-aided methods in reducing the infrastructure requirement can help in reducing the obstacles faced by new artists. While there is still a skill level required in terms of computer aided methods, the versatility has been shown to be large.

It is not an easy task to record and market one's own music. However, with the changes in information availability, as well as the improvements in computers, new methods of music recording, reproduction, marketing and sales have become available. For example, an aspiring musician can record their music and make it available for their fans without ever manufacturing a single physical CD. Artists can upload their own tracks onto their own websites and allow their fans to purchase their music online. They could even send their music to a music sales website that will provide the sales for them. Various payment methods are available. The music landscape has changed considerably. Independent recording studios are available for use. While the recording artist knows that it takes hours to record a single music track, the time spent in the studio is usually limited and costly, especially in the traditional type studios. If the same performance standard were to be obtained in the privacy of one's own home, this would be a favourable solution for many recording artists. The main obstacle in achieving this goal would be the acoustic reinforcement and treatment costs as well as the equipment costs.

7.2.1 A Personal Note

A motivating factor in deciding on this topic originated from the many frustrated musicians that the researcher came into contact prior to this project. There seemed to be an abundance of talented musicians who were not able to realise their goals owing to either economic reasons or lack of skill in terms of the recording process and/or music industry. After highlighting that there are options available for someone / some bands that has/have not been given a studio recording opportunity, the feedback was that the proposed methods were not the "correct" way to become a successful musician. While there are numerous reputable books on making a home studio or home recording studio methods, it seemed that those alternatives were perceived as inferior. This dissertation will hopefully add momentum to those stated alternative methods and is especially suited in a context where economic factors are of concern.

The use of AR to compensate for lively rooms is possible. Many aspiring musicians could convert an available room into their home recording studio. They do not need large lively rooms that are designed for a specific RT, rather they could crudely use a room whereby there are drapes and carpeting which would provide only a modest RT. The AR software would be able to add the additional reverberation. Added to this, many musicians use digital instruments, for example, electronic keyboard over a piano, which is then connected directly to the PC and has many options for different instrument sounds (drums, guitar, violin etc). There are software-based instruments that are also available and are independent of room type. Thus, a change in energy and focus from the complete reliance on the environment to the reliance on the computer is needed to improve versatility and reduce costs for musicians who have not got access to the professional recording studios.

7.3 Was This Study Successful?

Methods of assessing reverberation were looked at including a review of the new ISO 3382:2008 standard. Experimental chapters were then presented, firstly to measure and analyse the RT and EDT of four different locations and then secondly to be used as a basis for further experimentation. The measurements were conducted in accordance with the ISO standard and had many graphical presentations. This trend remained throughout Chapter 4 and 5 where numerous waveforms were compared and analysed with octave-filtered plots. The three different reverberant locations were then critically compared to an almost anechoic environment. This analysis was conducted according to the latest standards including the use of computer-aided software for computation and evaluation. The data obtained from this analysis was then used in two further applications, which were the central focus of this dissertation. An objective and subjective experiment was undertaken where the artificially reverberated samples were manipulated to mimic the naturally reverberated samples. In the objective test, the results were evaluated purely by mathematical methods and consisted of impulse samples. In the subjective evaluation chapter, vocal samples were used which were also artificially reverberated to mimic the naturally reverberated vocal samples. Statistical testing was incorporated to evaluate whether two sample groups could identify the artificially reverberated from the naturally reverberated vocals. The results confirmed that artificial methods are not easily discriminated from naturally reverberated sounds.

Throughout this study, an attempt was made to find a balance between the impersonal positivistic epistemology to that of the interpretivistic experiential paradigm. While the majority of this text would be classified under the former, at least some parts could be classified as the latter. The motivation for this goal was that this topic in acoustics consists of both objective aspects as well as subjective personal facets.

To answer the paragraph question, did this study achieve its goal? In terms of using digital artificial methods in successfully applying AR to impulse and vocal samples to the degree that they could be passed off as the naturally reverberated samples, then the answer is affirmative. In terms of providing the reader with a balanced two-sided approach to the subject matter, that could only be answered by the reader.

7.4 Further Study

Acoustic reverberation has many applications and several areas of interest were uncovered while conducting this study.

It was found that an anechoic environment could be matched to a lively room in terms of its RT. While reverberation is one of the acoustic parameters that make up the experience of sound, there are many other parameters that too are important and could be looked at in a similar manner. The goal could be to artificially copy the sound of sought after rooms/auditoriums and bring these widely cherished sound types to more people. For example, most people would not have the opportunity to experience and compare the great sounding halls that are available worldwide, albeit just attending one such performance venue. If these top halls are acoustically mapped and artificially reproduced to bring that sound experience to more people in a cheaper setting, this would be a great achievement for music. It is accepted that it would be near impossible to copy the sound exactly, however a close copy may be ascertainable.

In terms of recording studios, if well-known and liked studios have their reverberation characteristic mapped, the data can be applied to semi-anechoic rooms to provide them with a similar RT and EDT sound. These settings can be programmed into an acoustic software program that can be used to artificially reverberate the dry sound to sound like the pre-set well-liked studio. Thus, the famous studios can be copied in their RT and EDT and used as "plug-ins" for the RT option settings in programs such as Cubase and Cool Edit Pro.

Reverberation time and EDT are not the only parameters that are needed in order to match the sound in a room; further work can be undertaken in the field of AR applied to moderately dead recording rooms where other parameters can be controlled for as well. This would be in keeping with Bradley's (2005) recommendation of the study of sound level strength, G, early to late arriving sound ratio C_{80}, LF, and IACC. The ultimate goal would be to have these settings available in the recording software suites.

The computer-recording studio relies on various software packages to edit the recorded tracks. Most studios tend to use a particular version. There are a few professional recording suites available such as Sonar, Adobe Audition, FL Tools, Nuendo, Reason, ACID Pro and Cubase. There is a lot of overlap between these programs. There are also amateur suites that have an abundance of settings that are comparable with the professional versions. It would be interesting to perform a comparative analysis of the available software recording suites as well as provide a method of testing the acoustic quality of the settings that each program provide. For example, audio filtering (applying Bessel, Butterworth or Chebyshev to an audio sample) is a commonly used tool in studio editing. How accurate are these different programs in providing their stated parameters. Can the recording engineer who is accustomed to using program "X" immediately use program "Y" and apply the same settings to obtain the same sound output? If the outputs are compared, would they be identical and do they meet their specification when evaluated by acoustic MLS software [and when evaluated by listening tests]?

The subjective testing used in this study had mixed results for the first reverberation test. The participants had to choose the most reverberant of the vocal samples heard. Many participants were able to achieve 75% while others got a mark of zero. What was most surprising was that some participants obtained incorrect answers for what was seemingly clearly reverberant sounds compared to the less reverberant sounds played. The reasons for this finding remain unclear. One possible reason is that reverberation is a personal experience and is not easily objectified. There have been numerous studies on the optimum RT for various music and speech types. However, many studies have been dedicated to contexts such as classrooms, auditoriums and the like. With the recent changes in the options that people have available to listen to their music, further study can address the optimum reverberation and other acoustic parameters for mobile devices, including devices that rely on headphones, devices that do not have stereo output, devices that are used in varying environments such as mobile phones and tablets. Added to this, many of these devices are used where there is additional background noise. What effect does the background noise have on the experience of reverberation? The EDT is known to be a critical aspect in the experience of reverberation. In terms of the new technology available to consumers, what role does EDT have on the experience of sound for these systems?

Many music users listen to their music on a mobile device. These devices mainly rely on MP3 data. The bit rates are not fixed and range from less than 120kbps to 320kbps. Higher bit rates are associated with improved quality of sound. What is the effect on reverberation as well as other acoustical parameters in terms of changes in the bit rates and music types? A study was conducted to compare the performance of impulse response measures with and without MP3 data compressions. The finding was that the RT_{60} errors decrease with an increase in bit rate, which was expected. The researchers noted that bit rates equal to or higher than 128kbps had minimal error in RT (Paulo, Martins, & Coelho, 2002:8). While their research addressed the objective evaluation of RT through MLS, a subjective evaluation may provide a different result. For example, subjective testing on a group of mobile music users to determine whether they could notice a difference in terms of various acoustic parameters for different bit rates may provide interesting results. There is disagreement as to what is an acceptable bit rate in terms of sound quality and file size. Many forums on the internet have been dedicated to just this topic with mixed opinions. A scientific study in this field may provide further light on this topic.

The objective application section of this dissertation had the anechoic samples reverberated to sound like the naturally reverberated vocals. On many occasions, the applied AR was too strong in some octave bands while perfect in other bands. The *scientific filters* in Cool Edit Pro were used to reduce the excess AR in the problematic octave bands. The results were not always pleasing, however with trial and error testing there were positive results realised. It is not an easy task to remove reverberant sounds once they are combined with the wanted signal as they are not easily differentiated from the original, owing to their content being similar (Marro et. al.,1998). In terms of the application shown in this study, *scientific filters* were helpful in removing some reverberation, thus further work into this topic could still be completed in applying this method and finding out if it is a viable option.

Anechoic rooms, while allowing for crisp sound, lack running reverberation[49], which fills in the gaps between the notes played by the musicians. For example, a string quartet would be heard as a group of independent soloists. Dry rooms tend to create a sound field whereby the sound seems to be originating directly from the front of the listener, and the musicians themselves usually do not enjoy the experience of such rooms. This is the basis for the main objection of the dry recording method. A room that has running reverberation would allow for the quartet to sound as a unit with the harmonics intermingling as they play and feed off each other, including the feedback obtained from the room. If the musician who was being recorded has real-time reverberation applied and played back to them through headphones, this may improve their experience of the dry room recording.

[49] Running reverberation is the sound that is heard while the music is continuously playing.

REFERENCES

Adobe Systems Incorporated. (2008). *Cool Edit Pro is now Adobe Audition*. Retrieved 1 July 2008 from http://www.adobe.com/special/products/audition/syntrillium.html

Ahlzen, L. & Song, C. (2003). *The Sound Blaster Live! Book.* A Complete Guide to the World's Most Popular Sound Card. No Starch Press. San Francisco. USA

Aldred, J. (1971). *Manual of Sound Recording*. (2nd Ed). Fountain Press London

Ando, Y. (1998). *Architectural Acoustics: Blending Sound Sources, Sound Fields, and Listeners*. Springer-Verlag, New York

Arango, T. (2008, November 26). Digital Sales Exceed CDs At Atlantic. *New York Times*, Retrieved July 31, 2009, from Academic Search Premier Database

Balachandran, C.G (1976). *Journal of Sound and Vibration*. Pitch change during reverberant decay Elsevier Ltd. Vol 48, Issue 4, Pages 559-560. Retrieved on 7 June 2009 from http://www.sciencedirect.com/science

Bartlett, B. & Bartlett, J. (2009). *Practical recording techniques.* (5th ed). Elsevier. Oxford, UK.

Barron, M. (1993). *Auditorium Acoustics and Architectural Design*. Spon Press, London.

Beranek, L. L. (1954). *Acoustics.* McGraw-Hill Book company. York, PA

Beranek, L. L. (2004). *Concert halls and opera houses: music, acoustics and architecture.* (2nd ed). Springer-Verlag. New York pp 21

Berg, R. E. & Stork, D. G. (1982). *The Physics of Sound*. Prentice-Hall, Inc., Englewood Cliffs, N.J. pp 203-220

Bork, I. (2002) Simulation and measurement of auditorium acoustics - the round robins on room acoustical simulation. *Institute of Acoustics.* Vol. 24.

Boyk, J. (2000). *There's life above 20 kilohertz! a survey of musical instrument spectra to 102.4 kHz*. retrieved from http://www.cco.caltech.edu/ boyk/spectra/spectra.htm on 10 July 2009.

Bradley, J. S. (1996). Optimizing the Decay Range in Room Acoustics Measurements Using Maximum-Length-Sequence Techniques. *Audio Engineering Society*. Volume 44 Issue 4 pp. 266-273. April

Bradley, J. S. (2005). Using ISO 3382 measures, and their extensions, to evaluate acoustical conditions in concert halls. *Acoustical Science & Technology*. Vol 26, 2 pp 170-178

Browne, S. (2001). *Hybrid Reverberation Algorithm using Truncated Impulse Response Convolution and Recursive Filtering*. (Masters dissertation). University of Miami

Brüel & Kjær. (1984). *Measuring Sound*. (Rev. ed). BR0047-13. Brüel & Kjær. 2850 Nærum. Denmark

Brüel & Kjær, (1988). *Measurements in Building Acoustics.* BR0178. Brüel & Kjær. 2850 Nærum. Denmark

Brüel & Kjær. (2008a). BZ-7204 v.2: How *do I avoid overload in RT measurements with a pistol?* Retrieved on 15 June 2008 from: http://www.bksv.com/Support/FAQ.aspx

Roads, c. (1996). *The computer music tutorial*. MIT Press

Robinson, D. W., and Dadson, R. S, (1956). A Re-Determination of the Equal-Loudness Relations for Pure Tones. *British Journal of Applied Physics*. 7, No 5 166-181

Dragonetti, R., Ianniello, C., & Romano, R. (2009, January). Reverberation time measurement by the product of two room impulse responses. *Applied Acoustics*, *70*(1), 231-243. Retrieved April 17, 2009, doi:10.1016/j.apacoust.2007.12.001

Coetzee, C.H., Janeke, H.C., Terre'Blanche, M.J., Kruger, P. & Payze, C. (2001). Psychological Research. UNISA. Pretoria p134

Choi, Y., & Cabrera, D. (2005). Some current issues in computer modelling for room acoustic design. *Australian acoustical society vol. 33 no. 1 April pp 19-24*

Cox, T. J. and D'Antonio, P. (2004). *Acoustic absorbers and diffusers - theory, design and application.* Spon Press. New York

Eargle, J. & Foreman, C. (2002). JBL *Audio Engineering for Sound Reinforcement*. Hal Leonard Corporation. Milwaukee, WI.

Evanescence. (2003). *Fallen* [CD]. United States: Wind-up

Everest, F.A. (1981) *The Master Handbook of Acoustics*. Tab Books Inc. Blue Ridge Summit, PA

Everest, F. A (2001). *The Master Handbook of Acoustics*. (4th ed). McGraw-Hill. USA

Everest, F.A. (2007). *Critical Listening skills for Audio Professionals.* Course Technology Cengage Learning. Canada.

Gervais, R. (2006). *Home Recording Studios: Build it Like the Pros*. Thompson Course Technology

PTR. Boston, MA
Goodell, J. (1999, July 8). World War MP3. *Rolling Stone*, Issue 816/817. Retrieved July 31, 2009, from Academic Search Premier database.
Griesinger, D. (1996). Spaciousness and Envelopment in Musical Acoustics. *Audio Engineering Society*. Paper number 4401 AES Convention 101 (November).
Hak, C.C.J.M., & Vertegaal, J.S. (2007). MP3 stimuli in room acoustics. *International congress on acoustics.* Madrid, 2-7 September 2007
Hertz, K. (2009, July 15). *Ken Hertz about music industry today***.** PICNIC Salon New York [Seminar] [Webcast]. Retrieved 31 July from http://gabemac.com/2009/07/15/ken-hertz-about-music-industry-today/
ISO. (1997). *ISO 3382:1997, Acoustics — Measurement of the reverberation time of rooms with reference to other acoustical parameters*. International Organisation for Standardisation, Geneva, Switzerland.
ISO. (1999). *ISO 3741:1999 Acoustics - Determination of sound power levels of noise sources using sound pressure - Precision methods for reverberation rooms.* Standards South Africa. Groenkloof Pretoria
ISO. (2008). *ISO 3382-2:2008,* Acoustics — Measurement of room acoustic parameters — Part 2: Reverberation time in ordinary rooms. International Organisation for Standardisation, Geneva, Switzerland.
ISO. (2008b). *International Organization for Standardization*. Viewed on 16 June 2008 from: http://www.iso.org/iso/search.htm?qt=reverberation&searchSubmit=Search&sort=rel&type=simple&published=on
ISO. (2009). *ISO 226:2003*, Acoustics -- Normal equal-loudness-level contours. Retrieved 15 June 2009 from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue
Ivie. (2007). *IE-35 & IE-45 RT-60 Manual*. Ivie USA. Retrieved 4 July 2008 from http://www.ivie.com/products_inst_ie45intro.htm
Kang, J. (2002). *Acoustics of Long Spaces: Theory Design and Practice*. Telford Ltd. London.
Kinser. L,E., Frey, A.R., Coppens, A.B., and Sanders, J.V. (2000). *Fundamentals of Acoustics.* (4th Ed). John Wiley & Sons. New York, NY
Lee, S. K. & Ho. J, H. (2002). Reverberation time measurement of a structure or a room using modified wavelet transform (Abstract). *The Journal of the Acoustical Society of America*, vol. 112, iss. no. 5, pp. 2397-2397
Lee, S. K. (2002). An acoustic decay measurement based on time-frequency analysis using wavelet transform. *Journal of Sound and Vibration*, 252 (1), pp. 141-153.
Lee, S. K. (2003). Measurement of short reverberation times at low frequencies using wavelet filter bank. *KSME international journal.* vol. 17, no 4, pp. 511-520
Lee, S. K. &Yu. D. J. (2005). Measurements of reverberation time of a passenger car utilizing the wavelet filter bank. *Proceedings of the Institution of Mechanical Engineers, 219,3; pp 329-336*
Liebowitz, S. (2004). Will mp3 downloads annihilate the record industry? The evidence so far. In *Advances in the Study of Entrepreneurship, Innovation & Economic Growth*. Vol 15 Emerald Group Publishing Limited pp229-260
Liebowitz, S. J & Watt, R. (2006, August). How to best ensure remuneration for creators in the market for music? Copyright and its alternatives [Abstract]. *Journal of Economic Surveys*. Volume 20 Issue 4**,** Pages 513 - 545. Retrieved July 31, 2009 from www3.interscience.wiley.com/journal/118583569/abstract
Marro, C., Mahieux, Y., and Simmer, K.U. (1998). Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering. *IEEE Transactions on Speech and Audio Processing*. Vol 6, Issue 3, pp 240 - 259
Miller, R. 2007. *Why DSP Boxes Set the Same Way Differ*. Rane Note 167. Rane Corporation. Retrieved on 27 April 2011 from http://www.rane.com/note167.html
Morgan, D. R. (1997). A parametric error analysis of the backward integration method for reverberation time estimation. *The Journal of the Acoustical Society of America*, 101(5):2686-2693.
Morset Sound Development. (2008). *WinMLS 2004*. Retrieved on 4 July 2008 from http://www.winmls.com/
Murphy, D.T., Howard, D. M. & Tyrrell, A.M. Multi-channel reverberation for computer music applications. *Signal Processing Systems, 1998. SIPS 98. 1998 IEEE Workshop* Publication Date: 8-

10 Oct 1998. pp 210-219. Retrieved from http://0-ieeexplore.ieee.org.ujlink.uj.ac.za on 24 April 2009
Nisbett, A. (1995). *The sound studio*. (6th ed). Focal Press. Jordan Hill, Oxford
Nisbett, A. (2003). *The sound studio*. (7th ed). Focal Press. Jordan Hill, Oxford
Oberholzer - Gee,. F., & Strumpf, K. (2007). The Effect of File Sharing on Record Sales: An Empirical Analysis. *Journal of Political Economy* 115:1, 1-42 Online publication date: 1-Feb-2007.
Oguchi, K. (1998). Symposium on Measurement Methodology of Room Acoustical Parameters. *Nagata Acoustics News*. 98-1 No.121 Issued : January 25, 1998. Retrieved 6 July 2008 from http://www.nagata.co.jp/news/news9801-e.htm
Olsen, H.F. (1967). *Music, Physics and Engineering*. (Revised Ed). Dover Publications
Pohlmann, K. C. (2005). *Principles of digital audio*. (5th ed). McGraw Hill. New York pp 332-413
Rob, R. (2006, April). Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students. *The Journal of Law and Economics*, Vol. 49 number 1.
Roche, J.P. (2003). *USB 2.0 Sound: Creative Labs Audigy 2 NX .* Available: http://www.tomshardware.com/reviews/usb-2,711.html. Last accessed 28 June 2008.
SANS. (1999). SANS3741: Acoustics - Determination of sound power levels of noise sources using sound pressure – Precision methods for reverberation rooms. Standards South Africa. Groenkloof Pretoria.
Sayer, P. (2008, March). Third Major Music Studio Set to Go DRM-Free. *PC World*, *26*(3), 24-24. Retrieved July 31, 2009, from Academic Search Premier database.
Shure (2009). *Model SM58 Dynamic Microphone Specification Sheet*. Retrieved on 15 January from http://www.shure.com/proaudio/products/wiredmicrophones/us_pro_sm58-cn_content
Smith, B. J., Peters, R. J. & Owen, S.(1996). *Acoustics and Noise Control.*(2nd Ed). Adison Wesley Longman Ltd. Essex. England. pp 1 -
Suzuki, Y., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *Journal of the Acoustical Society of America*. Vol 116 (2), pp918-933
Syntrillium Software Corporation. (2002). *Cool Edit Pro Version 2 Users Guide*. Phoenix, AZ USA
Templeton, D. (Ed), Sacre, P., Mapp, P. and Saunders, D. (1997). *Acoustics in the Built Environment – Advice for the Design Team* (2nd ed). Architectural Press. Oxford
Tremaine, H.M (1974). *Audio Cyclopedia*. (2nd ed). Howard W. Sams & Co., Inc, Indianapolis. p44 White, H.E and White, D.H. (1980). *Physics and Music.* Saunders College. Library of Congress Cataloging in Publication Data. USA. pp 352-380
Tsutomu, O., Emi, N., Norie, K., Yoshitaka, F. & Hiroshi, I. (1991). High-Frequency Sound Above the Audible Range Affects Brain Electric Activity and Sound Perception. *Audio Engineering Society*. AES Convention:91 (October) Paper Number:3207
Watkinson, J. (1998). *The art of sound reproduction*. Focal press. Woburn MA
WinMLS, (2004). *WinMLS 2004 Reference Documentation*. Morset Sound Development. Retrieved on 3 July 2008 from www.winmls.com
Yoshimasa Electronic Inc (2005). *DSSF3 - RA - Impulse response Manual*. Retrieved on 3 July 2008 from http://www.ymec.com/manual/esa/appendix.htm
Zentner, A (2005). File Sharing and International Sales of Copyrighted Music: An Empirical Analysis with a Panel of Countries. *Topics in Economic Analysis & Policy*: Vol. 5: Issue. 1 pp63-90, Article 21. Retrieved on 31 July 2009 from www.bepress.com/bejeap/topics/vol5/iss1/art21
Zhang, Y. (2005). *A Method to Predict Reverberation Time in Concert Hall Preliminary Design Stage*. Doctoral Dissertation, Georgia Institute of Technology.

APPENDIX A

A.1 The Impact of a Person's Physical Position in a Moderately Reverberant Room and Its Effect on SPL

Procedure Overview

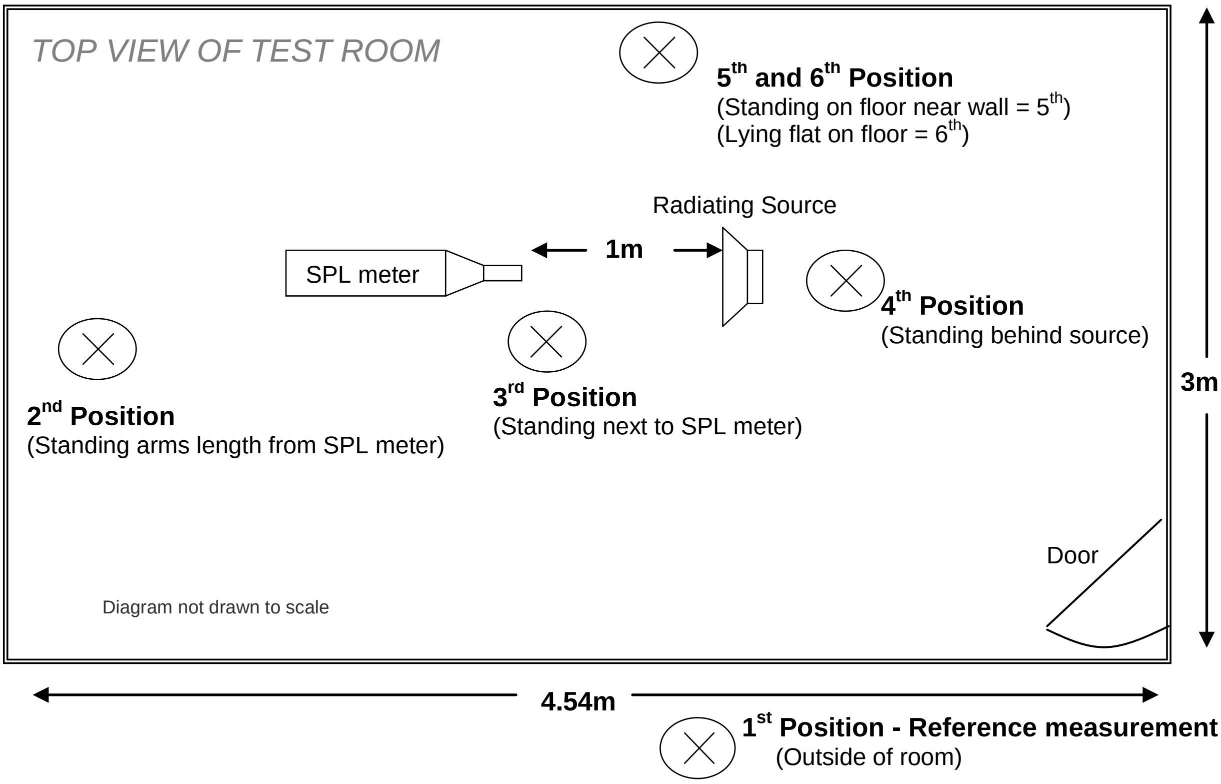
Using the same reverberant room as in Chapter 4 (small room), the researcher was positioned in different locations in the room while measurements of the SPL of a single radiating source were undertaken. The SPL value was recorded for each of the positions. The equipment set-up was the same as in Chapter 4, except that the source was changed from an impulse of a balloon to a 1kHz sine wave output from a 6W speaker. The SPL meter and radiating source were not moved during the experiment and only the researcher was changing positions.

Table A.1: Summary of test setup data for experiment.

Date:	27 June 2008
Site:	Doornfontein, JHB (UJ) Lab 4319 John Orr Building
Room dimensions and description:	Shown in Chapter 4.
Environmental conditions:	Clear sunny day
Equipment required:	SPL meter, oscilloscope and leads, laptop with media player, audio test CD, speaker, two stands, 2m ripcord, tape measure, spirit level, compass, thermometer and barometer.
Applicable standards followed: (Method reference)	ISO/SANS 3741; BR0047-13 (Brüel & Kjær, 1984) *Measuring Sound*. (Rev. ed).
Temperature:	23°C
Barometric Pressure:	834hPa
Background noise (dBA):	39,0-42dBA
Measuring height above ground plane:	1,3m
Distance between source and SPL meter:	1m
Signal type and value:	$3V_{p-p}$ @ 1kHz (Sine wave)

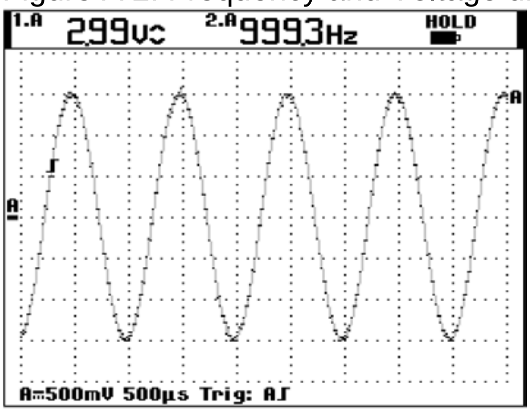
The floor plan (Fig. A.1) shows the respective positions that were used. Each position was taken as a single measurement and sufficient time was allowed for the signal to reach steady-state while the measurements were recorded. The positions were re-checked and confirmed to be accurate. The next figure shows the floor plan including the different locations that the researcher used. If the reader would like to see photos of the room, they are referred to Chapter 4 where there are pictures of this test site under the heading "*Reverberation Time Analysis for Small Room*".

Figure A.1: Floor plan of the echoic room showing the source, SPL meter and the sequential locations of the researcher (Top View).



The researcher used a portable oscilloscope to obtain the output signal that was present at the radiating source. The source was set to radiate at 3Vp-p with a frequency of 1kHz as shown in Figure A-2.

Figure A.2: Frequency and Voltage drop across the radiating source speaker.



The radiating source maintained its output throughout the measuring process. Each time the researcher's position changed, a period of 10 seconds was timed to allow for the signal to reach steady state. The SPL value was measured using a digital storage oscilloscope fed from the SPL meter as it was not possible to read the SPL meter's screen while taking the measurements.

Results

The SPL level for the first measurement was taken as the reference as the researcher was not present in the room and thus had no effect on the radiated sound pressure[50].

[50] Door of room was closed.

The different SPL values for each position and their respective deviation from the reference measurement are shown in the following table.

Table A.2: Results for the different physical locations and their respective dBA values.

POSITION OF RESEARCHER	1st attempt: dBA value	2nd attempt: dBA value	Average dBA value
1st (Outside of room)	84,4	84,3	**84,4**
2nd (Standing arms length from SPL meter)	81,4	80,7	**81,1**
3rd (Standing next to SPL meter)	84,4	85	**84,7**
4th (Standing behind source)	83,6	83,4	**83,5**
5th (Standing on floor near wall) [51]	82,8	83	**82,9**
6th (Lying flat on floor in same position as 5 above)	84,1	84,4	**84,3**

Discussion

The reference value of 84,35dB was the second highest value recorded which suggests that the presence of a person in this test room would only have a resultant absorptive effect on the SPL radiated and reflected in the room for a 1kHz pure sine wave. While it is known that people can also reflect sound and increase the SPL, particularly at 400Hz (Brüel & Kjær, 1984:25), this did not seem to be the case for this experiment. The overall effect of the researcher's presence was absorptive, as it reduced the SPL for all but one measurement position. The largest attenuation effect was 3,3dB while the least attenuation effect was 0,1dB. Position 3 showed a slight increase in SPL of 0,35dB. It would be interesting to see how these measurements would change if a lower frequency was used.

Most manufacturers of SPL and real-time analysers recommend that the user stand at least arm's length away from the measuring device. This was achieved in the 2nd position where the researcher stood arm's length with his arm stretched out so as to imitate a person holding the meter (the meter was already on a stand). No part of his body obscured direct waves. The reduction in SPL measured was 3,3dB. This was not expected. This result was the largest deviation from the reference value. It seems probable that by standing at arm's length in the test room may have coincided with the some of the room's modes, which were probably consequently reduced. Thus, reflected waves were obscured and absorbed and thus a reduction in SPL occurred. Interestingly, standing next to the SPL meter allowed for a higher SPL level than obtained for standing in the "correct" position of arm's length. Standing next to the SPL meter resulted in the highest SPL value, which was also higher than the value measured while outside of the room. Position 5 allowed for an average SPL of 82,9dB, which was 1,45dB lower than the reference. When the researcher lay down in that same position instead of standing, the SPL increased to 84,25dB, which was now only 0,1dB lower than the first position where the researcher was outside of the room.

Limits and Notable Conditions

1. The room does contain parallel walls and thus is not a pure reverberant room.
2. The researcher was wearing a thick jacket, which may have exacerbated the absorptive effect of his presence. His pants were normal jeans.
3. The height of the researcher is above average at 185cm.

[51] This measurement was particularly sensitive to even the slightest physical movement.

4. While taking measurements in each position, it was noticed that minor movements would change the SPL. The changes were slight but nevertheless there were fluctuations. The researcher tried to stay as still as possible during measurements, sometimes not breathing for the duration of the measurement. Changes in SPL were noted for both slight movements and for the noise related to the moving of clothing. This room was found to exaggerate even minor sounds.

Conclusion

Position changes in a room that is echoic can change the SPL by a considerable amount. Minor changes in movement were picked up by the SPL meter. The ability of reverberation to increase the SPL in the test room was moderated by the position of the researcher.

APPENDIX B

B.1 Test to Show That Reverberation Increases the SPL Measurement of a Sound Source.

Procedure Overview
Setup an SPL measurement of a steady state radiated acoustic signal in an area that is anechoic. Setup the same experiment in an area that is more echoic. Compare the SPL results and obtain a conclusion.

Figure B.1: Summary of test setup data for experiment.

	Measurement 1: Anechoic	Measurement 2: Echoic
Date:	23 June 2008	26 June 2008
Site:	Farm in Heidelberg (*Rooikraal Landgoed*)	Doornfontein, JHB (UJ) Lab 4319 John Orr Building
Room volume	NA	50,07m^3
Environmental conditions:	Clear sunny day	Clear sunny day
Equipment required:	SPL meter, oscilloscope and leads, laptop with media player, audio test CD, speaker, two stands, 2m ripcord, tape measure, spirit level, compass, thermometer, barometer and battery power source/inverter. Microphone cables used were XLR/F to XLR/M and all cables were balanced where applicable.	
Applicable standards followed: (Method reference)	ISO/SANS 3741; BR0047-13 (Brüel & Kjær, 1984) *Measuring Sound*. (Rev. ed).	
Measuring height above ground plane:	1,3m	
Distance between source and SPL meter:	1m	
Signal type and value:	3$V_{p\text{-}p}$ @ 1kHz (Sine wave)	
States of occupancy	unoccupied	

Procedure – Detailed
<u>Anechoic area</u>
The same environment that was used as the reference measurement in the RT measurement and analysis (Chapter 4) was used for this experiment. The echoic venue was the small room that was also used in the RT measurement and analysis section.

<u>Equipment setup for Anechoic Measurements</u>
The SPL meter and sound source (speaker) were put onto their own upright stands. The stands were spaced 1m apart and were spirit levelled to be parallel with the ground plane. It was important to get the SPL microphone directly in line with the speaker. A digital compass was used to point the SPL meter opposite the speaker. The speaker was positioned and the cardinal direction was noted[52]. The next figures show the alignment of the source and SPL meter.

The SPL meter and sound source were put onto their own upright stands. The stands were spaced 1m apart for SPL measurements, and then spirit levelled to be parallel with the ground plane. According to the ISO standard, the microphone should be directly facing the source, while the ANSI standard requires an angle of incidence of 70-80 degrees (Brüel & Kjær, 1984). A compass was used to calibrate the angle of incidence.

[52] The speaker's magnetic field was noted. The speaker's magnet did not affect the compass reading.

Figure B.2: Alignment of SPL meter using spirit level.

Figure B.3: Alignment of sound source.

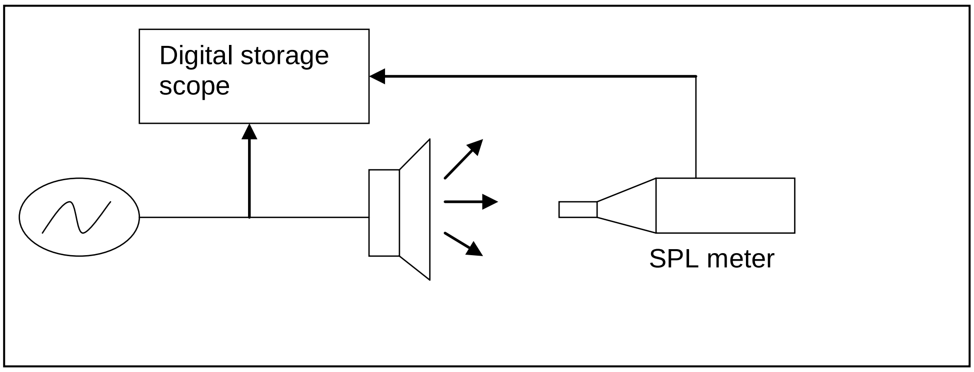
The surrounding environment had no sound reflective vertical objects except the SPL meter and vertical stand. There would be some reflections from the ground however. A blanket was used to minimize these reflections. There may also be some reflection from the test gear as they are made from reflective materials. The blanket was used to cover the test gear as much as possible.

The speaker's amplifier was powered by the laptop's USB port[53]. The audio signal fed to the speaker was from the mini audio jack from the laptop. Windows Media Player was used to play the steady state sine wave audio signals[54]. The outputted voltage across the speaker was measured for both frequency and voltage. The SPL meter was also connected to the oscilloscope.

[53] The USB power rail was found to be connected directly to the battery for this laptop thus the power available was more than 30W. The power required was low and did not exceed 7W.

[54] An audio test CD was used that consists of 99 tracks of various sine wave frequencies and noise types. The sine wave tracks used have been verified by checking them on an oscilloscope and frequency counter.

Figure B.4: Basic circuit layout of the test gear.



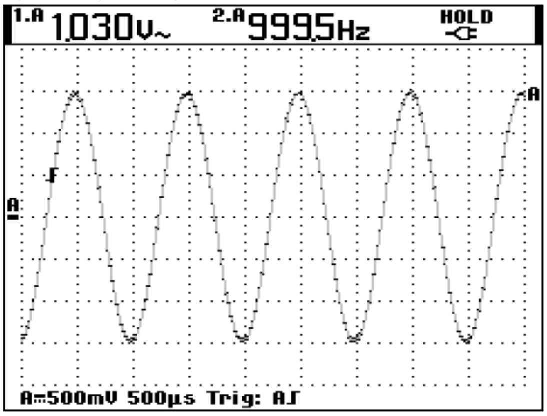
The experimental setup for anechoic measurements has been shown in Chapter 4 under the anechoic test section heading “*Reverberation Time Analysis for the Anechoic Environment (reference)”*.

The SPL was obtained for three different frequency values, 1kHz, 500Hz and 5kHz. Both dBA and dBC SPL values were recorded as well as the AC and DC output signals from the SPL meter.

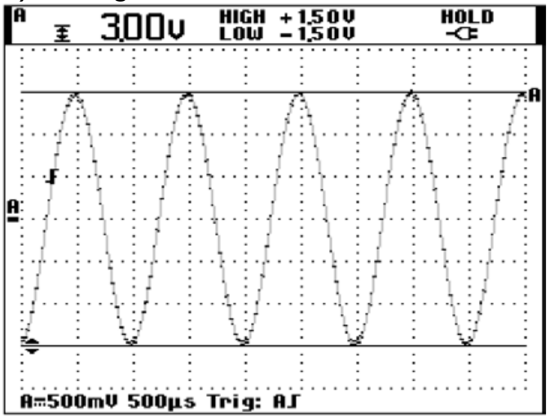
For the first measurement of 1kHz, the speaker was set to play at $3V_{p\text{-}p}$ ($1{,}06V_{rms}$) which became the reference for all the measurements both echoic and anechoic. Thus, for each measurement a $3V_{p\text{-}p}$ was obtained at the three different frequencies as shown in the oscilloscope screens that follow.

Figure B.5: Frequency and voltage drop across the radiating speaker source for 1kHz.

a) Frequency

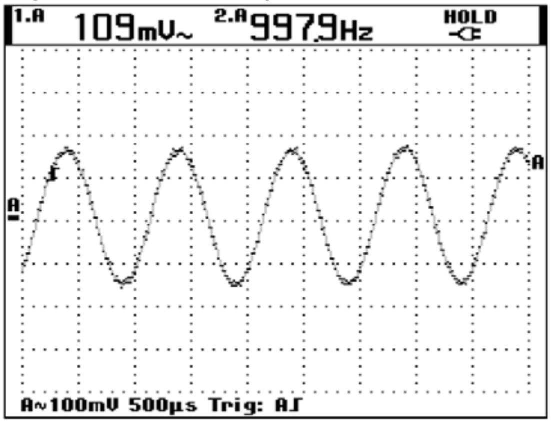


b) Voltage



A scope screen of the SPL meter’s AC output was also included for the 1kHz signal test. The SPL meter’s AC output is a measure of the sound level picked up at the SPL meter.

Figure B.6: AC output from the SPL meter.



The frequency that the SPL meter registered was 997.9Hz, which was close to 999.5Hz outputted by the radiating speaker, thus the electret microphone pickup in the SPL meter was accurate with reference to frequency. The displayed waveform was also a good representation of the sine wave outputted from the speaker, although a phase shift was noted.

Figure B.7: Frequency and voltage drop across the radiating speaker source for 500Hz.

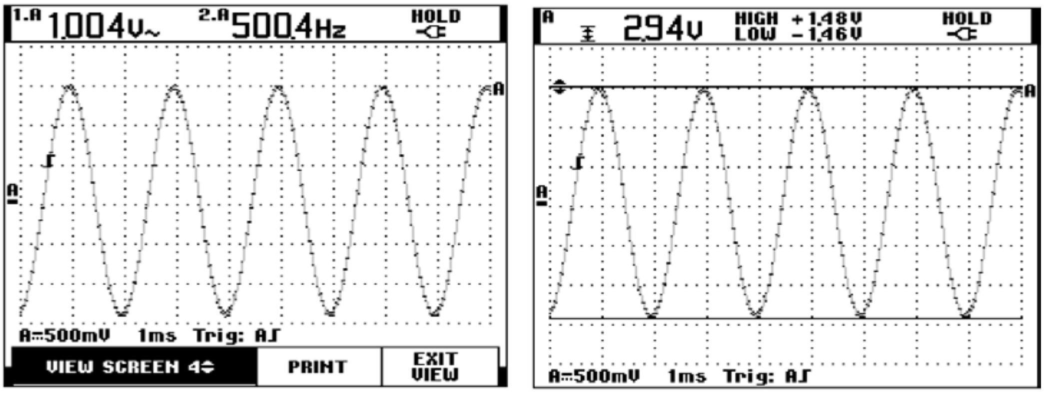
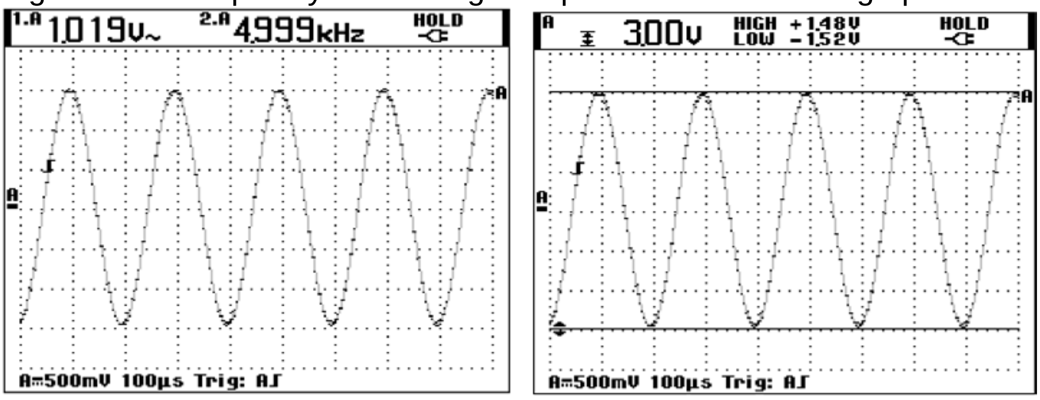


Figure B.8: Frequency and voltage drop across the radiating speaker source for 5kHz.



Each frequency was applied sequentially for at least two minutes and the dBA and dBC values were obtained. The measurements were obtained remotely so the influence of a researcher would be negligible. The air pressure, temperature and background noise were also taken down and tabulated as summarised in the following table.

Table B.1: Measurement 1: Anechoic Area.

Source signal	Distance from source	Height above ground plane	Temp & Pressure	Background noise (dBA)	SPL meter Frequency o/p	SPL meter DC o/p (dBA only)	dBA value	dBC value
3Vp-p @ 995.5Hz	1m	1.3m	23°; 841hPa	37-39	997,9Hz	780mV	**78.02**	78.3
2.94Vpp @ 500.4Hz	1m	1.3m	23°; 841hPa	37-39	500Hz	732mV	**73.2**	77.5
3Vpp @ 4.999Hz	1m	1.3m	23°; 841hPa	37-39	4,995kHz	714mV	**71.4**	70.3

Echoic Area

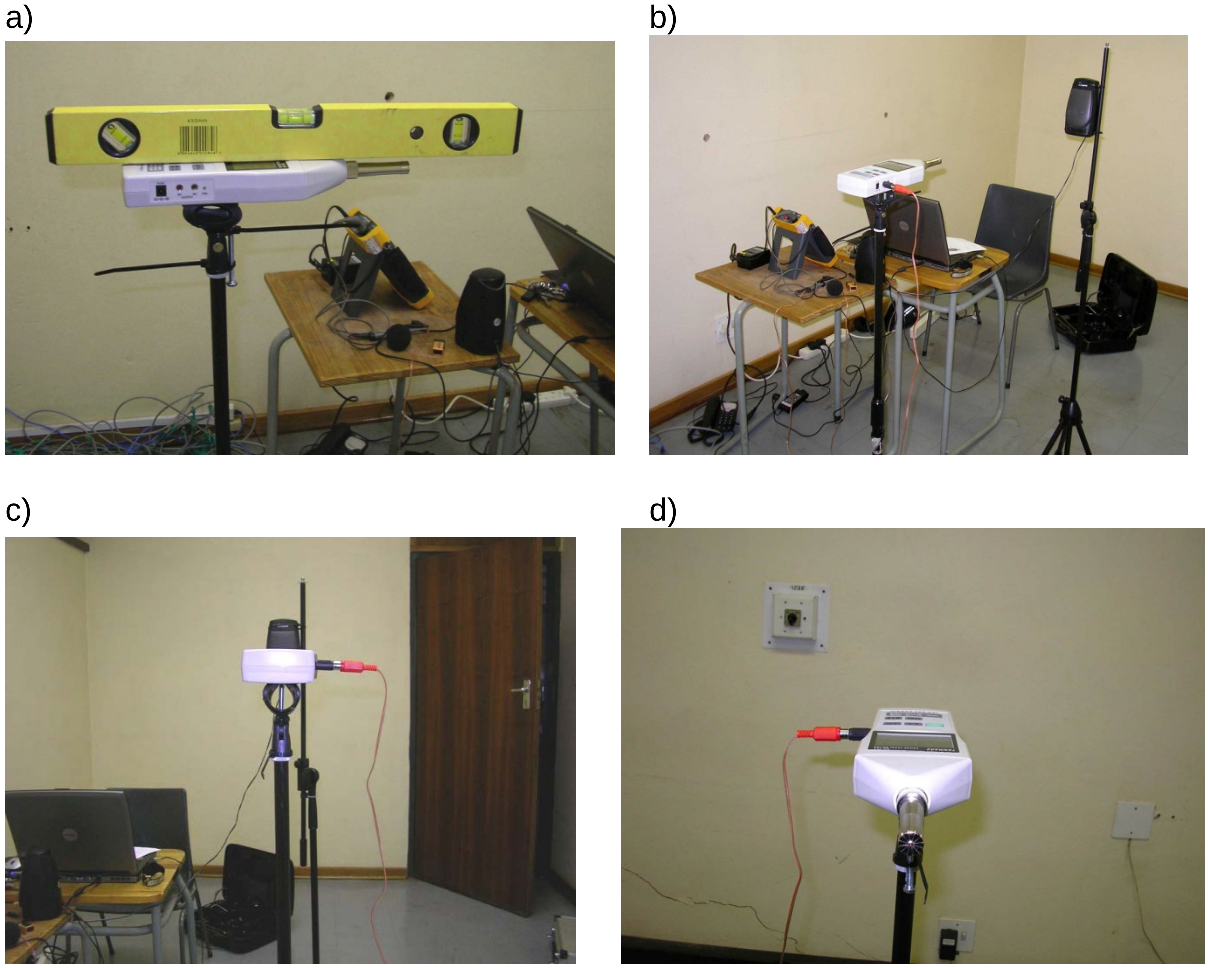
This room has already been described in Chapter 4 under the RT tests for the small room. A room that exhibited a higher reverberation than that of the anechoic area was located. An additional figure (Fig. B.9) is shown of the room and the measuring setup.

Figure B.9: Echoic room.

Equipment setup for the Echoic Measurements

The equipment was set up to match the anechoic layout. The height above ground and the separation distance between source and microphone were the same as for the anechoic setup. The oscilloscope and laptop were also connected as was done for the anechoic readings. The figures to follow show the setup and the room structure with reference to the SPL meter.

Figure B.10 a-d: Equipment setup and room structure from SPL meter's point of reference.



The measurements were taken in the same way as with the anechoic setup. The picture shows the door open but it was closed during the measurements. The waveform printouts are shown next.

Figure B.11: Frequency and voltage drop across the radiating speaker source for 1kHz.

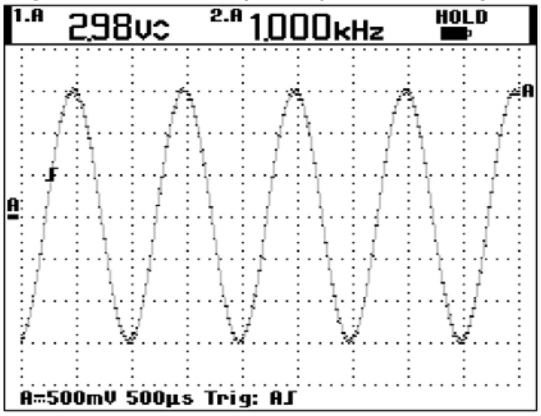


Figure B.12: Frequency and voltage drop across the radiating speaker source for 500Hz.

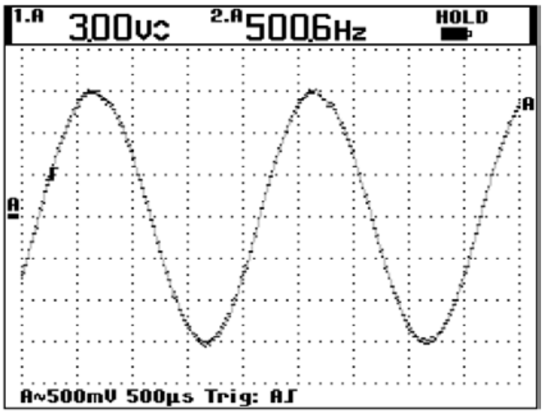
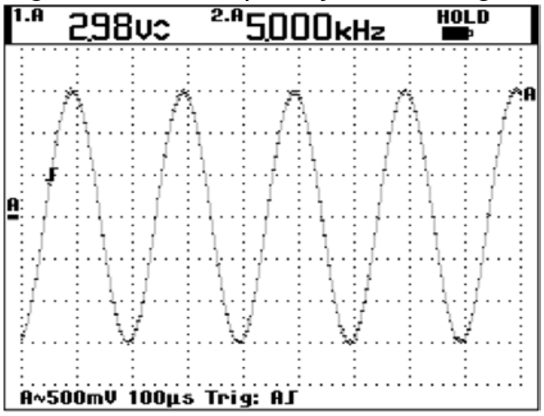


Figure B.13: Frequency and voltage drop across the radiating speaker source for 5kHz.



The results are tabulated in the Table B.2.

Table B.2: Measurement 2: Echoic Area.

Source signal	Distance from source	Height above ground plane	Temp & Pressure	Background noise (dBA)	SPL meter DC o/p (dBA only)	**dBA value**	dBC value
2.98V_{p-p} @ 1kHz	1m	1.3m	22°; 833hPa	34,5-35	845mV	**84,5**	84.3
3V_{p-p} @ 500.6Hz	1m	1.3m	22°; 833hPa	34,5-35	790mV	**79**	83
2.98V_{p-p} @ 5kHz	1m	1.3m	22°; 833hPa	34,5-35	762mV	**76.2**	74.7

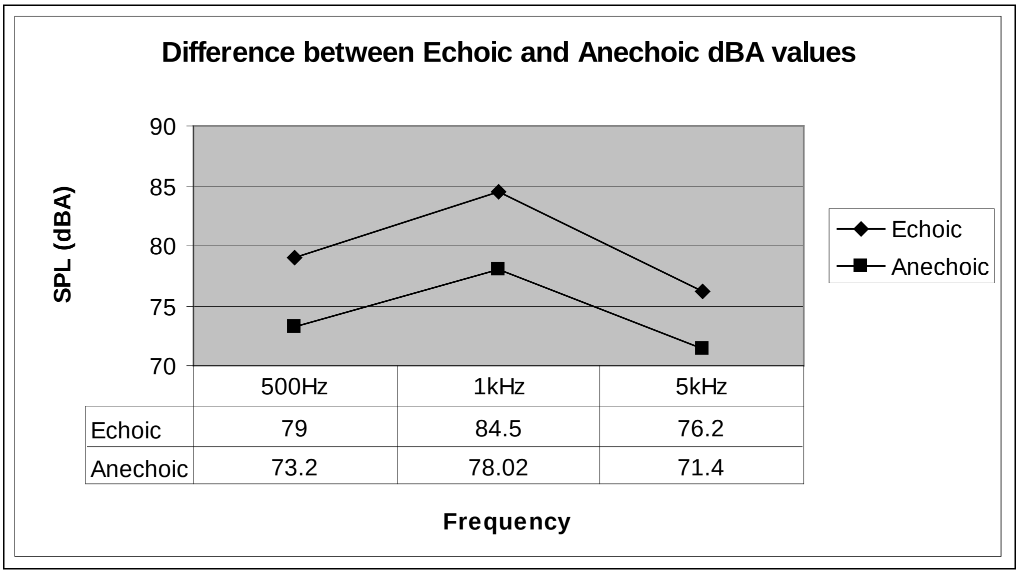
Results Discussed

There was a noticeable difference between the two measurement environments. A difference was expected, however the size of the difference was more than originally thought. The 1kHz signal in the echoic area was 6,48dBA louder than the anechoic measurement. The difference between the dBA and the dBC values were negligible which is what would be expected as both weightings are normalised at that frequency. The 500Hz signal had different dBA and dBC values for both echoic and anechoic measurements. The human ear has a reduced sensitivity to frequencies below 1kHz and thus it is understandable that the dBC value would be higher than the dBA value. The SPL meter may have also affected this difference as the accuracy of the meter is calibrated for 1kHz and has its best tolerance at that value. Thus, values outside of 1kHz are subject to higher measuring inaccuracy. For example, the SPL meter has a measuring tolerance of ±1,5dB at 1kHz but ±2,0dB at other frequency values. With this kept in mind, the measurements for 500Hz were still seen as representing a good description of the difference between A and C weightings. The echoic dBA value was 5,8dB louder than the anechoic value. The echoic room consisted of parallel walls, which probably assisted in creating resonant modes that increased the SPL.

Table B.3: Summary of main parameters measured.

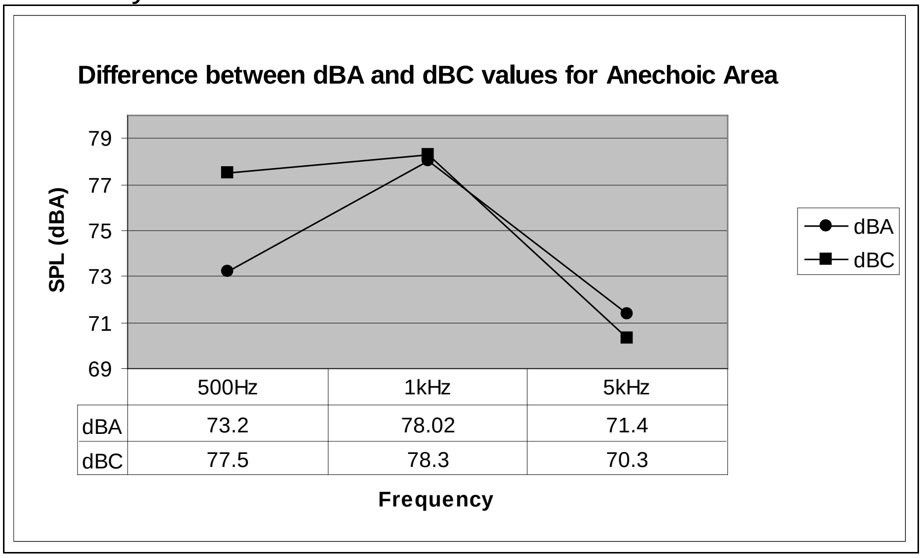
MEASUREMENT TYPE	SIGNAL VALUE	dBA value	dBC value
ANECHOIC	$3V_{p-p}$ @ 995.5Hz ≈ $3V_{p-p}$ @ 1kHz	**78,02**	78,3
	$2.94V_{p-p}$ @ 500.4Hz ≈ $3V_{p-p}$ @ 500Hz	**73,2**	77,5
	$3V_{p-p}$ @ 4.999Hz ≈ $3V_{p-p}$ @ 5kHz	**71,4**	70,3
ECHOIC	$2.98V_{p-p}$ @ 1kHz ≈ $3V_{p-p}$ @ 1kHz	**84,5**	84,3
	$3V_{p-p}$ @ 500.6Hz ≈ $3V_{p-p}$ @ 500Hz	**79**	83
	$2.98V_{p-p}$ @ 5kHz ≈ $3V_{p-p}$ @ 5kHz	**76,2**	74,7

Figure B.14: Graph illustrating the difference in dBA values for echoic and anechoic SPLs.



The graph shows that for each value of frequency, the dBA value for the echoic room was higher than the dBA value for the anechoic area. The graphs show a similar trend with a consistent difference between the echoic and anechoic measurements.

Figure B.15: Graph illustrating the difference between dBA and dBC values for anechoic SPLs only.



Problems Encountered

1. The anechoic measurements were conducted outside on a farm. There was an intermittent wind blowing. The measurements were taken over a two-hour period in order to take consistent measurements. It was interesting to note that even though some wind was blowing that day, the effect of this breeze on the SPL was negligible. The mic's windscreen or windshield was used for the measurements outside.
2. Crime posed a problem. The owner of the farm came and warned me of the danger I was in by the sheer presence of my car, which was parked near the road. I was told to move my car and watch out for thieves who were trying to steel the maize crop. This unfortunately made any further measurements difficult as I was nervous to stay on.
3. When taking measurements in the echoic room, my cell phone interfered with the speaker output even though the cell phone did not ring. I subsequently removed my phone from the room and it was no longer a problem. It was however interesting to see how the radio frequency signal added noise to the speaker's output, which increased the SPL level measured during that instant. This problem was solved early and did not affect the final readings.
4. In the echoic room, the influence of my jacket and my material bag was noticeable on the SPL level measured. When I removed my jacket and bag there was a change in the SPL measured. However, irrespective of whether I was in the room or not, wearing a jacket or not, the SPL level for the echoic measurements was far higher than the anechoic ones.

Limits

1. The relative humidity was not recorded.
2. A noise map was not completed as the anechoic area's background noise was between 37 and 39dBA irrespective of which direction the SPL meter was pointed.
3. A perfect anechoic area has not been found as the area chosen would still exhibit some reverberation owing to the ground, the nearby equipment and the researcher's presence.
4. The echoic room had two classroom tables and a chair present during the measurements.

Conclusion

Acoustically reverberant rooms do increase the SPL of the measured signal when compared to the same measurement conducted in an anechoic area. The increase in SPL is substantial and is mainly due to the additional reflected sound waves present at the microphone.

APPENDIX C

C.1 Comparison between SM58 and SPL IEC651 Type-2 Microphone

Introduction

This section consists of a short comparative test between the SM58 microphone to that of the SPL IEC651 Type-2 microphone used in the RT test and analysis section (Chapter 4). The measurement procedure and test setup was similar to that of the small room's test and was in fact conducted minutes after that test.

Results

No distortion was found on any of the sound impulses recorded from the SM58. The SM58 had a lower signal gain than the SPL meter and thus the gain was increased to match the sensitivity of the SPL meter microphone.

The test was conducted in the small room. A comparison of the RT_{60} and EDT obtained from each microphone are displayed in the next two figures. Note the close tracking of each plot to the other showing a close match between the two responses.

Comparing the RT_{60}

Figure C.1: Comparison between SM58 and SPL IEC651 type 2 microphone - RT60.

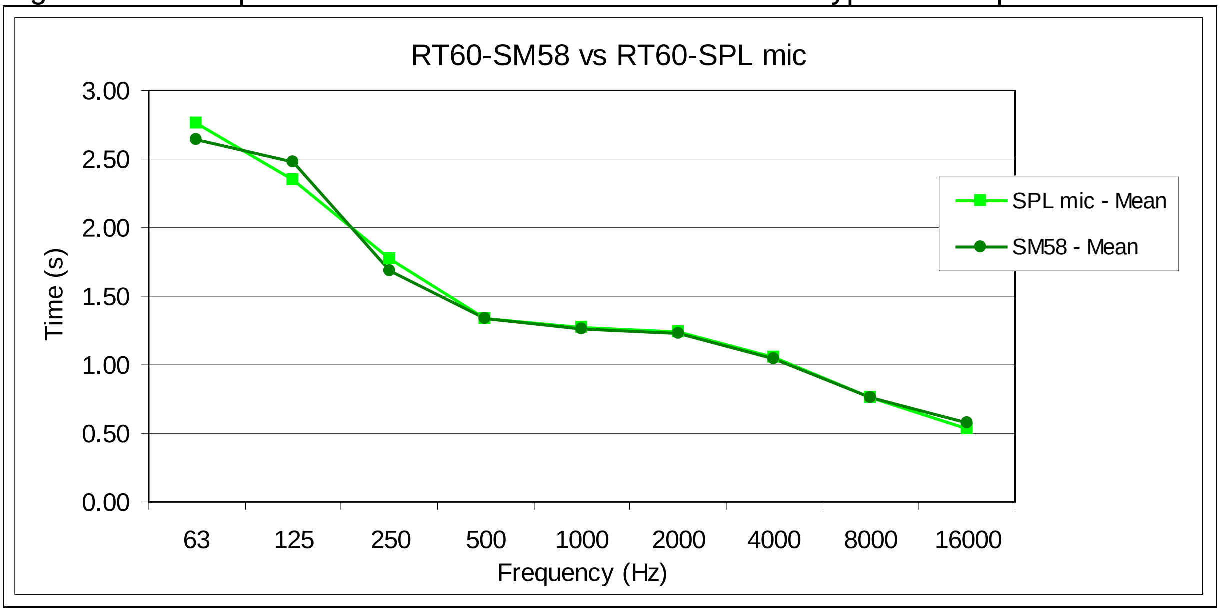


Table C.1: Numerical values of the octave specific RT_{60} plots for both microphones.

x-axis F[Hz]	SM58 – Mean (s)	SPL mic – Mean (s)	Percentage Difference (%)
63	2.64	2.76	4.44
125	2.48	2.35	5.38
250	1.69	1.77	4.62
500	1.34	1.34	0.00
1000	1.26	1.28	1.57
2000	1.23	1.24	0.81
4000	1.04	1.06	1.90
8000	0.76	0.76	0.00
16000	0.57	0.53	7.27

Comparing the EDT

Figure C.2: Comparison between SM58 and SPL IEC651 type 2 microphone - EDT.

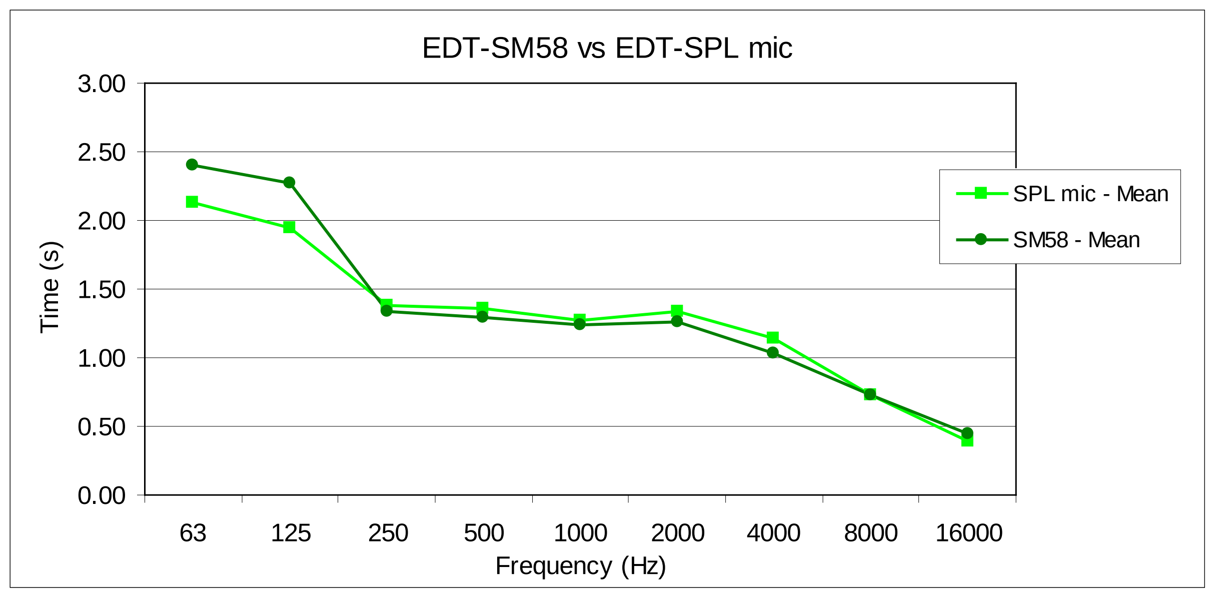


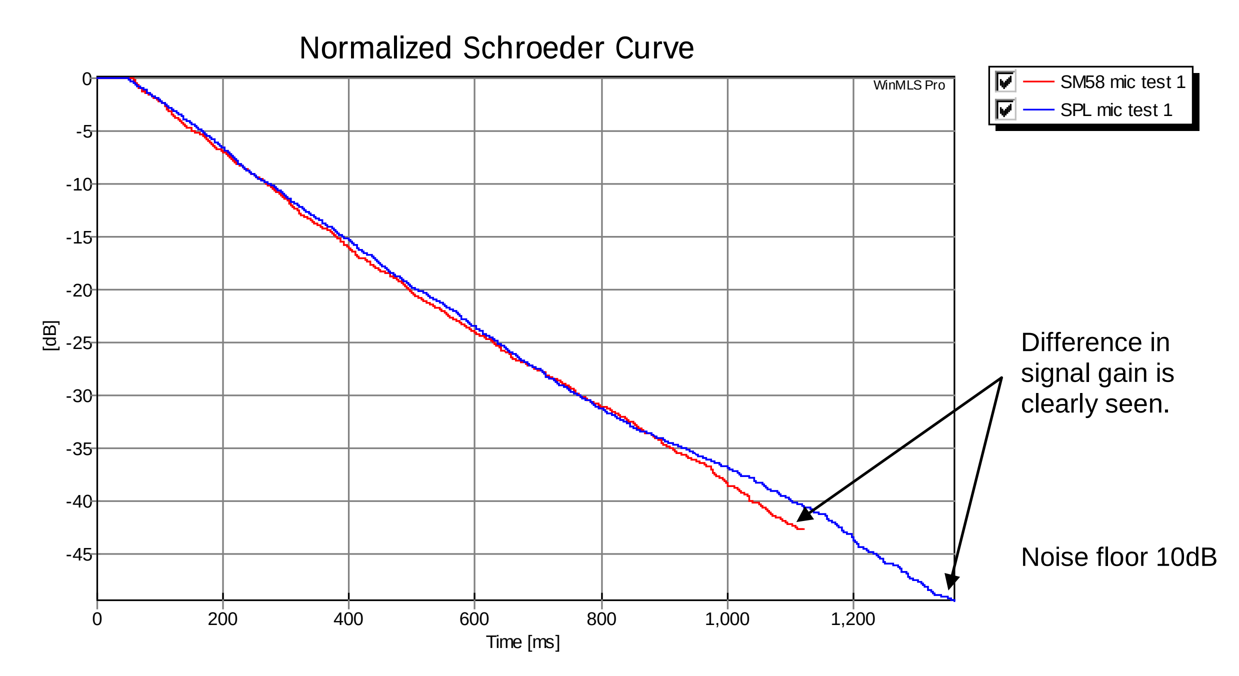
Table C.2: Numerical values of the octave specific EDT plots for both microphones.

x-axis F[Hz]	SM58 – Mean (s)	SPL mic – Mean (s)	Percentage Difference (%)
63	2.41	2.13	12.33
125	2.27	1.95	15.17
250	1.33	1.38	3.69
500	1.29	1.36	5.28
1000	1.24	1.27	2.39
2000	1.26	1.34	6.15
4000	1.03	1.15	11.01
8000	0.73	0.73	0.00
16000	0.45	0.39	14.29

The SNR was found to be good for all measures at each frequency band; however, the SPL microphone had a higher SNR as well as a higher sensitivity. The EDR was high for all readings and the correlation coefficients were all close to -1, thus these results were found to be acceptable in terms of meeting the requirements set out by the analysis software.

The difference in SNR can be seen from the Schroeder curves of two of the impulses.

Figure C.3: Schroeder plot of the two microphones.

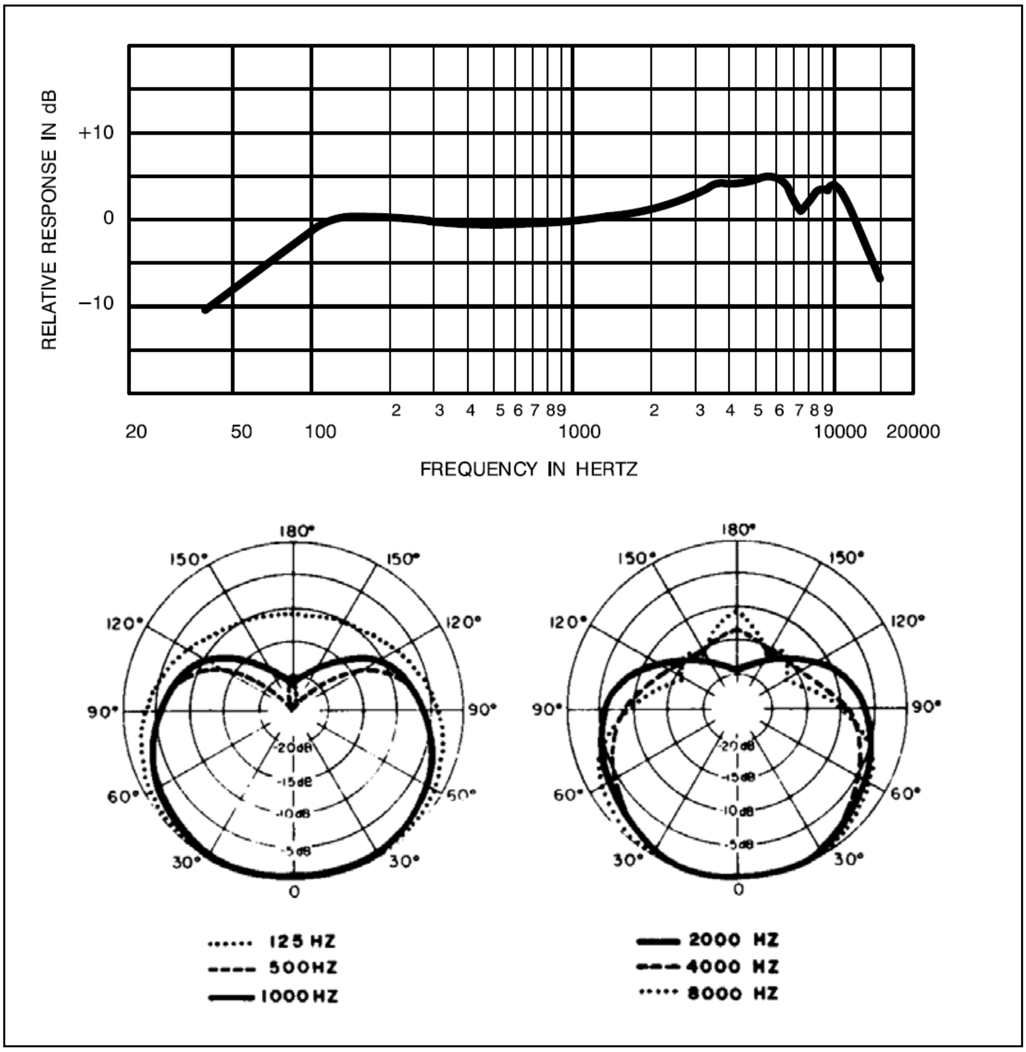


From the above figure, it is shown that the gain obtained from the SPL microphone was superior to that of the SM58, which is what was expected as the SPL microphone has an electronic gain adjustment built in and has a higher sensitivity specification.

Discussion

The SM58's frequency response and polar pattern are shown on the next page with a discussion following.

Figure C.4: Frequency response and polar pattern of SM58 (Shure, 2009).



The pass band has some higher frequency pre-emphasis especially after 2kHz. This is mainly owing to the main application of the microphone being that of vocal use. In terms of RT tests, the increased sensitivity shown in the higher frequency range did not make much difference to the RT when compared to the SPL meter microphone response. In further testing, amplification adjustments were found to have minimal effect on RT. If say for example, frequencies of 200Hz to 1kHz were amplified by 10dB on the raw impulse wave, minimal change in RT was noted (less than 2%). To change the RT, the frequency band needs to be adjusted in their time domain and lengthened. This was discussed in Chapter 5. Thus, surprisingly adequate results were obtained using the SM58 for this experiment. Ideally, the researcher would have liked to use an IEC651 type 1 microphone, but owing to budget constraints, a type 2 was accepted for this first study.

A noticeable difference in the SM58 and the SPL meter microphone was that of sensitivity. The SPLs response was far more sensitive but posed a problem of being overloaded easily. A higher SPL range can be chosen but at the expense of a lower sensitivity at the bottom dB range. Thus, for this experiment the SM58 provided a better dynamic range of measurement with a good SNR. The microphone should be omnidirectional for RT tests. The SM58 falls short in that area while the other testing microphone was omnidirectional. However, the RT_{60} was closely matched, especially for frequency range 250Hz to 8kHz. The EDT was also similar but did show deviation in the lowest and highest octave bands. This is possibly due to the differing polar-pattern and sensitivity.

APPENDIX D

D.1 Qualifying Information of Purposively Sampled Participants Used in the Reverberation Listening Test Questionnaire

Sample Group 2: Purposive

In alphabetical order:

Participant 1: Carol Baron
Carol Baron International Concert Pianist and winner of six Gold Medals in Open Competition. Studied under *Harold Rubens, Adolph Hallis* and *Isidor Epstein* in South Africa. Carol Baron continued her studies in Italy under Arturo Benedetti-Michaelangeli and Renate Borgatti.

Carol Baron, holds several music degrees (D.MUS, FTCL, LRSM, L.T.C.L. UPLM, U.T.L.M.). She had her own Radio Programme for two years in Zimbabwe where she played and sang and gave music appreciation classes on the air. She has appeared in Opera. Carol Baron's career only matured in the 1980's due to the isolation of the then Rhodesia and her first tour was informal performances in the United States, playing to selected audiences in Washington, New York, Michigan, Cedar Falls, Iowa and Newport News in Virginia in 1982.

In 1983 Miss Baron was the first pianist from Africa to be invited to play in the *Peoples Republic of China* and was given rave reviews by the Chinese Critic in Beijing and by the Director of the Shanghai Conservatoire where Miss Baron gave a Master Class and Recital. Miss Baron was invited to the Soviet Union by Professor Mikhail Vokresensky, Senior Professor, of Piano at Moscow State Conservatoire and a concert pianist in his own right, who had heard her play in Zimbabwe, and she gave outstanding recital performances in UFA and Recital and Concerto performances in Kazan and in the Great Hall of Leningrad, where she was given the Honour of Opening the Winter Music Season.

She played *Beethoven's 4th Piano Concerto Op 58 in G Major* and all the critics gave all her performances rave notices. Miss Baron's second Tour of the Soviet Union was in 1988 where she played in Moscow, Kiev, Tallin, Riga and Leningrad and the Southern States of the Soviet Union.

In October1986 Miss Baron had another successful tour of the then German Democratic Republic of Germany (G.D.R.), playing in Berlin, Leipzig, Dresden, Karl Marxstadt, Magdaburg, Gentin and Augustusburg and her second Tour was confirmed for 20-30 January 1989, where she performed Solo Recitals and Piano Concertos with Orchestra.

Participant 2: Daniel Baron
Daniel Baron is a South African born singer/songwriter. He is in his early twenties and is already a celebrity musician in South Africa. He took part in the 2009 *M.Net Idols* South Africa competition where he placed 9th in the competition. His debut single, "Paint My Blue Sky" has taken the radios by storm. Daniel travelled to the US in 2010 to perform and promote himself. Whilst there, he wrote the soundtrack song for upcoming worldwide documentary film, Wrongfully Detained (song and music video still to be released). He has performed internationally and has recently released his first CD. He has been on various South African radio stations.

Participant 3: Steven Baron

Steven has been a disk jockey for several years. He provides his clientele with the latest computer aided DJ methods. Steven has a large knowledge base of various genres of music. He performs for various audiences and thus has experience in how the different equipment types sound in different locations.

Participant 4: Anne de Klerk

Anne is a professional musician and voice trainer. She has worked live performances as well as in studio applications. She has been singing professionally for 17 years. Anne performed the vocal tracks for the subjective testing

Participant 6: Brian De Klerk

Brian is a self-taught musician. Brian has played in a band as well as being an active member of his church choir. He has played various instruments (guitar, accordion, drums and harmonica) but focuses on the piano. He has composed numerous songs which have been used by many musicians.

Participant 5: Ilan Vardy

He is the owner of the home automation company *IVSystems* and performs audio visual design and installation for a vocation. He has had experience in using, testing and evaluating various audio and visual products amongst his other service offerings.

Sample Questionnaire

REVERBERATION LISTENING TEST QUESTIONNAIRE

The aim of this questionnaire is to determine how well a group of people are in their ability to identify an artificially reverberated audio track out of a sample of two normally reverberated tracks.

Participation is voluntary.

Key terms explained: reverberation, artificial reverberation, dead sound, live sound.

Name and surname (optional): ____________________________

Instructions:

Read the question and then circle the most correct option after listening to the audio tracks. An explanation will be given to you during this questionnaire.

Part 1: Information about you

1. Which of the following phrases best describes your view of music?
   - A: I love it and can't live without it.
   - B: I enjoy it.
   - C: I could live without it.
   - D: I am indifferent.
   - E: I don't like it.

2. How much time on average do you spend listening to music per day?
   - A: less than 25minutes.
   - B: More than 25 minutes but less than 1 hour.
   - C: More than 1 hour but less than 1,5 hours.
   - D: More than 1,5 hours but less than 3 hours.
   - E: More than 3 hours.

3. Which of the following phrases best describes your musical ability?
   - A: I play or used to play a musical instrument.
   - B: I play a musical instrument/perform and I consider myself a professional musician/performer.
   - C: I don't play any instrument, but can appreciate it when someone plays one well.
   - D: I like music but am no good at evaluating whether someone is good at music.
   - E: I have no musical ability.

4. Which of the following phrases describes your music listening ability? A good ability would mean that you are able to hear minor changes in sounds while a weak ability means that you are not aware of minor sound changes.
   - A: I have been told that I have a good ear for music.
   - B: I have been told that I have a good ear for music and I have told myself that as well.
   - C: I am useless at listening.
   - D: I am average at listening.

Part 2: Listening skill test

An explanation of how the room size and layout affect reverberation will be presented.

1. Two sound tracks will be played to you. Decide which one is more reverberant.
   A: First track
   B: Second track
   C: Not sure / they sound the same

2. Two sound tracks will be played to you. Decide which one is more reverberant.
   A: First track
   B: Second track
   C: Not sure / they sound the same

3. Two sound tracks will be played to you. Decide which one is more reverberant.
   A: First track
   B: Second track
   C: Not sure / they sound the same

4. Two sound tracks will be played to you. Decide which one is more reverberant.
   A: First track
   B: Second track
   C: Not sure / they sound the same

Three track samples will be played for each of the next questions.

5. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?
   A: The first track
   B: The second track
   C: The third track
   D: They all sound similar

6. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?
   A: The first track
   B: The second track
   C: The third track
   D: They all sound similar

7. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?
   A: The first track
   B: The second track
   C: The third track
   D: They all sound similar

8. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?
   A: The first track
   B: The second track
   C: The third track
   D: They all sound similar

9. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?
   A: The first track
   B: The second track
   C: The third track
   D: They all sound similar

10. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?

A: The first track

B: The second track

C: The third track

D: They all sound similar

11. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?

A: The first track

B: The second track

C: The third track

D: They all sound similar

12. Listen to the three successive track samples carefully; one of them was not recorded in the same location as the other two and should have a different reverberation sound. Which of the three tracks has a different reverberation sound?

A: The first track

B: The second track

C: The third track

D: They all sound similar

End of questionnaire. Thank you for your participation.