Perception and Quantification of Reverberation in Concert Venues

Studying Reverberation Level, Spatial Distribution and Dynamics using Room-Enhancement Environments

Von dem Fachbereich 3 der Hochschule für Musik Detmold
zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN

genehmigte Dissertation

vorgelegt von
Diplom-Tonmeister, Mag. art.

Winfried Lachenmayr

aus München,
Dezember 2016

1. Gutachter: Prof. Dr.-Ing. Malte Kob
2. Gutachter: Prof. Dr.-Ing. Dr. Tech. h. c. Jens Blauert

Tag der mündlichen Prüfung: 17.2.2017

Abstract

Reverberation is an important factor of the acoustics in a room. It influences the acoustic perception of the listener and the performer. Each concert venue has its specific acoustic properties. Numerous studies regarding these properties have been conducted, mostly in real world or fully synthesized environments. However, both acoustic quality and perception in concert spaces are still not satisfactorily explained.

The present thesis contributes new findings in the field of reverberation (late energy) for concert spaces. Previous concepts are further refined and novel approaches suggested. Several experiments are conducted in *semi-virtual* acoustics, namely real rooms whose existing acoustics is altered by means of an electronic reverberation system with loudspeakers. Thus, the possibility of changing the acoustic situation at the push of a button is offered, while the listeners' visual and tactile perception remains the one from the real world environment. A lecture hall and a medium-sized concert hall equipped with enhancement systems are the test environments. Three aspects of reverberation are studied using this technique among others: reverberation level, spatial distribution of reverberation and the connection between signal dynamics and acoustics. The related perceptual attributes reverberance, listener envelopment and perception of dynamics are investigated by means of listening experiments.

Following a qualitative investigation on enhancement systems, it is observed that reverberance depends highly on reverberation level. The method of only assessing decay time is not sufficient. An energy parameter such as strength must be included to predict reverberance. A loudness-based reverberation analysis is further explored and found to perform well in principle, however the three loudness models investigated differ noticeably. The direction of late reverberation in concert halls and the influence on the feeling of envelopment is further specified. Several tests show that the current measure neglects late reverberation from behind and above which contribute to listener envelopment. Lastly, the connection between signal envelope or dynamics and room acoustics is investigated, specifically regarding reverberation. Studies are conducted using, for example, a constant virtual orchestra source or a large pool of audio recordings from concert halls and opera houses. It is observed that reverberation alters the signal dynamic considerably, which is vital both in the context of acoustics and performance practice.

Kurzfassung

Der Nachhall des Raumes ist ein wichtiger akustischer Faktor, der sowohl die akustische Wahrnehmung des Zuhörers, als auch die des Musikers beeinflusst. Jeder Aufführungsraum hat einzigartige akustische Eigenschaften. Zahlreiche Studien wurden hierzu durchgeführt, meist in realen oder vollständig synthetisierten Konzert-Umgebungen. Trotzdem sind die akustische Qualität von und Wahrnehmung in Konzerträumen nach wie vor nur unzureichend erklärt.

Diese Arbeit liefert neue Erkenntnisse zum Themengebiet Nachhall (späte Energie) in Konzerträumen. Bestehende Konzepte werden entwickelt und neue Ansätze vorgeschlagen. So werden in dieser Arbeit Versuche in *semi-virtueller* Akustik durchgeführt, d. h. in realen Räumen, deren bestehende Akustik durch ein elektronisches Nachhallsystem mit Lausprechern verändert wird. So kann die akustische Situation auf Knopfdruck beeinflusst werden, während sich der Proband visuell und haptisch in der realen Konzertumgebung befindet. Ein Hörsaal und ein mittelgroßer Konzertsaal, ausgestattet mit elektronischem Nachhallsystem, dienen als Umgebung. Drei Teilaspekte des Nachhalls werden unter Benützung dieser und weiterer Techniken untersucht: Nachhallpegel, räumliche Verteilung von Nachhall und die Verbindung zwischen Signaldynamik und Akustik. Drei verwandte Wahrnehmungsattribute werden mittels Hörversuchen untersucht: Halligkeit, Umhüllung und wahrgenommene Dynamik.

Ausgehend von einer qualitativen Untersuchung elektronischer Systeme wird beobachtet, dass die Halligkeit stark vom Nachhallpegel abhängt. Die alleinige Betrachtung der Nachhallzeit ist nicht ausreichend zur Beschreibung der Halligkeit, ein Energieparameter wie das Stärkemass muss berücksichtigt werden. Die lautheitsbasierte Nachhallanalyse wird weiter untersucht und scheint grundsätzlich anwendbar. Bei dem Vergleich dreier Lautheitsmodelle werden jedoch deutliche Unterschiede sichtbar. Der Einfluss der Richtung des Nachhalls auf das Gefühl der Klangumhüllung wird präzisiert. In mehreren Tests zeigt sich, dass die derzeitige Beschreibungsgröße späten Hall aus den Raumrichtungen hinten und oben vernachlässigt, die jedoch zur Umhüllung wesentlich beitragen können. Zuletzt wird die Verbindung zwischen Signaldynamik und raumakustischen Einflüssen untersucht, speziell für Nachhall. Versuche u. a. mit einer konstanten virtuellen Orchesterquelle oder einem Korpus an Audioaufnahmen von Konzertsälen und Opernhäusern werden durchgeführt. Hierbei zeigt sich, dass der Nachhall die Dynamik des Signals deutlich verändert, was sowohl für die Akustik als auch für die Aufführungspraxis wesentlich ist.

Preface

The research work for the present thesis has been conducted between 2014 and 2016 at the Hochschule für Musik Detmold (Erich-Thienhaus-Institute), Müller-BBM acoustic consultants from Planegg/Munich and the Aalto University, Helsinki. The project has received funding from the European Commission within the Initial Training Network (ITN) Marie Curie Action project BATWOMAN under the 7th Framework Program (EC grant agreement no. 605867). My gratitude also goes to Müller-BBM for accommodating the project and general support.

I would like to thank my thesis supervisors Prof. Dr. Malte Kob and Prof. Dr. Tapio Lokki for their support. I am also deeply grateful to my instructor at Müller-BBM, Dr. Eckard Mommertz, for supervision, feedback and for general interest along the way.

My fellows from the ITN offered me an interesting view in neighboring disciplines. Collaboration came about especially with Sebastià V. Amengual from HfM Detmold and Alejandro Osses from TU Eindhoven. I very much appreciated the help and collaboration with the experienced concert hall-researchers at Aalto University Jukka Pätynen, Aki Haapaniemi and Antti Kuusinen. The audio engineering colleagues and musicians at HfM Detmold are to be thanked for their openness and support in experiments. The link to the real world of acoustics was kept through my colleagues at Müller-BBM.

Finally, I would like to give my sincere gratitude to Naomi and my parents who greatly helped me along the way.

Detmold and Munich, March 3, 2017,

Contents

Preface 1

Contents 3

1. Introduction and background 7

1.1 Performance venues, acoustics and reverberation 7

1.2 State of the art methods in research for concert acoustics 9

1.3 Intention and scope of the thesis 14

1.4 Structure of the thesis 16

2. Methods and materials 17

2.1 Electronic room enhancement system 17

2.2 Test environments 18

2.3 Acoustic recordings and measurements 21

2.4 Perceptual measurements: listening experiments 25

3. Reverberation level in real and enhanced rooms 27

3.1 Introduction 27

3.2 Investigations on preferred reverberance 35

3.3 Influence of reverberation level on reverberance 40

3.4 Equal reverberance matching 48

3.5 Loudness-based reverberation analysis 56

3.6 Reverberation gain in electronically enhanced rooms 63

3.7 Discussion and conclusion 68

4. Spatial distribution of reverberation and envelopment 71

4.1 Introduction 71

4.2 Small venue: lecture hall with semi-virtual acoustics 74

4.3 Further observations for interaural cross-correlation over time 81
4.4 Medium-sized venue: chamber hall with semi-virtual acoustics 84
4.5 Large venues and comparison between rectangular and vineyard design: reproduced Berlin halls 90
4.6 Influence of different seat backrest height on directional late energy . 101
4.7 Discussion and conclusion 105

5. Levels and dynamics of music in reverberation 107
5.1 Introduction 107
5.2 Initial considerations 110
5.3 Relevance for music performance demonstrated with Wagners' Lohengrin 115
5.4 Constant orchestra source in simulated rooms 123
5.5 Typical levels and dynamic ranges in venues for classical music: analysis of complete concert and opera performances 127
5.6 Level and dynamics when judging "impact" of crescendos in different acoustics 132
5.7 Discussion and conclusion 140

6. Summary and general discussion 143

Appendix 147

Bibliography 155

List of publications 165

List of frequently used abbreviations

L_{eq}	Equivalent Pressure Sound Level
3D	three dimensional
A/D	Analog-to-digital
D/A	Digital-to-analog
DLM	Dynamic Loudness Model
EDT	Early decay time
EQ	Equalizer
Fig.	Figure
FS	Full scale
IACC	Interaural cross-correlation coefficient
IR	Impulse response
ISO	International Standards Organization
JND	just noticeable difference
LEV	Listener envelopment
RMS	Root-mean-square
SPL	Sound pressure level
Tab.	Table
TVL	Time Varying Loudness

1. Introduction and background

1.1 Performance venues, acoustics and reverberation

Performances of music were historically often tied to certain performance venues. For western music, rooms such as churches, opera houses, and later, concert halls served as the spiritual or social framework for these events. Certain types of music were mostly performed in specific venues or even required certain sonic characteristics about the room – concert acoustics. The link between music and room was likely a practical one initially. Spiritual music was performed in churches along with the service, early opera venues were built out of the need for a space to house the performance for several shows and seat paying customers. Early on, an understanding of sound seems to have evolved – always with the objective of enabling the listener to perceive, receive and understand the music well. This is documented through works of literature (e.g. [1]) where music with its emotional impact is studied alongside acoustic phenomena. In addition, other references show the relevance of sound, such as composers of the Viennese classic era discussing the acoustic conditions of venues and changing the orchestra accordingly [2] or depictions of absorptive carpets being hung in churches for important musical events [3]. This practice suggests that music can be enhanced by the appropriate acoustic situation provided by the architecture – a topic that is also investigated by architectural historians [4].

1.1.1 Terminology

The word *acoustics* stems from ancient Greek and describes the science of sound as a whole, as well as sonic properties of rooms. After a sound is emitted by a sound source, the energy starts propagating in the room. When a boundary is reached, the sound energy is reflected via adjacent particles, a process that repeats several times on different surfaces (early reflections of increasing order), see Fig. 1.1. The room acoustic terminology differentiates accordingly between early reflections (discrete reflections) and *reverberation*, the sum of late sound reflections. These late reflections would be reflected several times and are not audible as individual sound events. In this thesis, *reverberation* is used equivalently to *late reverberation*. The root of this word lies in Latin *reverberatio*, the event of rebounding (*re-*: backwards, and *verberare*: to beat). Under ideal conditions with uniformly reflecting boundaries, the energy from the source is the same everywhere in the room beyond a certain distance: the state of a diffuse sound field [5, p. 116]. The measured sound level decay after an impulsive signal is uniform and linear over time. However, this is not the case in reality where the initial decay can differ from the later decay.

Different measures for analyzing the acoustics of a room have been suggested over the decades. The measures are currently derived by recording and processing the impulse of a sound source in a room. By applying time windows, parts of the so called impulse response are weighted. The early energy from the sides, for example, was found to correlate with what a listener perceives as a "widening" of the source. The findings have been standardized in an international standard (ISO 3382, [6]).

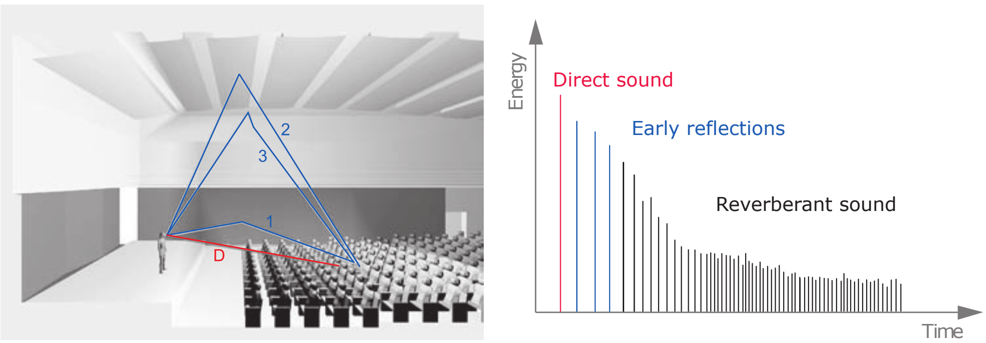


Figure 1.1. Schematic representation of sound built-up, Figure from [7, p. 12]

The relevant sound source in a concert hall: the musician with an instrument, does not emit an ideal impulse, but rather a continuous acoustic signal with varying levels and frequency content. This constantly changing combination of direct sound and reflections now reaches the listener. There are numerous steps involved from the point of sound reaching the listeners ears to the stages of auditory processing and perception. This transition between the physical and perceptual domain shall not be discussed here in detail. The auditory perception can be characterized by attributes: grouped or categorized descriptors of *auditory events* (Blauert, [8, p. 2]).

1.1.2 Perceptual attributes

A number of main attributes exist with sub-attributes and connections between each other, a small hint of the complexity of acoustic perception and quality. Depending on the set of stimuli and the presence or absence of perceptual cues, attributes might be inactive, others might gain more importance or weight. Within the framework of concert room acoustics, the main perceptual attributes are well established through several studies [9] [10][11] [12]. Transitions between groups are sometimes blurred, and attributes change names and weight between studies. For large concert halls 3-6 grouped attributes describe most of the variance [13, p. 470]. An example with individual vocabulary elicitation is shown in Fig. 1.2. For a specific set of concert halls three major attribute groups are found (bold text and red/green/blue), which were *grouped* from individual vocabulary (vertically arranged). Here, the importance and distribution of attributes change somewhat, even between two musical pieces (top and bottom half) in the same study. This variation shows the above mentioned dependency on the stimulus set. Level, spatial, and spectral information together with 1-2 attributes related to the reverberation balance seem to explain most of the acoustic quality. Two of the attributes that are often present and mostly influenced by late sound energy are *reverberance* and *listener envelopment*. There are certainly more descriptors for the influence of reverberation in rooms [14] or for assessment of sound field representations [15]. Vocabulary for describing acoustic perception in rooms of many different sizes, for speech and music, is currently under investigation [16].

Overall, it seems that there is sufficient knowledge to describe auditory perception in concert halls. However, understanding of the individual attributes and related physical measures are still incomplete as is shown later. This issue becomes particularly obvious when new concert hall designs are successful without being in line with the current theory. Christchurch town hall of New Zealand is an example where the acoustical theory was changed after its successful execution. Recently, the Philharmonie de Paris was inaugurated, exhibiting a non-linear level decay as a result of the acoustic design choices (see Section 3.6). This happening does not follow current acoustical practice, where it would be considered an anomaly, but might be good or necessary in that case. In general, there has been a slow but constant change of preferred hall shapes over time, for example, as shown by Meyer [17, p. 318]. This gap between theory and practice is not surprising. Most studies from which the theories and quality descriptors were deducted are either based on simplified sound fields or already established hall designs. Hence, the following section reviews and discusses methods in concert hall research.

1.2 State of the art methods in research for concert acoustics

To gain insight into auditory perception and preference of acoustics empirically, an acoustic situation can be assessed through a listening test. The two established concepts for this are laboratory listening tests, and in situ questionnaires following a performance. As in other fields of research, the laboratory offers more control and freedom for manipulations, whereas in situ testing gives realistic results and

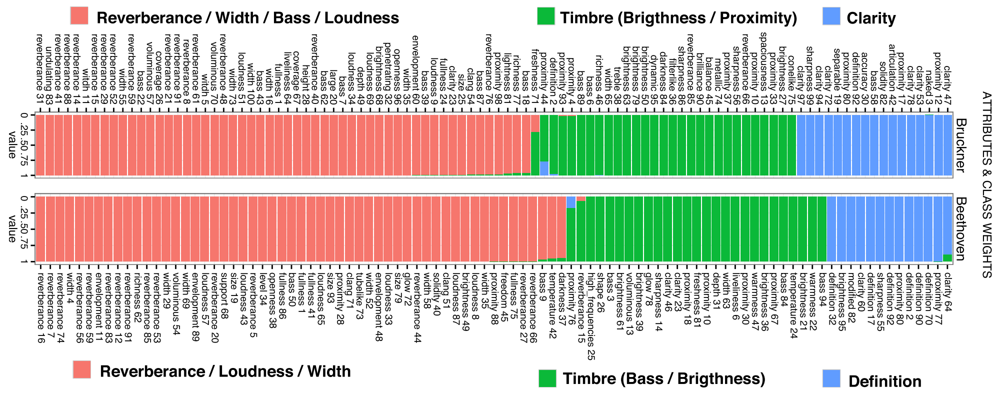


Figure 1.2. Perceptual attributes from a concert hall evaluation test: Words used by the individual participants are arranged vertically. From this vocabulary main attribute groups are deducted (in bold, red/green/blue). Two different music stimuli were tested (top half, bottom half) [12].

more practical correlations. In the context of concert venue research this translates to recreating the sound field of a performance either in a controlled laboratory environment or testing on-site such as a concert hall (*in situ*). A recent overview of studies was given by Lokki ([18], Table 1), showing that all relevant experiments were done using one of these two techniques. Both approaches come with different possibilities of employing a sound source, assessment procedure and capturing or recreating the sound field (see Tab. 1.1).

Table 1.1. Comparison of established methods in concert hall research with typical properties. *Real* refers to the topic under investigation, i.e. an original musical performance in a concert venue.

Testing Method	Laboratory	In Situ
Source	Artificial	Real
Room	Virtual	Real
Sound field	(Re-)Synthesized	Real
...Presentation	Headphone or loudspeaker	–
...Reproduction	Convolution (measured or simulated IR), Recording	–
psbl. bias due to		
...Environment	Simplified, alien	Preoccupating (but real)
...Comparability	– (Variety of stimuli)	Limited stimuli, memory

Studies based on many loose questionnaires without temporal connection to a specific concert give a broad picture of overall quality and opinions on concert halls as well as a lot of data to compare different rooms in terms of physical measurements [19]. However, the approach is scientifically less stringent as both the selection of participants as well as the context in situ and awareness of the space likely biased the outcome strongly. Other more scientifically well-founded in situ studies offer a lot of insight and resulted in important contributions [10]. Yet, it is known that humans are relatively poor at making absolute perceptual judgments, memorizing them and are, furthermore, strongly influenced by other senses. Evaluating a single situation without reference for comparison, possibly including memory, is fallible, when using humans as a measurement tool. Therefore, the approach of in situ testing has some noticeable shortcomings as it is almost impossible to switch to other rooms or sound fields for comparison. Furthermore, the quality of human judgment might be influenced by non-acoustic factors such as public opinion or visual bias.

Laboratory testing was certainly identified as an option, and was used for fundamental perceptual experiments early on. For evaluation of a complete acoustic scene, however, the sound field needed to be recreated properly to avoid bias through presence of artifacts. Noticeable effort has thus been made in capturing and (re-)creating sound fields over the decades through different approaches: Lehman and Wilkens first employed artificial heads (so called dummy heads) for comparing real orchestra recordings from different halls [9]. Individualized, head-tracked, binaural synthesis has since been seen as one promising environment [20, pp. 24-27]. Practical challenges remain with the implementation of head-related-transfer-functions, filtering of the headphones and tracking performance. Room acoustic simulation has undergone major developments recently, making it more suitable for testing as real-time multi-channel performance was made available and practical at the time of writing [21]. However, as of now there is still a gap when comparing the simulation to a real situation [22, p. 187]. For the multi-channel loudspeaker representation with methods such as Ambisonic and VBAP, Nearest-Loudspeaker Panned (NLP), has been proposed as an alternative. This method includes an accurate enough representation of relevant room directions while at the same time avoiding across-loudspeaker-panning artifacts [23]. Although, the method is successful only when combined with the use of an artificial, practically motivated source (loudspeaker orchestra [24]), a good sample set and refined evaluation and test methods [25],[26]. A number of studies were made possible that refined and enhanced knowledge in concert hall acoustics, an overview of which is recently given by Lokki et al. [12].

A realistic production of a test environment has clearly been the target of researchers. Ideally, the test environment should be so realistic that a stimulus cannot be distinguished from the original situation. Yet, realism of the environment was usually neither asked for nor checked formally. While this is hard to achieve, it should ultimately be the goal [1]. Still, in the above mentioned study Lokki concludes [12]:

> [...] the differences between listening to live concerts *in situ* [...] and reproduced room acoustics in controlled laboratory conditions may shift the focus of the subjective observations and, thus, influence the results.

[1] Some proof towards the realism of NLP was shown when listening test participants (recording studio owners) correctly guessed their own resynthesized studio rooms [27].

Though possibly not necessary for fundamental investigations regarding auditory perception, the more complex aesthetic questions such as music perception in concert halls seem to emphasize the role of the quality of the production. Performing a laboratory test with any method, the participants rate what they perceive. For example, participants might judge the reproduction (technology) instead of what is really under investigation. Or, in the view of Blauerts' perceptionist [28]: the current reality of the listening test participant is only the laboratory reproduction.

One possible extension might thus be to combine the two environments as depicted schematically in Fig. 1.3 and discussed in the next section. First steps in this direction have been taken by Kob et al. when implementing a virtual acoustic system in a proper concert hall [29], with one intention being the application for experiments [30].

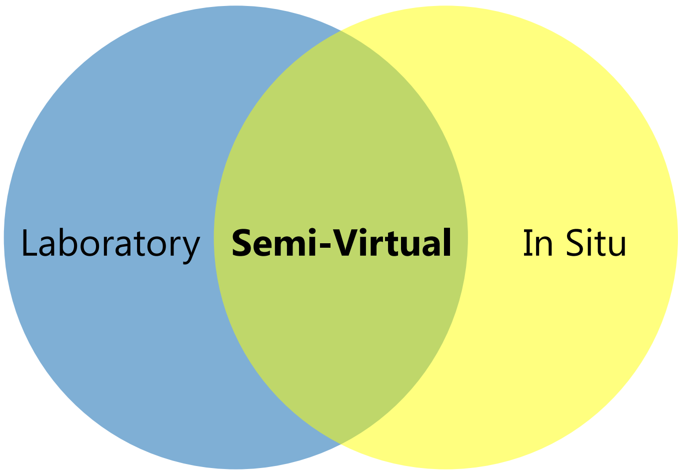


Figure 1.3. Conventional experimental methods in room acoustic research and semi-virtual combination chosen for this thesis.

1.2.1 Semi-virtual environment as a bridge between laboratory and in situ testing

Over the last few years, electronic room acoustics have been applied increasingly outside of research. In its classic form, as a *reverberation enhancement system*, it describes a system installed in a real room, where the sound is recorded by microphones, processed and reverberated, then played back through a great number of loudspeakers typically hidden in the ceiling and walls, see Fig. 1.4. Therefore, additional artificial late reverberation or singular reflections are added to the real room, in other words a *semi-virtual environment* (or augmented reality) is created. Traditionally, the technology comes into use in acoustically *dry* rooms, such as heavily upholstered opera houses, to create appropriate reverberation for the performance of certain musical repertoire. It would then become similar to the acoustic situation of some intended space (e.g. a big concert hall) or a certain sound aesthetics for different genres such as romantic opera. The systems are generally not intended as a replacement or correction of inferior acoustics but as an aesthetically motivated addition. The application is often controversial as some installations of older and newer dates were not sounding satisfactory. Also, the possibility of changing the acoustics by the push of a button is a somewhat frightening thought to some performers.

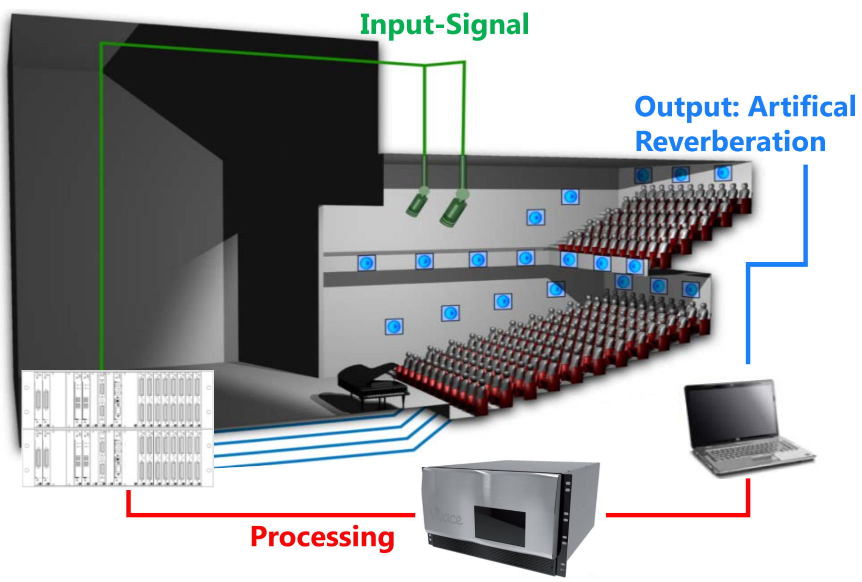


Figure 1.4. Schematic for room enhancement system *Vivace*, Figure adapted with permission from Müller-BBM ASO.

An overview of the history of active enhancement systems is given in [31]. Pioneering systems with analogue technology are shown and current DSP-based techniques explained: Systems with feedback loop (regenerative), without loop (in-line) and hybrid approaches are compared. For the task at hand, the performance and tuning of these systems would be of interest, however, not too much is found about the individual systems or sonic differences. In general, mostly case studies of completed projects are published [32], [33], [34], [35], seldom going into more practical detail regarding challenges faced or tuning procedure [36], [37]. This deficiency is partly because of company confidentiality but also due to a lack of thorough investigation.

There are few scientifically well founded publications with an enhancement system as a test environment and proper listening tests to evaluate the sound field or specific sound field properties. Klatte et al. used a VRAS enhancement system to assess classroom acoustics with pupils seated in the laboratory [38]. However, the fairly dry room with only eight loudspeakers cannot be considered enhanced in the previous sense, but served as big enough dry laboratory for the simulated virtual rooms. Still, it is an example for using an enhancement system in a scientific study. Watanabe et al. [39] enhanced a real concert hall to have a church sound, of which a virtual reproduction was then compared to the target simulation. It showed perceptual similarity between the two; the quality of the enhancement of the room was not evaluated. An attempt to qualitatively compare two different commercial enhancement systems installed in an orchestra rehearsal room occurred [40]. The testing itself happened in a laboratory room with a re-synthesis of the enhanced rooms through ambisonics, not directly in a performance venue as with the real application. Only small differences, or no differences at all, were found between the two systems for criteria such as reverberation quality and within a direct comparison of the systems. Recently, a study of a large group of listeners in a concert hall about preferred sound levels was conducted [41]. Here, a Meyer Sound Constellation system was used to alter the reverberation time, but nothing regarding this topic was discussed in the results.

To sum up, very little formal in situ testing from audience perspective with an enhancement system has been done, partly due to its limited availability. Now, instead of re-creating a sound field in the laboratory the experiments could be conducted in the real room. The possibility of manipulating the room on-site in real time is one of the main motivations for using the technology in this research work. Recording and reproduction for presentation in a test are not necessary, which gives a higher quality for the overall listening experience. Audience members (or musicians) can be in an appropriate setting and are not influenced by the visual or tactile appearance of a laboratory room. Of course, the visual appearance of the concert space can present a bias unless it is part of the test design. The acoustics of the existing room are always present, which guarantees a certain degree of realism for the experiments while at the same time limiting the possibilities, as the existing sound field cannot be suppressed.

The system in use for this work (Section 2.1) allows for the addition of reflections at earlier and later points in time, singular or multiple reflections, with the limitation of a computational delay. As there has already been a lot of research in early reflections, the focus of this work when applying room enhancement systems is its main usage: reverberation (late sound energy) with affiliated perceptual attributes and measures.

1.3 Intention and scope of the thesis

The aim of this work is to review and extend the current state of knowledge on perception and objective descriptors of reverberation using state of the art technology such as reverberation enhancement systems and other forms of auralization. The research focus is on understanding perception and objective descriptors of reverberation with some thematic overlap on early reflections. Three related aspects (see Fig. 1.5):

1. Level of reverberation and reverberance,
2. Spatial distribution of late energy and envelopment,

3. Connection between signal envelope/dynamics and reverberation.

These are analyzed by means of literature study and several experiments applying new and established techniques: for instance, room enhancement technology, namely changing the existing room acoustics to a semi-virtual condition. This technique is presently used as a commercial tool but hardly in the framework of research, as shown above.

Likewise, the analysis of the acoustical signal in the listeners ear is done in other acoustic disciplines but is seldom required for comparison of concert hall acoustics. In this situation, comparable audio material from many different halls is collected as the basis for the analysis, and to enable future assessment of auditory models.

Current state of the art conditions for assessing room acoustics are documented e.g. in ISO 3382-1:2009 [6]. With the parameters defined in this standard, room acoustic quality and related phenomena should be well explainable. It is postulated that this is not the case for predicting reverberance and envelopment. A number of hypotheses can be derived, which have been investigated in the thesis by means of literature assessment and perceptual experiments:

1. ISO Standard 3382 parameters are insufficient for predicting reverberance.
2. ISO Standard 3382 parameters are insufficient for predicting envelopment.
3. If the ISO parameters are not satisfactory, improvements can be incorporated.
4. The connection between room acoustics and dynamics of the signal in the listeners ear is not sufficiently explained by the state of research.
5. Reverberation enhancement and semi-virtual acoustics are useful tools for bridging the gap between laboratory and in situ studies.

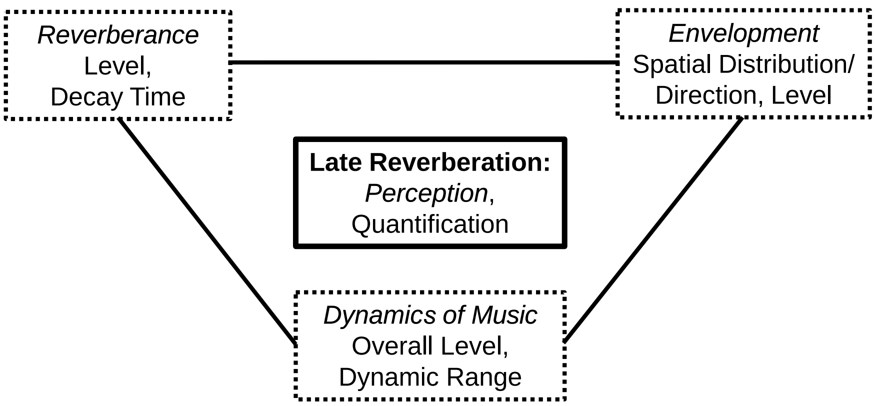


Figure 1.5. Overview of the three chapters of the thesis, centered around the main topic.

1.4 Structure of the thesis

The thesis is structured as follows: the second chapter introduces technologies with methods applied for the experiments within the present work, state of the art for measurement and reproduction of sound fields in room acoustic research.

The third chapter explains the current theory and approach for quantifying reverberation. Next, the importance of assessing the level of reverberation is discussed. The method is further refined to incorporate acoustic strength of the room. A psycho-acoustic approach is evaluated and is tested for a variety of stimuli and three loudness models. As electronic room enhancement is used for several experiments, the advantages revealed and limitations of this approach are discussed.

Chapter 4 focuses on the spatial distribution of (late) reverberation and its link to listener envelopment (LEV). The current LEV measure, derived from somewhat oversimplified laboratory situations, emphasizes lateral energy. This phenomenon is thus questioned and explored for two different electronically enhanced rooms: A semi-virtual lecture room and a semi-virtual chamber hall. Lastly, it is shown that an inclining vineyard concert hall offers less late sound energy from behind and to the side of the listener and therefore less envelopment.

In chapter 5, it is argued that differences in levels and dynamics are an important aspect influenced by acoustics, explaining differences among hall types. A partly historically motivated study of romantic Wagner opera and venue acoustics further encourages a more detailed investigation of music dynamic and signal envelopes: the link between different acoustics and different reverberation, and the dynamic of the music signal is then analyzed by extracting short and long term envelopes of audio signals for different halls and receivers, from actual recordings or simulations.

Finally, chapter 6 provides an overall discussion and summary with major results achieved within the thesis.

2. Methods and materials

The following chapter introduces the technology and methodology used throughout the thesis.

2.1 Electronic room enhancement system

The system in use is a commercial product available under the name *Vivace*, developed and distributed by Müller-BBM ASO. Vivace is a room enhancement or active variable acoustics system, applied to change the acoustics of rooms. These rooms would traditionally be opera houses or temporary concert venues [32]. Up to 192 input and output channels are supported, the setups in this thesis used 64 channels through MADI-protocol. The version of the product used throughout in the following is the convolution based in-line version. The signal is reverberated and played back without incorporating a feedback loop. Vivace impulse responses (IRs) are multi-channel impulse responses measured in real rooms which then were further selected as well as de-correlated. The present sound field of a typical room usually provides early reflections and (late) reverberation already. The environment together with the enhancement could thus be characterized as semi-virtual. Other systems or techniques for generating multi-channel reverberation could be equally suited (WFS, regenerative/feedback loop enhancement etc.). However, the technique and configuration used here is specialized for this purpose and minimizes spectral coloration, also for higher reverberation gains, which were found to be one of the main limitations of regenerative systems [36].

In the general application, the music signal from the stage is recorded similarly to a main microphone approach in classical music recording. It is then processed and convolved in the Vivace processor and played back through loudspeakers distributed in the room. For most studies in this work, a line signal or audio file was used as the input for the Vivace system providing better repeatability, more freedom in manipulating the sound field, and higher suppression of feedback. The control of the acoustic situations is done with a remote laptop. A flow chart is shown in Fig. 2.1, a schematic was also illustrated earlier in Fig. 1.4. The direct sound loudspeaker was accessed with the anechoic file either from the Vivace routing matrix, as shown in the flow chart, or directly from a sound card at the input stage. The system latency introduced by Vivace and the D/A-conversion typically added up to about 6-7 ms.

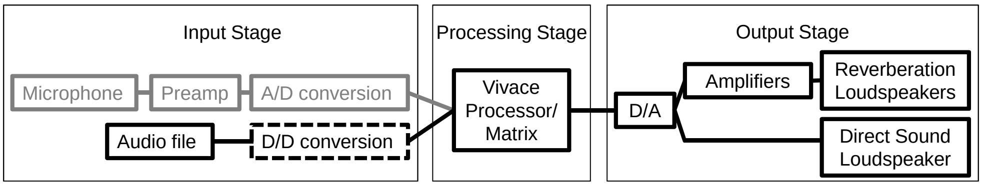


Figure 2.1. Flow chart with the working principle of an enhancement system as used in the thesis (black).

2.2 Test environments

2.2.1 Semi-virtual or enhanced rooms

2.2.1.1 BBM Lecture Hall

The lecture hall at Müller-BBM, Planegg/Munich, has room dimensions of 20.5 x 11.5 x 3.3 m offering a room volume of approx. 770 m^3 (see Fig. 2.2). Together with the absorption of the surface materials, this results in a reverberation time T_{30} of 0.7 s at mid-frequencies, optimized mainly for speech with occasional smaller music performances. The main surface materials are perforated metal on the ceiling, perforated metal and gypsum board on three walls, glass for the remaining wall/facade and wooden flooring. The perforated wall and ceiling panels are built partly sound absorptive and partly reflective. Behind others there are 52 loudspeakers in total, hidden in the dry wall cavity, see Fig. 2.3. These can be addressed individually by the Vivace system through a Nexus matrix. The loudspeakers are K&F Sona5 powered by Lab.gruppen amplifiers, with ± -3 dB points at 110 Hz and 22 kHz. Lightly upholstered variable seating is arranged. The orientation of the listeners during experiments was towards the short side of the room, see Fig. 2.4.

2.2.1.2 Konzerthaus Detmold

The Konzerthaus Detmold is a space for chamber music and small to medium-sized orchestra performances. The reverberation time T_{30} is about 1.6 s at mid-frequencies, rising slightly towards lower frequencies. According to the computer model the room volume adds up to 4150 m^3 (34 m x 18 m x 6 to 9 m, Fig. 2.7), which results in a volume of 7 m^3/person (590 seats). The literature suggests values of 7-12 m^3 per person for music halls [42] and a reverberation time of 1.6 to 1.8 s is generally considered

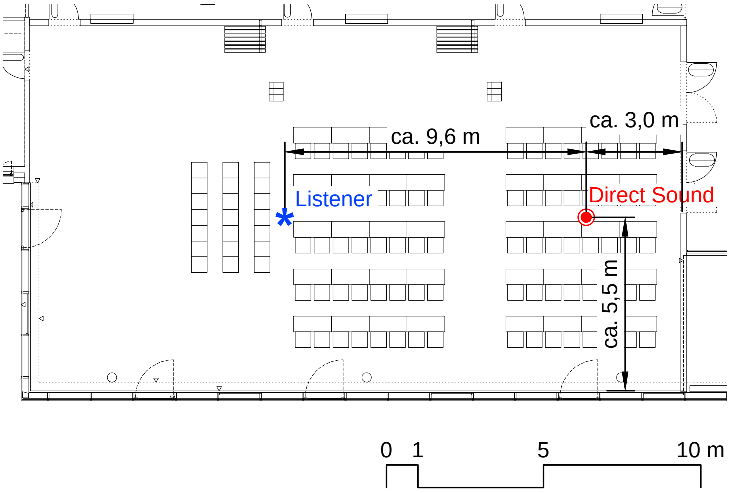


Figure 2.2. Plan of the lecture hall with a typical test arrangement.

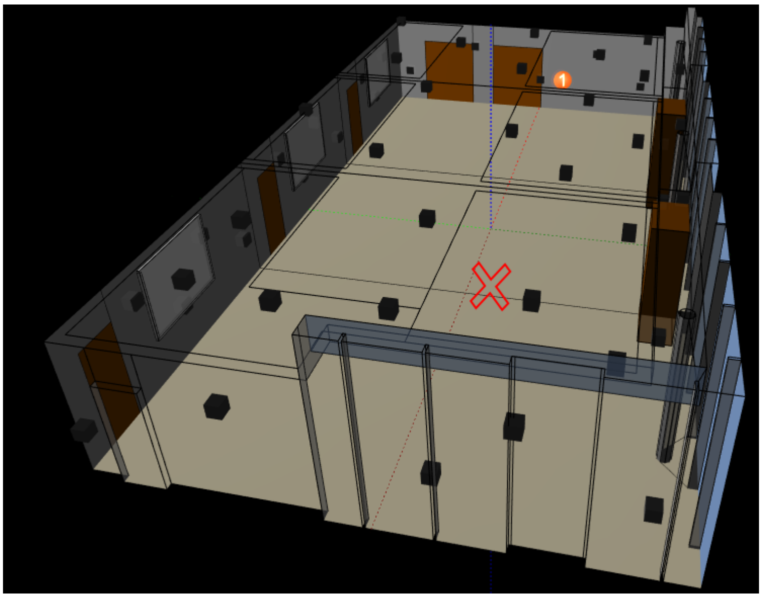


Figure 2.3. 3D-model view of the lecture hall. Black rectangles represent enhancement loudspeakers. The orange circle (1) shows the position of the direct sound speaker for a test, "X" of the listener.

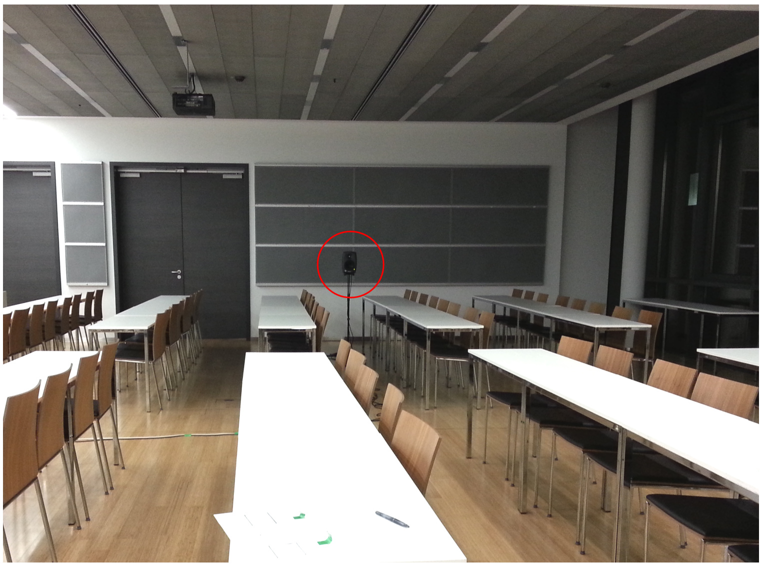


Figure 2.4. Lecture hall with direct sound loudspeaker in the front from a listeners' perspective. Enhancement speakers are hidden behind the grey panels in the walls and ceiling.

appropriate for chamber music [19, p. 536, 551]. Hence the Konzerthaus is built at a suitably lower end of possible sizes. The surfaces of the stage house and ceiling are wooden, the remaining walls are painted concrete, the floor parquet. The lightly upholstered variable seating in front of the stage was present for all experiments (see Fig. 2.6). The remaining seating area is raked with heavily upholstered, fixed chairs.

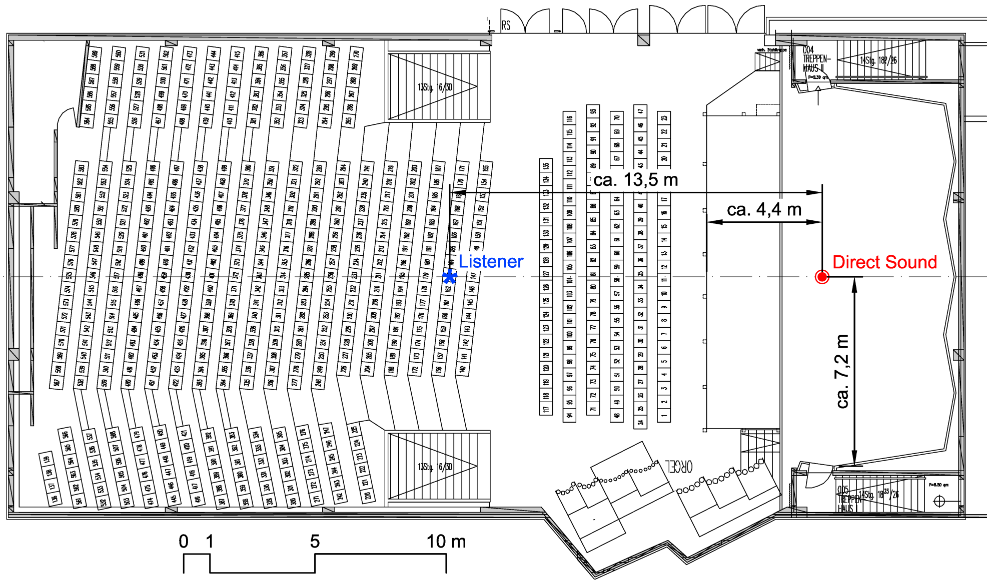


Figure 2.5. Plan of the Konzerthaus Detmold with a typical test arrangement.

Figure 2.6. Panoramic view of the Konzerthaus Detmold with elevated stage, stage house and the organ on the right. Loudspeakers are positioned in a band along the wall and the ceiling.

There are three loudspeaker groups or types: Discrete wall speakers in the rear of the hall, discrete ceiling speakers and the Wave-Field-Synthesis (WFS) loudspeaker band along the walls. Of all these, 64 channels of real and virtually created WFS loudspeakers were used for the output of the enhancement. The resulting distribution of loudspeaker channels can be seen in Fig. 2.7.

2.2.2 Reproductions of real concert halls

When investigating fundamental perceptual differences, a small set of stimuli can be appropriate. However, it may be that the topic under investigation is only explained

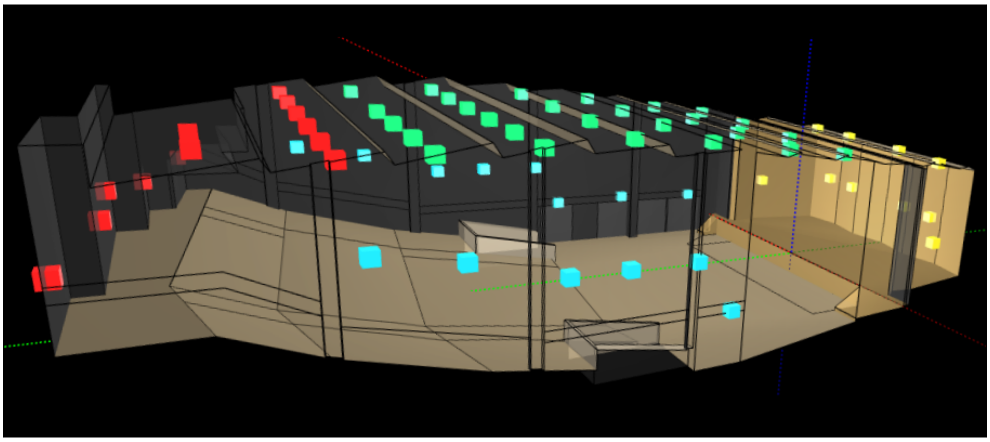


Figure 2.7. 3D-model view of the Konzerthaus Detmold. Rectangles in different colors represent loudspeaker channels used for the artificial enhancement.

well for the given set. Thus, for validating measures and gathering data from a wider set of situations, real world sound field situations from existing concert spaces should be analyzed as well. Where in situ testing is not possible or feasible for the question under investigation, reproductions are used by means of:

- Computer simulations of existing venues (Section 5.4),
- Real halls measured and recreated using the spatial decomposition method [25], Section 4.5),
- In situ recordings of occupied and unoccupied real halls, Section 2.3.1.

Several studies with comparison of multiple venues have been conducted before. In this thesis, it is thus applied only for selected questions. The venues are partly anonymized. All of the venues are measured with the same system and software.

2.3 Acoustic recordings and measurements

2.3.1 In situ recordings

The majority of experiments in this thesis are conducted in situ. In the case of capturing the in situ sound field for later reproduction or analysis, a recording technique has to be chosen. There are a vast number of different recording techniques in audio engineering. In acoustics research, binaural recording has been established. Mainly for practical reasons, two different systems have been used in this thesis to capture the signal at a listeners position.

Figure 2.8. Pseudo-Dummy head (left) and location of miniature microphone on human listener (right): The microphone is marked with red, no tape was attached for the actual recordings.

2.3.1.1 Pseudo-Dummy Head

The first method is adapted from a successful approach with audio recording. This method is the use of a separating disk in a head-like recording setup (OSS technique). The distances between microphones is 18 cm. Two omnidirectional miniature microphones DPA 4060 are used (same model as in the FABIAN head and torso simulator [20, p. 39]). In addition to these two front channels, two additional rear channels are recorded with the internal microphones of the recording device (Zoom H4N), offering a better representation of the horizontal sound field distribution. The sensitivities of microphones of a single system, and among systems, were matched by adjusting the microphone pre-amplification. For this, a reference tone calibrator was used, as well as the diffuse sound field in the reverberation chamber. The setup is shown in Fig. 2.8 (left) and described in more detail in [43] and in Section 5.6.

It is noted that this approach is not a dummy head in the strict sense as the head is only approximated. However, it is necessary to enable the possibility of instantaneously recording multiple receiver locations. The material is not used for experiments that focus on aspects where correct and accurate localization is essential. Also, it was shown that proper dummy head recordings with headphone reproduction are not necessarily superior to other recording- and reproduction methods in reproducing auditorium sound quality either [44].

2.3.1.2 Human head

For live recordings with audience present two miniature microphones DPA 4060 were worn by the author, slightly above and in front of the ear (see Fig. 2.8, right). It was applied to collect a large corpus of in situ data from a multitude of auditoria where, due to practical reasons, minimal visual and practical obstruction were needed. Other systems were not possible to use as they require at least one seat for the measurement device and another seat close by for safety and overseeing. As in the previous setup, this arrangement is not strictly binaural as the outer and middle ear are not taken into account.

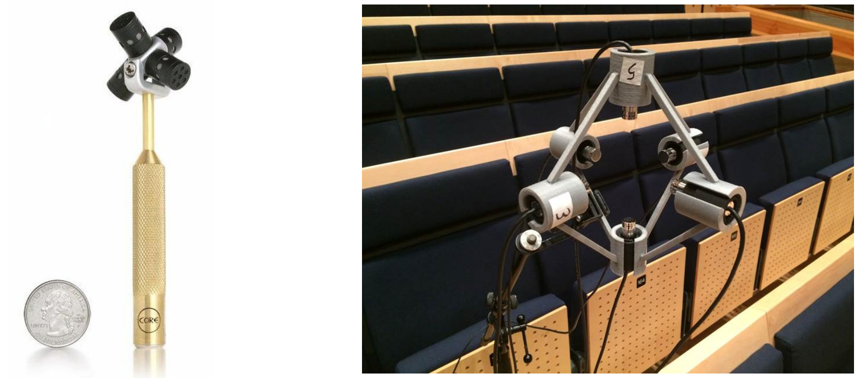


Figure 2.9. Arrays for directional measurements: TetraMic (left, photo from Core-Sound) and 6-ch array (right) with 10 cm distance between capsules

2.3.2 Monaural and binaural measurements

Measurements according to ISO 3382 need to be done with an omnidirectional or binaural system. A class-1 microphone type was used with the software mReverb from Müller-BBM for measurements and calculation. Binaural impulse responses were recorded with the dummy head Cortex MkII which follows the requirements of the ISO standard and a Neumann KU100 which does not comply with the standard but is still well established in binaural technology.

2.3.3 Array measurements

Directional measurements are more widely applied in concert acoustic research in recent years. Computing differences in level and time of arrival between multiple microphone capsules, the incident direction of direct sound, and reflections can be measured. The two systems employed in this thesis are the IRIS system from Marshall-Day (www.iris.co.nz), utilizing a Core-Sound TetraMic (similar to Soundfield microphones, Fig. 2.9, left) and a 6-channel array with omnidirectional capsules (see Fig. 2.9, right), approach used and code developed by Aalto university [25]. Both systems can be used to compute directional parameters such as lateral fraction (LEF). According to the standard, this measure would require a figure-of-eight microphone (directivity), but the computation from the array signals is equally valid as was shown for the IRIS system in [45].

2.3.4 Measures of level, loudness and dynamic

In chapter 5, several experiments explore how the acoustics influence the envelope of the music signal, requiring metrics for level, loudness and dynamics which shall be briefly introduced (recent overview in [46]). Perception and measurement of level, loudness and dynamics are research fields of their own. Here, only the two established approaches, sound pressure level and loudness are considered.

For sound pressure level (SPL), the squared acoustic pressure is given in logarithmic form. Maximum SPL values (L_{max}) as well as percentiles (e.g. L_5) can be measured and, more commonly, integrated and averaged over time (L_{eq}). Different time windows, (sizes of analysis windows) can be set, e.g. 1 s or 10 s for L_{eq} and an integration constant approximating the properties of the ear, 0.125 s (fast) or 1 s (slow). Weighting filters can be applied to the signal, the established curves are all altering low and high frequency content in various degrees. This change is denoted as indexes “A”, “C” and more recently “RLB” weighting curves. Linear frequency-weighting (unfiltered) is referred to as “Z” (zero weighting).

SPL is widely used as a physical indicator, but there are shortcomings: important features of the auditory system such as masking or compression are not accounted for. Hence, the perceived loudness of two signals with the same SPL can be unequal. Loudness N, modeling the different auditory stages, was developed to predict this discrepancy. Different auditory based models exist such as the Dynamic Loudness Model (DLM) and Time Varying Loudness (TVL) with differences regarding the loudness of instationary sounds [47]. The models are potentially computationally time consuming and do not always yield better results for predicting perceived loudness. For music recordings and broadcast content in a large study ([48], [49]) several predictors were compared and loudness measures did not perform better than the simple level measures. This discrepancy was similarly noted in a more recent publication specifically analyzing loudness of music [50]. The results were incorporated in the EBU-R128 program loudness recommendation [51]. Here, also a *Loudness Range* was defined using a RLB-filtered, percentile level range (L_5-L_{90}). In audio engineering the ratio of peak to average level is sometimes utilized (Crest-Factor). Overall, how to measure dynamic range is not clearly and uniformly defined.

In this thesis statistical level analysis and level histograms are employed. The Matlab tool Sound Logger, from the free Matlab toolbox Aarae v8, was used [52] for computing different levels and percentile levels such as L_5 (level exceeded 5% of the time) and calculations of level ranges (difference between maximum and minimum values). Energy-equivalent sound pressure level L_{eq}, L_5 and maximum-minimum ranges were chosen for analyzing signals. More detail and an example is given in the introductory section of the respective chapter (Section 5.1.3), closer to the actual analysis.

2.4 Perceptual measurements: listening experiments

Listening experiments or tests are a way to measure auditory perception by presenting selected stimuli and collecting the responses of subjects. There is a large variety of listening test procedures for different situations. Most of the tests performed in the thesis were designed as paired comparison tests. This indirect scaling method facilitates the judgment of differences between stimuli and can exhibit even small changes. In the task, participants have to decide which stimulus they find dominant regarding the attribute under investigation (typical interface see Fig. 2.10). A third option, "sounds the same/undecided" was introduced for most tests.

Individual judgments are put into a choice matrix and converted into a preference scale. Two types of conversion were applied in the thesis: for an overview on the judgments the relative frequencies following Thurstone's Case V, here called "LCJ Scale", can be plotted. It gives information about the perceived magnitude of the sought after attribute and the distances between stimuli. Relative frequencies are the column sums from the normalized choice matrix [53]. The BTL model computes probabilities for the *maximum* value of the sought after attribute for each stimulus [54],[55, pp. 176]. A Matlab implementation ([56]) and an *R* implementation were used [57]. A Chi-Square test can be performed to test for overall differences in the set. Significance between pairs is checked with a multi-comparison test. Bonferroni-corrected significances, applied to account for multiple comparisons, are only reported for information as the reduced probability level increases chance of false negatives, and seems too strict for the mostly exploratory experiments [58][59].

The smaller the differences between stimuli, the more often listeners might be undecided or unreliable. One suggested method of evaluating the consistency of the data is to check for inconsistent judgments where listeners judged in a conflicting order (*triads*) [60, pp. 489-491]. This consistency coefficient ranges from 0 to 1 and different limits can be set, e. g. a value >0.85 is consistent. However, the smaller the differences between stimuli and the more challenging the test, the more likely inconsistencies can be found. A reason for triads can be the change of internal decision criteria between pairs, which is somewhat expected as has been pointed out by Cremer [13, p. 468]. This issue was discussed with fellow researchers using the method actively in listening tests also related to concert acoustics [12]. When excluding the inconsistent participants, the statistical noise is reduced but the magnitudes of the ratings stay fairly similar. The consistency measure thus appears to be too sensitive, as many subjects would have to be rejected and is therefore not applied throughout.

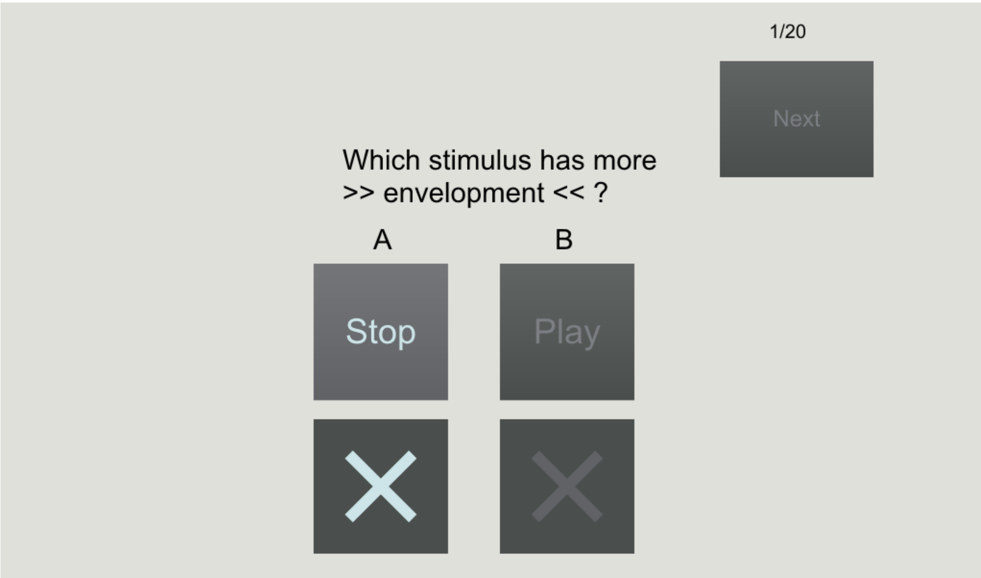


Figure 2.10. Graphical user interface for a paired comparison test: Envelopment, i.e. being surrounded by sound, has to be judged between the stimuli A and B.

3. Reverberation level in real and enhanced rooms

3.1 Introduction

Reverberation describes a physical quantity of the room being filled with energy after excitation with an acoustic signal. When perceived by a human, this evokes *reverberance* among other percepts. As shown in the first chapter, reverberance is considered one of the main auditory attributes in performance venues. Suitable, lacking, or abundant reverberance is noticed by listeners and performers. A number of physical measures exist that were shown to be related to this perception. The most well known measure being the reverberation time introduced by W.C. Sabine. The purpose of the following investigations is to enhance knowledge about the relation between physical measures of reverberation and the perception related to it.

3.1.1 Impulse response measures, reverberation time and level dependency

The standardized method for assessing acoustics of performance venues is described in ISO 3382-1 from 2009 [6]. Here, measurement and processing of impulse responses as the most commonly used method, is defined with the backward integration for deriving decay times. Annex 1 introduces further quality criteria belonging to certain *subjective listener aspects*. For example, above mentioned perceptual attributes, and how to calculate the criteria from the impulse response.

Measures such as the Reverberation Time (T_{30}) and early decay time (EDT) *translate* the decay process into a measure of time. It captures how long it takes for the energy to decrease by a certain decibel value. EDT is listed as more subjectively relevant in the standard and as a predictor for reverberance, as backed up by several studies [9],[61],[62],[63]. However, T_{30} is widely referred to, as in the German standard for acoustics, in smaller to medium-sized rooms [42]. For the second approach, time windows of the impulse response are analyzed and energy relations built. Clarity C_{80}, where the sound energy of the first 80 ms is compared to the late part of the impulse response (IR) is thus a level-descriptor of the balance between early reflections and late reverberation. A stronger late part, or a weaker early part decrease clarity C_{80}. Similar criteria with other time limits are strongly inter-correlated and perform better or worse depending on the situation (C_{50}, C_{50}, T_S).

None of the relative measures give information about the sound pressure level in the room: If the reverberation is twice as loud, the same T_{30} is calculated under diffuse field conditions (see Fig. 3.1, solid blue and dotted blue line). A bathroom might have the same decay time as a concert hall, even though the impression of reverberation is different. On the other hand, two rooms can have similarly perceived reverberation (Fig. 3.1, blue and yellow line) despite noticeably different reverberation times. Strength G_{inf} does give information about the amplification of the room, in the standard it is used to describe the *subjective level of sound* [1].

Several studies have shown that the standardized parameters could be improved by modifications: Late Strength G_{late} (80 ms-infinity) was seen to be important for analyzing hall acoustics [64] and stage acoustics [65]. Strength G_{inf} [11] and late lateral sound level L_J [66] where found to be relevant for reverberance. Also, more frequency bands were recommended to be used for averaging EDT [63]. Reverberance is mostly associated with temporal effects of reverberation. There is some evidence that it is also influenced by spatial aspects [67], i.e. more enveloping reverberation gives a stronger feeling of reverberance.

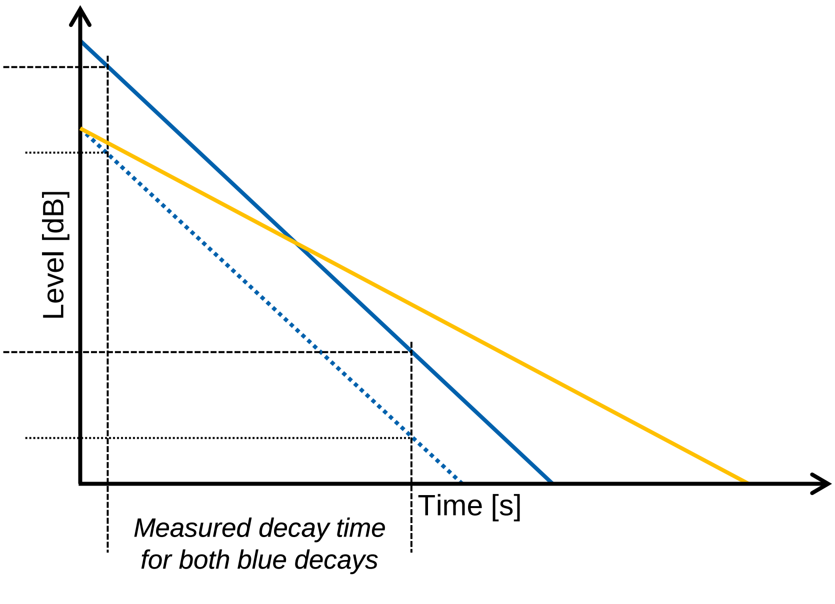


Figure 3.1. Schematic representation of two decays with the same (relative) decay time but different reverberation levels (solid and dotted blue line). Also, a longer, quieter decay (yellow) might evoke the same reverberance as a louder, but shorter decay (solid blue).

[1] Note that the Strength values in this thesis are often not referenced according to standard as the sound power of the source was unknown. In these cases the level differences between acoustic conditions are discussed, not the absolute values.

3.1.2 Frequency characteristics

Spectral differences are regularly elicited in listening tests (see Fig. 1.1.2) and thus certainly relevant, but are only discussed on the side in this thesis. Related terms include *Timbre*, with e.g. *Warmth* and *Brilliance* related to low and high frequency content. For simplification, octave bands are standard for reporting e.g. room acoustic criteria, with the disadvantage of possibly overseeing narrow-band effects. Analysis using fast-Fourier-transformation (FFT) offers more resolution: Recently, time-frequency analysis (*waterfall* diagram known from studio acoustics) has been applied to concert hall measurements [68].

Regarding the spectrum of decay times, a previously cited hall comparison study found a slope of -0.1 to -0.2 per octave [9] for halls to be preferred over other halls with a more flat EDT [2]. In the survey of Finnish halls ([11], Table 5) an increase of ca. 10% for 125/250 kHz compared to 500/1000 KHz and ca. 20% shorter EDT at 2/4 kHz were shown. In the study with several European halls ([12], Fig. 9), it can be noticed that in the stalls the decay times at very low frequencies (63 Hz-octave) are often shorter than at 125 Hz or 500/1000 Hz, which has not been observed previously as this octave was rarely reported. Note that all of these were in unoccupied condition. Similarly, data from Beranek showed a ratio between reverberation times for the averaged octave bands 125 Hz compared to 500/1000 Hz: A mean value of 1.09 (SD 0.18) for 75 *unoccupied* halls is found, namely 9% longer reverberation time at lower frequencies. For 45 *occupied* halls, a bass ratio of 1.11 (SD 0.11, averaged T_{30} for octave bands 125/250 Hz compared to 500/1000 Hz) was found [19, p. 517, Fig. 4.11], slightly increased due to mid-high frequency absorption of the audience but still in the range of 10% difference.

Absorption in the room limits the respective reverberation times. Air and residual absorption is effective mainly in the high frequency range, the audience absorbs sound also in the low to mid frequency range. Thus, by default a longer decay can be expected in the low frequencies. Up to 50% rise at 125 Hz is considered appropriate for orchestral music [69, p. 29]. Together with the recommended values at mid-frequencies of 1.8-2.0 s, this results in a *tolerance area* suggested in different standards [42].

The appreciation or acceptance of the increase in low frequency decay was partly attributed to the lower sensitivity of the auditory system in this area ([13, p. 368]). When utilizing loudness models for decay analysis, as shown later (Section 3.1.3), this property as well as temporal and frequency masking would be accounted for. An additional possibility is to filter the input signal according to the source/stimulus spectrum [70]. For example, with a generalized music filter or standardized curves

[2] However, the exact slope of EDT spectrum was dependent on the music presented and was not well explained by the statistical model.

such as IEC-filter. For certain tasks, such as the detection of echoes from IRs this was suggested as well [71] and better performance was indeed observed recently by Rauber and this author (successive work to [72], unpublished). In the acoustic assessment of open offices a speech filter is nowadays utilized to better characterize the actual source (ISO 3382-3, 2012).

The strength criteria give information about the absolute frequency content of the impulse response. *Warmth* is supposed to be best predicted by Strength G at low frequencies [19, pp. 512-513], but not all relations are as simple. All frequencies are attenuated over the distance and more so higher frequencies.

3.1.3 Approaches including auditory modeling

Even with level information considered there might still be insufficient prediction regarding how sound is perceived by the listener. Impulse responses describe the system (room) without considering the signal (music). Properties of the auditory system such as compression, masking etc. are not accounted for either. Lately, psychoacoustic and room acoustic research are approaching each other. This collaboration is partly motivated by the need for better algorithms in challenging acoustic situations, as the understanding of room acoustics cannot progress without taking into account hearing and perception.

3.1.3.1 Impulse response based approach

A combined approach of system and signal domain has been proposed by Lee et al. where calibrated impulse responses are put into a loudness model [73], followed by a decay calculation similar to conventional reverberation time. Fig. 3.2 shows three loudness-processed impulse responses schematically. Compared to the previous Fig. 3.1, it can be seen that the initial decay is more smooth. The high early energy is perceived as being louder than the late energy due to model's level sensitivity. Compression leads to the previously identical decay slope of the louder impulse response (solid blue) to be slightly more flat than for the quieter but otherwise similar IR (dotted blue).

The method showed good results for several experiments where listeners matched the reverberance of different stimuli. Calculating a reverberation time and early decay time (T_N, [73] and EDT_N, [74]) clearly showed the influence of level on reverberance. For calculation, the reference IR was given a fixed level of L_{AFmax} = 75.5 dB as an input into the loudness model. EDT_N is then the early decay time of this "Loudness-IR" [75]. After re-analyzing several experiments, a simpler relationship was found with T_L or EDT_L using only the listening level as a simple correction-factor to T and EDT, yielding similar or even better performance than EDT_N [76]. The previously mentioned experiments were done on headphones only and have not been evaluated in further practical scenarios as yet. Also, information about reverberation level as measured by strength G for example was not discussed as a "competing" parameter. Thus, these new approaches need further validation. Also, the additional prediction efficiency needs to be weighed against the "cost" of making necessary assumptions and additional computation.

T_N performed well for predicting reverberance in a different study where the reverberation time of simulated rooms of different volumes was kept constant by adjusting the absorption [77].

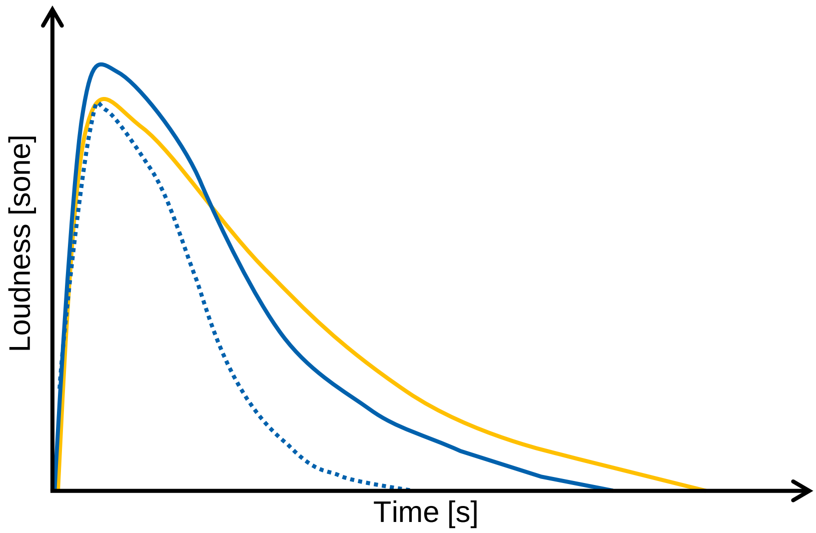


Figure 3.2. Schematic representation of three impulse responses after loudness processing. Compares to decays in Fig. 3.1.

3.1.3.2 Signal based approach

The recent progress in auditory modeling is focused around speech intelligibility, localization and de-reverberation, topics all related to reverberation. One study, has specifically investigated the use of an auditory model for assessing concert acoustics from music in rooms [78], [79]. Four parameters related to ISO such as perceived clarity, were deducted from a calibrated binaural recording and outperformed their impulse response-based counterparts. The model was trained and optimized with different sets of data, both in rooms with distinct differences and rather simple stimuli (solo cello or single speaker samples). The model analyzes a binaural recording and extracts dry/early and reverberant streams using peak detection. Instead of actually recording on-site, a hybrid approach is possible by measuring IRs as usual and convolving with the given reference stimulus. The work seems to be a first successful attempt, which needs further application and testing before utilizing it for empirical studies. By default, the model is not freely available and open. This fact, and the overall lack of experience regarding the usage, discouraged its use for the investigations planned in this thesis.

3.1.4 Slope of the sound level decay

Previously it was argued that the sound level decay is uniform and linear under ideal diffuse conditions. However, in practice there are several situations where this might not be the case. When an additional room volume is connected or coupled to the main volume, the reverberation energy built-up changes as the sound from the coupled volume is reflected back into the main volume. This effect was incorporated deliberately in several acoustical designs as a means to increase reverberation (so called reverberation chambers or involuntarily e.g. for opera stage towers). A theoretical description has been derived [13, pp. 219-224]) as well as experimentally investigated [80], where decay models were compared with scale models. Furthermore, listening tests were conducted showing a noticeable effect only with the maximum setting of 10% coupling surface area. The main volume had a reverberation time of T_1 = 1.4 s, the coupled volume had a T_2 of 2.2 seconds [80, pp. 126-129]. Another interesting outcome was that double-slope decays are preferred over linear decays for solo instruments/choir but not so for orchestra. A very recent example for reverberation volumes is the design of the Philharmonie de Paris [81], [82]. Reverberation energy from coupled spaces are said to be most noticeable after full stops or pauses. Whereas, for *reverberance*, the reverberation audible throughout the music is more important (running reverberation). Thus, reverberation volumes are controversial among practitioners considering the costs and feasibility (see also [19, p. 505]).

But even in a room without reverberation volumes deviations can be found between early and late decay. Measured differences between EDT and T_{30} of up to 20% were found for 17 unoccupied halls (hall means), with an on average 6% shorter EDT. This value is slightly larger than the JND of decay times [83]. The effect can be partly attributed to early reflections changing the decay and overall insufficient diffusion (reflective walls, absorbent floor/audience)[84]. O'Keefe [85] also analyzed EDT/T_{30}-ratio for one hall before and after renovations. In this case the goal was to reach higher EDT/T_{30}-values by directing reflections onto the audience to create more reverberance in the previously too "dry" music theater. It was also noticed that EDT/T_{30} is strongly correlated with the ratio of room height to room width.

In conclusion, there is some evidence that slightly non-linear level decays are common in concert halls and possibly even desirable. The studies discussed above did not consider loudspeakers for creating reverberation. This fact is important to point out as the scenario of a room enhancement system manipulates the reverberation decay. Different from the reverberation chamber, the gain of an enhancement system can be much higher (especially for systems avoiding the feedback loop). The loudspeakers are more directed onto the listener and manipulated easily. Some theoretical consideration was conducted after which an enhancement system is compensating or overcompensating for the present absorption surface [13, pp. 230-231].

3.1.5 Electronic room acoustics

The possibly non-linear level decay of enhancement systems has not been documented for actual installations and is investigated in the following sections, specifically in Section 3.6. In the first chapter, the basics about enhancement systems were introduced and a brief literature overview given (Section 1.2.1). Further questions remained unclear:

3.1.5.1 Is an enhancement system beneficial?

Among professional musicians and cultural managers/orchestra directors electronic enhancement is seen controversially. At a recent orchestra manager conference (Deutscher Orchestertag 2015), an informal poll by the author showed that around 50 % of the directors would be accepting towards electronic enhancement as a tool and if necessary. The fact that the acoustics of a room could change depending on technical gear is worrisome to some professionals (against the time invariance of the acoustics). A world-class conductor, Valery Gergiev, mentioned that he needs the room to be absolutely constant and have no doubt about it[3]. According to Beranek [19, p. 505], other musicians found the innovation desirable. Some might have heard a bad installation, others are opposed by the idea of electronic manipulation. These factors are difficult to evaluate and not the subject of this study. The other main point of critique is the artificiality of the sound, unnatural or detrimental reverberation which can be due to different factors.

A high gain of the enhancement system can potentially lead to an increasingly non-uniform decay, which becomes somewhat unnatural as most halls have only small differences in the decay slope (discussed in the previous section). Also, by lengthening the decay successive notes are less separated over time because the initially desired increase in running reverberation (“carrying the tone”) can become too much when clarity and definition suffer. Two more phenomena appear which are not entirely the focus of this study. When amplifying and thus overlaying reflections from the system and the room, comb filtering can occur due to interference which is generally not appreciated (spectral distortion). Similarly, audible feedback of single frequencies is detectable as artificial.

On the other hand, these systems are used more and more often to alter the acoustic

[3]This statement was made verbally and aimed at mechanical variable acoustics such as curtains running behind perforated panels. But just as curtain motors can stop working, the fear exists that other technical equipment could stop as well.

situation. If the change is well realized, and not too obvious, the systems seem to be gladly accepted as a tool. The number of installations in the last 20-30 years backs this up [31]. Although, a scientific preference test has not been conducted.

3.1.5.2 Return from experience

As mentioned in Section 1.2.1, the configuration procedure of an enhancement system by the system engineer is not clear. Final adjustments would often be done in agreement with the conductor or audio engineer of a specific production. Sometimes manufacturers publish a graph with different T_{30} reverberation times achieved in a certain venue, but it is hard to judge the quality of sound from these results. David Griesinger, developer for Lexicon reverberation, mentions interesting experiences regarding setting up LARES systems on a few occasions:

> The Staatsoper has a natural reverberation time at 1000Hz of 0.9 seconds, and Barenboim wanted something closer to Bayreuth, 1.7 seconds. ... I adjusted it till I liked it with my own singing [...] but Barenboim was NOT delighted. Horrible, he said. Good on the orchestra, horrible on the singers. [...] I installed a shelving filter in the microphone inputs, which reduced the reverberant level – not the reverberation time – by 6dB above 500Hz. Barenboim was delighted. [34]

> We carried a remote control that allowed us to vary the D/R [Direct to reverberant ratio] in half-dB steps. We lowered the D/R gradually, and the sound took on definite richness and depth. [...] But at one point Peter said STOP – that is too much! I could not hear the difference. Listen, he said. With that one extra half-dB the singer moved back 10 feet! [...][34]

> The music director in Amsterdam was conducting in the pit when I chose to raise the reverberation level by 0.5 dB. He immediately waved to me [...][86]

> The microphones are close enough to the stage that we can achieve independent control over reverb time and reverb level. [...] that the optimum reverb level for speech is about 6dB lower than the optimum for symphonic music. Opera requires intermediate values, with dialog being close to speech, unaccompanied singing requiring about 2 dB more, and accompanied singing about 2 dB more than that [...] The time-energy curve of the reverberators has been tailored to provide relatively high RT-20 vs. RT-60, high diffusion,[...] [36]

Comments such as these and from another system engineer Gunter Engel (Vivace) deliver numerous hints that the increase in reverberation level is more audible or critical to set than the reverberation time. However, measurements and verification by formal listening tests are mostly missing. Different hypotheses could be derived, for example, whether the system gain or reverberation level can be varied without changing the reverberation times in a real room with the enhancement system. Also, can these presumably small differences be in fact perceived and measured? There seems to be a rather slight margin when too much reverberation is present in an enhanced room and is quickly doing more harm than good.

3.2 Investigations on preferred reverberance

In this experiment, the rating of different gains of enhanced reverberation is investigated. Different amounts of artificial reverberation are added to the sound field of a real chamber concert hall and analyzed regarding decay times and strength. Experts trained in sound balancing, and laymen, were asked in an in situ test if the presented auditory situation would aesthetically need more reverberation, sounded appropriate or too reverberant. One possible outcome could be a general dislike for the artificial add-on.

3.2.1 Setup

The listening position in Konzerthaus Detmold was located at a distance of 19.7 m from the sound source, see Fig. 2.7 (room description on p. 18). The loudspeaker on stage (type Neumann KH120A, height 1.5 m) was repeatedly playing a violin sample, recorded anechoically with a piezoelectric microphone (usage permitted by Prof. Mores, HAW Hamburg). The violin piece was a solo work by J. S. Bach, Partita No.1 in B minor (BWV 1002), 6. Double, with a duration of ca. 18 s (see Fig. 3.3).

Figure 3.3. Score for the violin excerpt used as a test stimulus. Edition Breitkopf und Härtl, from IMLSP music library (public domain).

The enhancement system Vivace was used to create artificial reverberation with the audio file as a line input source. A Vivace preset was used with digital reverberation times of 2.1, 2.3, 2.4, 2.5, 2.1, 1.7 seconds for the octaves 125 to 4000 Hz. Seven acoustic conditions were selected: the real environment and six artificial reverberation settings with steps of 2.5 dB gain, covering a range of 12.5 dB digital gain. Energy before 80 ms was not altered noticeably with 0 dB change above 500 Hz and at most 1.5 dB at 125 Hz. The resulting measured reverberation times and late strength values can be seen in Figs. 3.5 and 3.6. The frequency excitation of the violin signal is limited by the lowest note of the violin at approx. 200 Hz (most energy in the anechoic audio signal was around 500 Hz). Thus, the reverberation at lower frequencies is not excited.

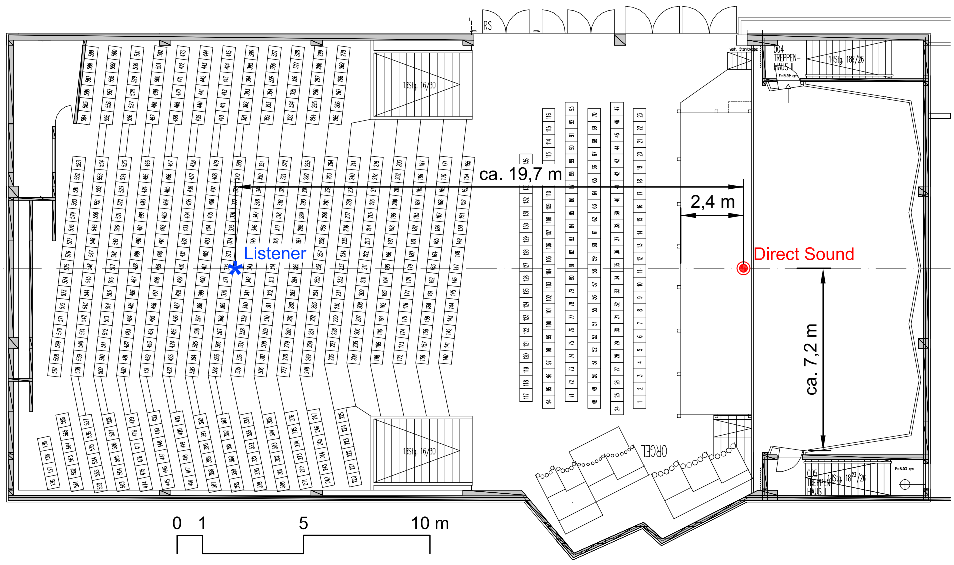


Figure 3.4. Groundview of the Konzerthaus Detmold with the experimental setup.

24 volunteer participants completed the test of which 9 were students of the audio engineering program, of advanced experience, and can be considered expert listeners. The other 15 participants were normal listeners. The test was introduced outside the concert hall orally from a written template. It was explained that different (room) sound situations would be presented with a reproduced violin playing in it and that it should be judged "off the top of one's head" if the presented scene could use more reverberation, sounded appropriate/good, or had too much reverberation. The scale included two intermediate steps, a total of 5 values. The participant was led into the dimmed hall to the listening seat with a sleeping mask over the eyes to prevent visual bias for the whole experiment or, if discomforting, with the eyes closed. No extra training was given to get an uninfluenced first opinion and to avoid familiarization with the room. For the same reason, each sample was tested only once. The stimulus could be played again if desired. The stimuli were played one after the other in randomized order with the participant giving the judgment orally.

3.2.2 Results

There was a statistically significant difference in the judgment for preferred reverberation, $\chi^2(6) = 75.093$, p < 0.000. Post hoc analysis with Wilcoxon signed-rank tests was

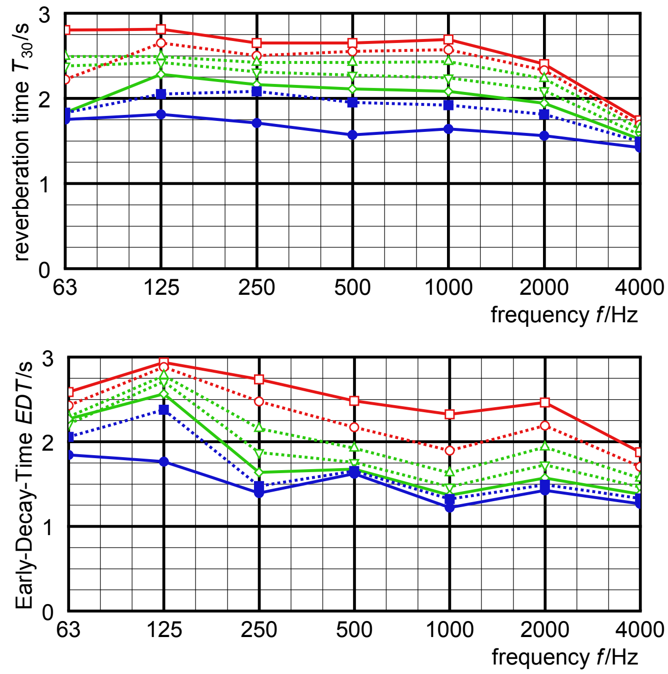
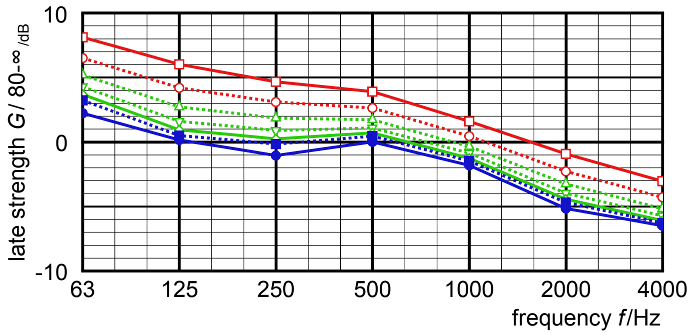


Figure 3.5. Measured reverberation time T_{30} (top) and early decay time EDT for the seven acoustic conditions, with and without electronic room acoustics system. Colors indicate the overall judgment in the listening test (green: “good, appropriate”).



8.1	6.0	4.6	3.9	1.6	-0.9	-3.1
6.5	4.2	3.1	2.6	0.4	-2.3	-4.3
5.2	2.7	1.8	1.7	-0.4	-3.3	-5.2
4.3	1.6	0.9	1.1	-0.9	-4.0	-5.7
3.7	0.9	0.2	0.7	-1.3	-4.5	-6.1
3.2	0.5	-0.2	0.4	-1.5	-4.7	-6.3
2.2	0.1	-1.1	0.0	-1.8	-5.2	-6.5

Figure 3.6. Values of late energy G_{late} for the seven conditions, relative to the 500 Hz octave of condition “natural acoustics”. Colors indicate the overall judgment in the listening test (green: “good, appropriate”).

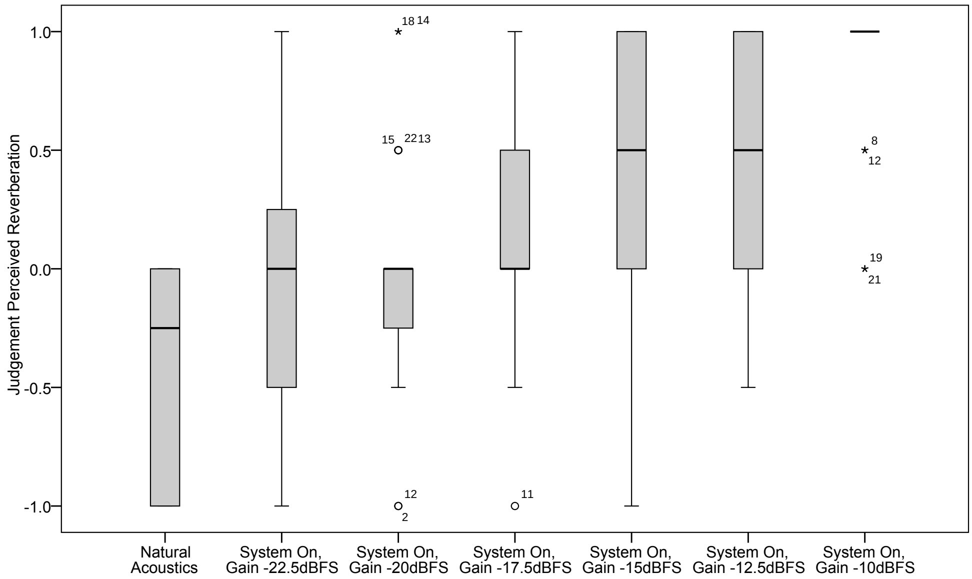


Figure 3.7. Boxplots for all participants (n=24) rating the presented reverberation. “-1” corresponds to “too little reverberation”, “0” – good/ appropriate, “+1” – “too much reverberation”.

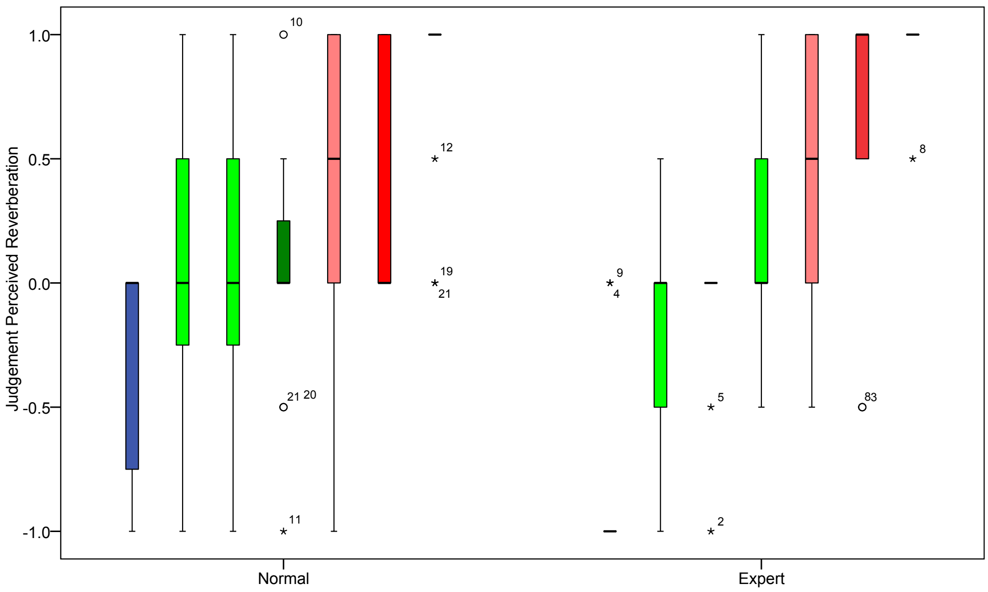


Figure 3.8. Boxplots for normal (n=15) and expert (n=9) listeners. “-1” corresponds to “too little reverberation”, “0” – good/ appropriate, “+1” – “too much reverberation”.

conducted without and with Bonferroni correction applied, resulting in a (fairly low) significance level set at $p<0.0024$. The results are shown as boxplots in Fig. 3.7 and can be interpreted as follows: Listeners preferred the stimulus "Gain -20dBFS", i.e. some added artificial reverberation (closer to 0) as more appropriate (Z=-3.265, p=0.001), as the real situation leans towards having too little reverberation. When separating the groups (Fig. 3.8) it can be seen that the responses of the normal listeners are more spread out but there is still a significant difference in the judgment (normal listeners: $\chi^2(6) = 39.688$, $p<0.000$). The acoustic conditions seem to be mostly accepted by the normal listeners, only one stimulus is judged as too extreme by most ("Gain -10dBFS"). With Bonferroni applied, only the fifth stimulus ("Gain -15dBFS") would be sig. different from the natural situation (Z=-3.066, p=0.002), without correction already the stimulus "Gain -20dBFS" (Z=-2.251, p=0.024), as above. For experts, the results are more accurate ($\chi^2(6) = 38.044$, $p < 0.000$). The third setting, judged to be the most appropriate, is the stimulus at Gain -20 dBFS. Interestingly the margin seems very narrow. The natural reverberation is clearly judged as acoustically too "dry". It can be concluded that a certain amount of artificial reverberation is appreciated.

The influence of the additional energy can be seen in the room acoustic measures above (Figs. 3.5 and 3.6). T_{30} increases from 1.6 to 1.9 s between the real and the first artificial situation and subsequently by approx. 0.15 s. Assuming a just noticeable difference (JND) of 5 % this corresponds to slightly bigger steps than the JND of around 0.1 s. The stimuli cover a range of roughly one second (1.6 to 2.65 s) with the most preferred setting at 2.1 s. Analyzing EDT reveals that the early decay time is somewhat shorter in this hall with 1.4 s at mid-frequencies. EDT is raised to 1.5-1.6 s at the most preferred setting with an increase at low frequencies. The reason for differences between T_{30} and EDT is discussed in Section 3.6. Late Strength G_{late} shows that the changes in gain translate to differences of 0.5-1 dB between stimuli. The most preferred setting was an increase in late energy of around 0.5-1 dB. For G, a just noticeable difference (JND) of 1 dB is reported[4]. Clarity C_{80} (not shown here) was lowered from +2.5 dB to around +1.5 dB for the most preferred case. Optimal values for C_{80} range from -2 to +2 for chamber music [19, p. 536]. To summarize, T_{30} was increased by 0.5 s, EDT by 0.2 s and the late energy by 0.5-1 dB.

3.2.3 Discussion

Even in a venue that is already optimized for chamber music, some artificial reverberation is appreciated. However, the amount of additional reverberation seems critical which agrees with a recent study investigating optimal reverberation levels in

[4] As Strength G is not calibrated according to standard, no absolute comparison with values from literature for late energy are possible.

audio mixing [87]. There, it was also found that too much reverberation is more easily disliked than too little reverberation - an important factor to consider for enhancement systems as well.

At first, the results are only valid for the presented signal and in fact, the signal characteristics are important when adding reverberation both in audio mixing and room enhancement. However, the stimulus in use can be characterized as typical for the venue and classical chamber music, solo violin music being a common genre with Bach partitas among the most played by violinists. From further tests by the author in the same room a slight increase in late energy was preferred also for solo voice, cello, organ and a medium-sized orchestra. None of the expert listeners mentioned a lack of realism or artificiality. Given that they were aware that the sound source was artificial and the acoustics were changing, this is notable.

Small deviations in T_{30}, EDT and G were well audible although close to the just-noticeable-differences (JND). A combined change both in reverberation time and strength appeared. The resulting values of EDT and C_{80} are closer to what would be considered optimal for chamber music whereas T_{30} was set as long as in a big concert hall. Other music styles might give different results.

3.2.4 Conclusion

Reverberation enhancement was tested in a chamber hall in a listening experiment. A slight increase through artificial reverberation was found desirable. This is a formal proof that electronic reverberation can enhance the listening experience. It became also clear that enhancement systems are sensitive to set properly which is in agreement with subjective experiences through practice; an optimal setting is important. This optimal setting was preferred both by experts and normal listeners with the experts being more accurate and critical.

3.3 Influence of reverberation level on reverberance

In the previous experiment, both reverberation level and reverberation time changed as the gain of the enhancement was increased. In order to investigate the importance of reverberation level, reverberance is judged for stimuli with varying decay times and level or strength. Two experiments are conducted, the first one with a set of stimuli covering a wide range, the second one with smaller differences among stimuli.

3.3.1 Variation of reverberation time and level: experiment 1

3.3.1.1 Setup

The experiment was conducted in the lecture hall at Müller-BBM in Munich, described in Section 2.2.1.1. A group of 23 listeners took part in the experiments, mainly students with some theoretical background in acoustics but no listening training. The students were visiting on a tour, participation in the test was voluntary. The average age was 23.9 years (median: 23 y) with 5 female and 18 male participants, no hearing loss was reported by the participants. All participants were tested at the same time, as a whole group. They were seated in a listening area of 4.0 m width by 2.4 m length (blue rectangle), see Fig. 3.9 [5]. The sound source was a Genelec 8030A loudspeaker at 9.6 m distance in 1.4 m height from the first row of listeners. The sound sample was a 12 second long anechoic guitar audio file with a listening level of ca. L_{Aeq} = 60 dB.

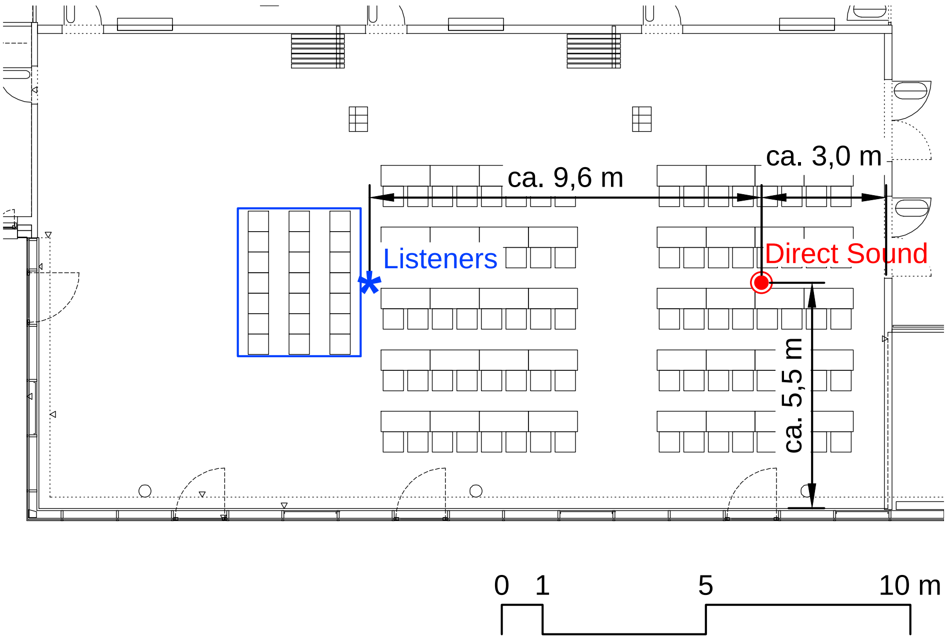


Figure 3.9. Ground view of the lecture hall with the test arrangement. Listeners were seated in the area marked in blue.

Different acoustic conditions were set with the enhancement system covering a range of reverberation times. The real acoustics as a reference ("Real room"), an artificially enhanced medium reverberation ("LowRT") and a long reverberation ("HighRT"). Included within each was a low and a high reverberation gain setting ("LowGain" and "HighGain"). The impulse response set was from a Vivace preset. While keeping the same reflection pattern the IR envelope was changed to exponentially fade out after 1.7 s or 4.0 s, thus reducing reverberation. Secondly, the overall digital gain was changed. The intention was to include two pairs with each similar reverberation times but different reverberation levels. The stimuli were not too difficult to distinguish

[5] Additional tables were placed behind the listeners which are not shown here.

from each other. It can be seen in Fig. 3.10 that the two pairs of stimuli (LowRT and HighRT) have similar measured reverberation times T_{30}. The stimuli could have been further matched by changing the filter of the IRs. Initially we decided against this in order not introduce another variable. The resulting deviations are discussed below. Fig. 3.11 shows the corresponding strength values, revealing a difference of around 1-1.5 dB in reverberation level between stimuli, except for the conditions “High RT, Low Gain” and “Low RT, High Gain” which have almost the same energy.

A paired comparison test design was chosen for presentation (for method, see Section 2.4). Participants were asked to choose the more reverberant stimulus (“Is stimulus A or B more or equally reverberant”), further defined as “having more reverberation, room sound”. Each participant filled in their answers on a paper form, which were then coded into a choice matrix. After calculating the choice frequencies, each stimulus was attributed with a certain *reverberance*. The test duration was approximately 11 min with 5 min of introductions and two pairs presented for training purposes. The complete comparison matrix was tested, but each pair was tested only once in the same order which could introduce some sequence effects. As all participants were tested at the same time, this does result in a different listening situation for each listener. Thus, impulse responses were measured at all listener positions (without the listeners present). The difference in T_{30} between individual listeners was always below the JND of 5% at 500 and 1000 Hz for all stimuli but could be greater than the JND for other octave bands. The directivity of the loudspeaker was homogeneous for the angle in which the group was seated. Overall it can be assumed that the listening condition for judging the reverberation was similar enough among listeners.

3.3.1.2 Results

Calculation of circular triads yields a consistency value >0.8 or 80% which according to current literature can be considered a fairly consistent performance of judgment. Hence all participants can be kept in the analysis.

Results for the first experiment are shown in Fig. 3.12. The overall order for the reverberance judgment is as expected, “natural acoustics/real room” as the least reverberant and the condition with high reverberation time and level as the most reverberant. The perceptual difference between stimuli, the “step” in reverberance, is equally spaced. This effect seems to be not as well reflected in the measured reverberation time T_{30} (Fig. 3.10, top). For instance, the two stimuli with long reverberation times (“HighRT,LowGain” and “HighRT,HighGain”) are noticeably different in reverberance even though the reverberation times T_{30} are fairly similar (both 4.3 seconds, except at high frequencies).

A correlation analysis of the relationship between listening test data and ISO 3382

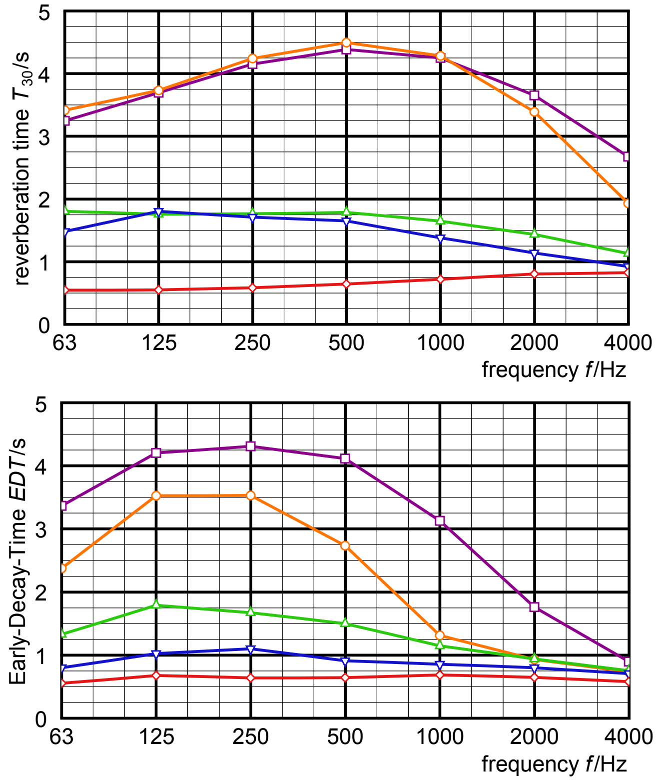
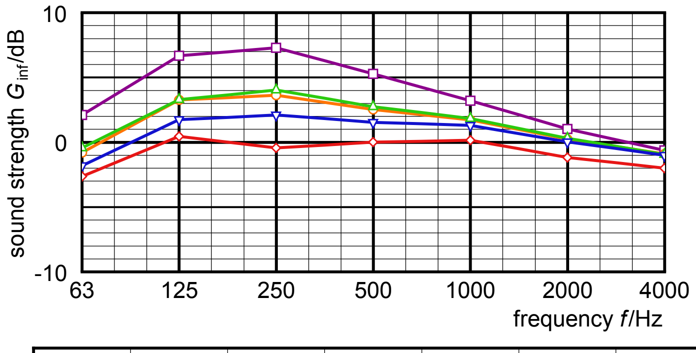


Figure 3.10. Reverberation time T_{30} (above) and early decay time EDT (below) for the five conditions (Experiment 1).



□—□	2.1	6.7	7.3	5.3	3.2	1.0	-0.6
○—○	-0.8	3.3	3.6	2.5	1.7	0.2	-0.9
△—△	-0.4	3.3	4.0	2.7	1.8	0.3	-0.9
▽—▽	-1.8	1.7	2.1	1.5	1.3	0.0	-1.0
◇—◇	-2.6	0.4	-0.4	0.0	0.1	-1.2	-2.0

Figure 3.11. Strength values G in dB for the five acoustic conditions (Experiment 1), normalized to the 500 Hz octave of condition “natural acoustics”.

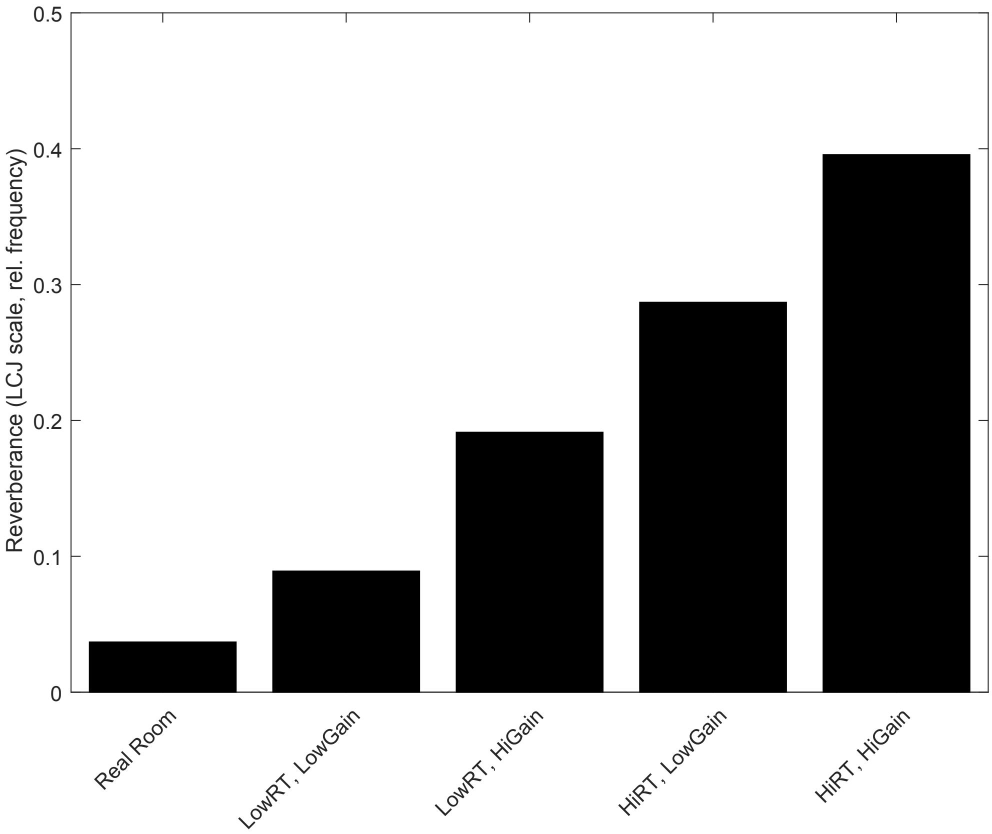


Figure 3.12. Reverberance judgments for the five acoustic conditions (Experiment 1).

parameters and extended parameters was conducted. Pearson correlation is applicable as the paired comparison data has an interval level. Table 3.1 contains values for the present experiment. It can be seen that decay times T_{30} and EDT, the standardized predictor for reveberance, perform similarly as the energy parameters that are including late energy. Overall the parameters fit well, with the exception of ratio EDT/T_{30} and early energy G80.

Table 3.1. Correlation coefficients between mean reverberance estimates in Experiment 1 and room acoustic parameters for 500 and 1000 Hz octave bands as suggested in ISO 3382. Strong correlations (|r|>0.9) are highlighted in bold.

f [Hz]	T_{30} [s]	EDT [s]	EDT/T_{30}	C_{80} [dB]	C_5 [dB]	G_{inf} [dB]	$G_{5-\mathrm{inf}}$ [dB]	G_{late} [dB]	G_{80} [dB]
500	**0.93**	**0.98**	0.04	**-0.96**	**-0.94**	**0.94**	**0.94**	**0.95**	0.41
1000	**0.94**	0.89	-0.44	**-0.94**	**-0.92**	**0.93**	**0.93**	**0.93**	0.44

3.3.2 Variation of reverberation time and level: experiment 2

Since the range of stimuli was rather large in the previous experiment with reverberation times ranging from approximately 0.7 to over 4 seconds and strength differing by 5 dB, a second experiment was conducted with smaller differences between stimuli.

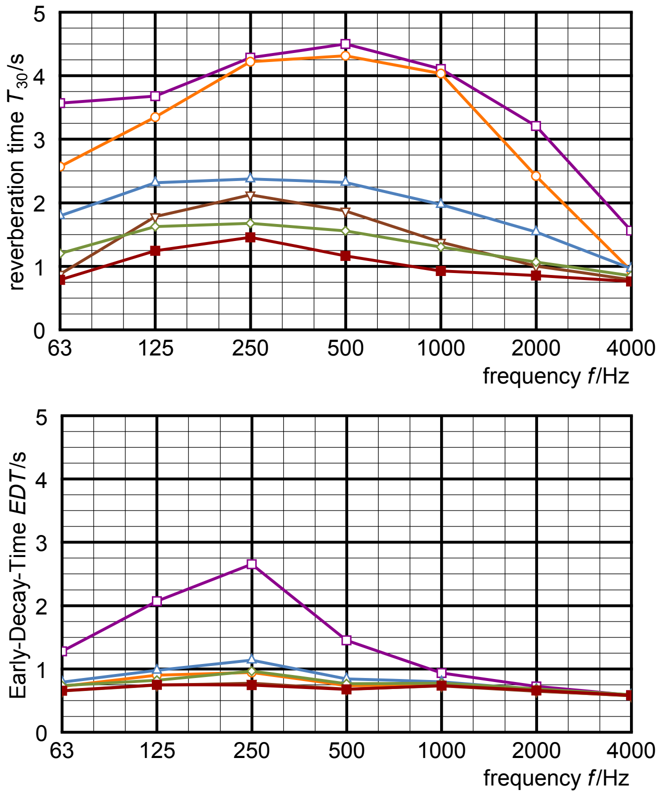
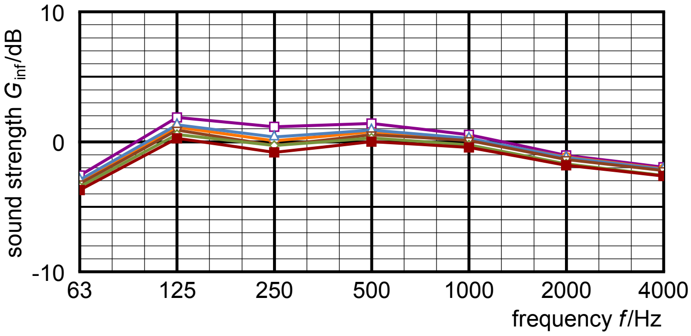


Figure 3.13. Reverberation time T_{30} (top) and early decay time EDT for the six conditions (Experiment 2).

□ High RT, High Gain
○ High RT, Low Gain
△ Medium RT, High Gain
▽ Medium RT, Low Gain
◇ Low RT, High Gain
■ Low RT, Low Gain



□	-2.6	1.9	1.2	1.4	0.5	-1.1	-2.0
○	-3.1	1.1	0.1	0.8	0.2	-1.2	-2.1
△	-2.9	1.3	0.4	0.9	0.2	-1.2	-2.1
▽	-3.2	0.9	-0.3	0.5	0.1	-1.4	-2.2
◇	-3.5	0.5	-0.3	0.3	-0.3	-1.7	-2.6
■	-3.7	0.2	-0.8	0.0	-0.4	-1.8	-2.7

Figure 3.14. Strength values G in dB for the six acoustic conditions (Experiment 2), normalized to the 500 Hz octave of condition "LowRT,LowGain".

3.3.2.1 Setup

The same arrangement and test design was used with a different set of stimuli and different participants. Again, a student group participated voluntarily with 23 persons and an average age of 26 years (Median 23 y, 10 female and 13 male). The subjects could be characterized as normal, non-trained listeners with some theoretical background in acoustics. No hearing problems were reported. Six stimuli were chosen this time, more closely related and therefore intended to increase the difficulty. Three pairs of impulse responses with the same digital impulse response length were set to two different reverberation gains. The resulting stimuli differ from the previous test. Only the two “HighRT” conditions are similar in terms of reverberation time T_{30} compared to the previous experiment (Fig. 3.13, top). The lowest T_{30} is around one second. More importantly, in the previous experiment, strength values for the acoustic conditions were distributed over a range of approximately 5 dB. Now, the difference between least and most reverberant condition is only 1.5 dB approx. at mid-frequencies, a much smaller change in reverberation level (Fig. 3.14). As the additional reverberation is quieter and does not alter the evaluation range noticeably between 0 to -10 dB, EDT does not differ much (Fig. 3.13, bottom).

3.3.2.2 Results

For all 13 subjects the consistency measure is only >0.4, likely due to a more difficult discrimination overall and less participants compared to experiment 1. To reach a level of consistency >0.6, four out of 13 subjects can be excluded from the analysis to form a consistent listeners group (n=9). The results for both are shown in Fig. 3.15. As in the previous experiment, the overall trend for the reverberance estimates are according to expectations. However, the estimates are not separated as clearly. The judgment between all listeners and only consistent listeners is fairly similar except for the two leftmost, least reverberant stimuli.

As in Experiment 1, there is at least one situation where reverberation time T_{30} fails to represent the perceptual measure. The amount of reverberance is judged to be almost equal between the fourth stimulus “MidRT,HighGain” and the following stimulus “HighRT,LowGain” while differing by almost T_{30} = 2 s. Early decay time EDT, the standardized predictor for reverberance, does not work well for this set of stimuli. This issue becomes apparent also from the correlation table 3.2. Only the energy parameters including late energy maintain very good correlation values. Whereas, decay time parameters perform worse.

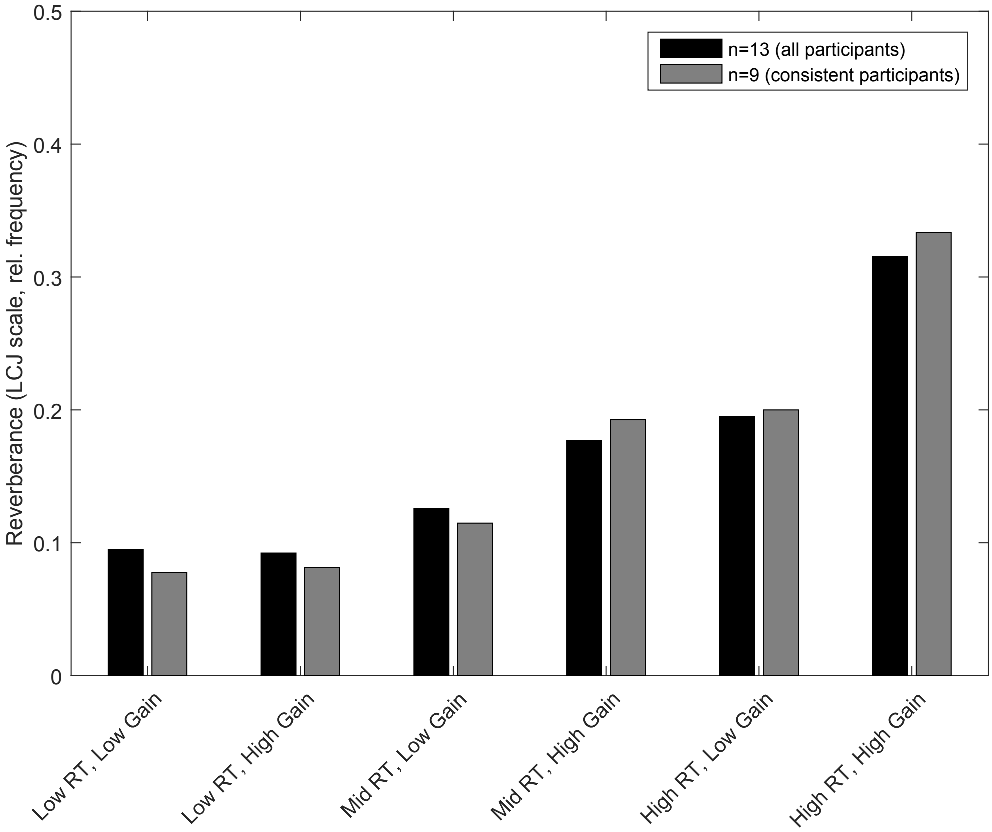


Figure 3.15. Perceived amount of reverberation for experiment 2 with all 13 listeners (consistency T>0.4) in black and more consistent 9 listeners (T>0.6) in grey.

Table 3.2. Correlation coefficients between reverberance estimates in Experiment 2 and room acoustic parameters for 500 and 1000 Hz octave bands as suggested in ISO 3382, consistent listeners (n=9). Strong correlations (|r|>0.9) are highlighted in bold.

f [Hz]	T_{30} [s]	EDT [s]	EDT/T_{30}	C_{80} [dB]	C_5 [dB]	G_{inf} [dB]	$G_{5-\mathrm{inf}}$ [dB]	G_{late} [dB]	G_{80} [dB]
500	0.89	0.88	-0.63	-0.84	**-0.97**	**0.97**	**0.97**	**0.93**	0.85
1000	0.87	0.86	-0.84	-0.86	**-0.94**	**0.92**	**0.92**	**0.95**	0.78

3.3.3 Discussion

When comparing the outcome of the two experiments, it is observed that only the energy parameters including the late energy performed well in both cases. Whereas, the currently standardized parameter for reverberance, EDT was not as consistent. This importance of the reverberation level is in line with findings discussed in the introductory section (e. g. [64], [65], [11] and [66]). At the same time, the outcome of a correlation analysis always depends on the individual stimulus set of a study, and should thus not be generalized.

It was observed in both experiments that EDT and T_{30} became different from each other. Namely, the early level decay is different from the later level decay. In other words, deviations from the linear level decay or *multi-slope decays* appeared. This finding will be discussed more in Section 3.6 of the chapter. The occurrence somewhat hinders independent manipulation of reverberation level and reverberation time. Although the listening situation, or sound field, might be artificially created, the sensation rated by the participants is real and the judgments valid for this situation.

3.3.4 Conclusion

Enhanced acoustic conditions with varying reverberation levels and reverberation times were rated in listening tests regarding reverberance estimates. Reverberant conditions with a T_{30} difference of up to 2.5 sec were perceived as similar and others with the same T_{30} were judged differently. Only the energy parameters including late energy performed well consistently. This observation suggests that influence of reverberation levels is as important as changes in reverberation time and should be analyzed in room acoustics.

3.4 Equal reverberance matching

In the previous two sections, reverberation was compared and rated by means of magnitude estimation, in terms of quality (Section 3.2) and quantity (Section 3.3). In this section, a magnitude adjustment test is performed where different reverberation is matched to evoke *equal reverberance* in order to investigate further quantitative aspects (see also [88]).

3.4.1 Setup

The experiment was conducted in the lecture hall equipped with an enhancement system described in Section 2.2.1.1. The direct sound was played back by a single speaker (Genelec 8030A at height 1.4 m). The listener was seated at a distance of ca. 9.6 m from the direct sound loudspeaker (see Fig. 3.16). The speaker was set off the middle axis 0.5 m to the right as it was found to better integrate in the enhancement sound field. The direct sound stimulus in this experiment was a solo saxophone audio excerpt, a melodic line with a duration of ca. 5 s anechoic/“dry” signal and 2.5 s for decay and silence, no fade-out, playing on an endless loop that would remain audible until the Vivace Stop-Button on the laptop in front of the listener was pressed. In the excerpt breaks, the reverberation tail was audible to fade into silence. The lowest and the highest played note of the saxophone example were at ca. 190 Hz and 330 Hz respectively.

The listening level was set to be an average L_{Aeq} = 64 dB at the listeners position as measured for a loop of 30 sec with artificial acoustics at an expected target setting for the experiment (B&K 2250 SPL-meter). If the electronic reverberation was to be turned up all the way, a L_{Aeq} = 68 dB would have been reached (L_{AFmax} = 73 dB). The background noise was measured at L_{Aeq} = 24 dB.

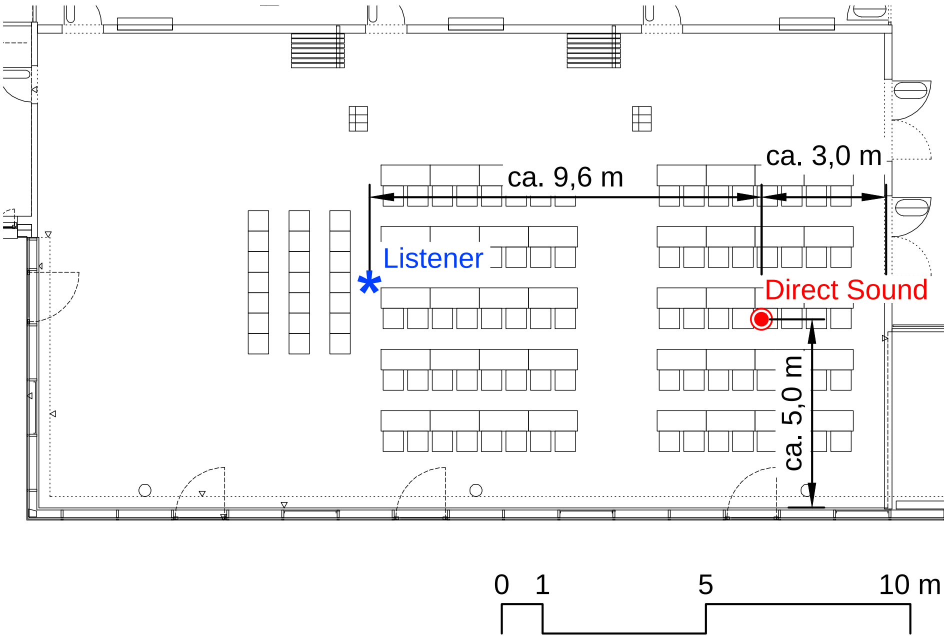


Figure 3.16. Ground plan of the lecture hall with the test setup.

The room acoustics of the real room served as a zero condition from which the participants started (T_{30} = 0.7 s). Electronically, four acoustic engines were available on the main screen in the Vivace system. Four different reverberation settings (“reverbs”) could be accessed in parallel. The four reverbs were derived from a fixed Vivace preset by manipulating the impulse response. An exponential fade-out was applied without otherwise altering the reflection pattern. “Reverb1” was set with a digital reverberation time of 1.6 seconds, the “Reference Reverb” to 1.3 seconds, “Reverb3” to 1.8 s and “Reverb4” 1.3 seconds with modified frequencies. The resulting measured reverberation times after matching will be presented in the *Results*-section (Fig. 3.18).

The levels of three reverbs had to be adjusted successively to a reference reverb by moving the faders. It was advised to listen for the decay and reverberance. The smallest possible fader step in the software was 0.4 dBFS (equivalent to a change of ca. 0.15 dB in Strength *G* in the target area, see also Fig. 3.17). When no *Solo* button was pressed, the reference was active and could be switched to another reverb by pressing "S". The task was repeated twice by the same participant. Since there was no training apart from the introduction to the user interface and method, the first try was only kept in the analysis after asking the listener if he or she felt comfortable with it as a valid try and comparing the performance later on. The matching task presented in this section was one of three tasks within the same test run. The average length for the matching was 9 min without introduction.

The level data from every fader position on the Vivace surface for each participant was put into SPSS. Boxplots were generated and the median calculated for 1st and 2nd tries separately and both tries together. Both tries were then combined. For the measurements, the faders were set to the position of the median values as well as minimum and maximum responses.

The participants were 14 colleagues from Müller-BBM of which at least half could be considered skilled listeners with on average >10 years of experience in room acoustics, listening tests and/or formal training in sound recording. The average age was 46 years with 12 male and 2 female participants. They were informed orally. Pre-written instructions about the structure and task of the study, expected test duration and sound pressure levels occurring were read aloud. It was noted that stopping or pausing was possible at any time.

Impulse response measurements were taken at the listening position with a monaural measurement microphone (Microtech-Gefell) and a commercial Cortex dummy head. The sweep signal was played from the measurement software mReverb via a Fireface UC through the Vivace System and Nexus converters to the direct sound speaker and reverberation speakers. The measurements were conducted at the median level values found in the experiment as well as minimum and maximum levels for each trial and fader as set by the participants. Parameters were measured in octave bands and later on, are given as an average of the frequencies 500/100 Hz according to standard. Additionally, the frequencies 125-4000 Hz were analyzed, as this frequency range offers the best correlation for reverberance according to [63]. Note that strength *G* values are referenced to the condition "real room/ natural acoustics".

3.4.2 Results

Figure 3.17 shows the fader levels required to perceive the same reverberation as set by the participants. The deviation within a "reverb" might seem large at first, yet this is an

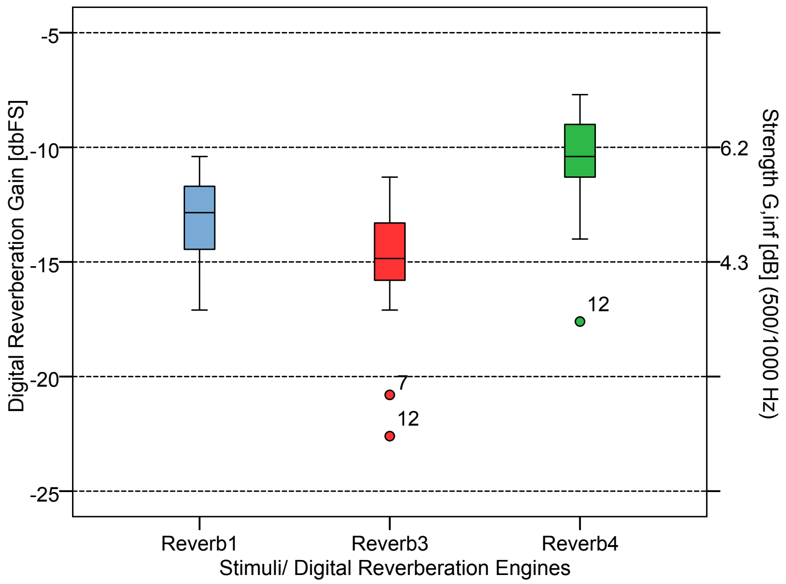


Figure 3.17. Boxplot with reverberation gains as set by 14 participants for *equal reverberance*. The median values were used for the measurements presented in table 3.3 and following.

arbitrary digital dBFS level. Strength values have been given for comparison, from which it becomes clear that all the answers lie within less than ca. 2 dB. Taking into account a just noticeable difference (JND) of 1.0 dB for G this is an additional deviation of ± 0.5 dB, which appears to be an overall good agreement between participants.

As each participant repeated the task it is interesting to see if the values set by individuals are stable, in other words if the fader position was set consistently by the individual. This factor is analyzed by calculating the individual standard deviations from the first to the second try. The mean individual deviations are only 0.8 dBFS. Participants were consistent overall with their fader setting, especially since the accuracy or smallest step was only 0.4 dBFS when moving the fader up and down.

Frequency dependent reverberation times T_{30} are shown in Fig. 3.18. There are clear differences of around 0.2 seconds between Reverb 1, 3 and 4.

As the task was to adjust the reverberation to have the same reverberance as the reference, it was expected to find a measurement parameter where the values for reference, reverb 1, 3 and 4 (grey area in the following table) are the same. Analyzing the conventional and extended IR measures in Table 3.3 it can be seen that T_{30} and EDT are quite different between stimuli in the grey shaded conditions. Decay time does not act as a good predictor here as both T_{30} and EDT differ between the four conditions by several JNDs. From the energy-parameters, C_{80} provides the closest matches with a standard deviation of 0.5 dB. This finding is interesting as it is not used to describe reverberance but perceived *clarity* or *definition* (though there is usually strong inter-correlation). C_{80} as a measure of the balance between early and late energy in general might serve this purpose here, or, participants listened for clarity as an

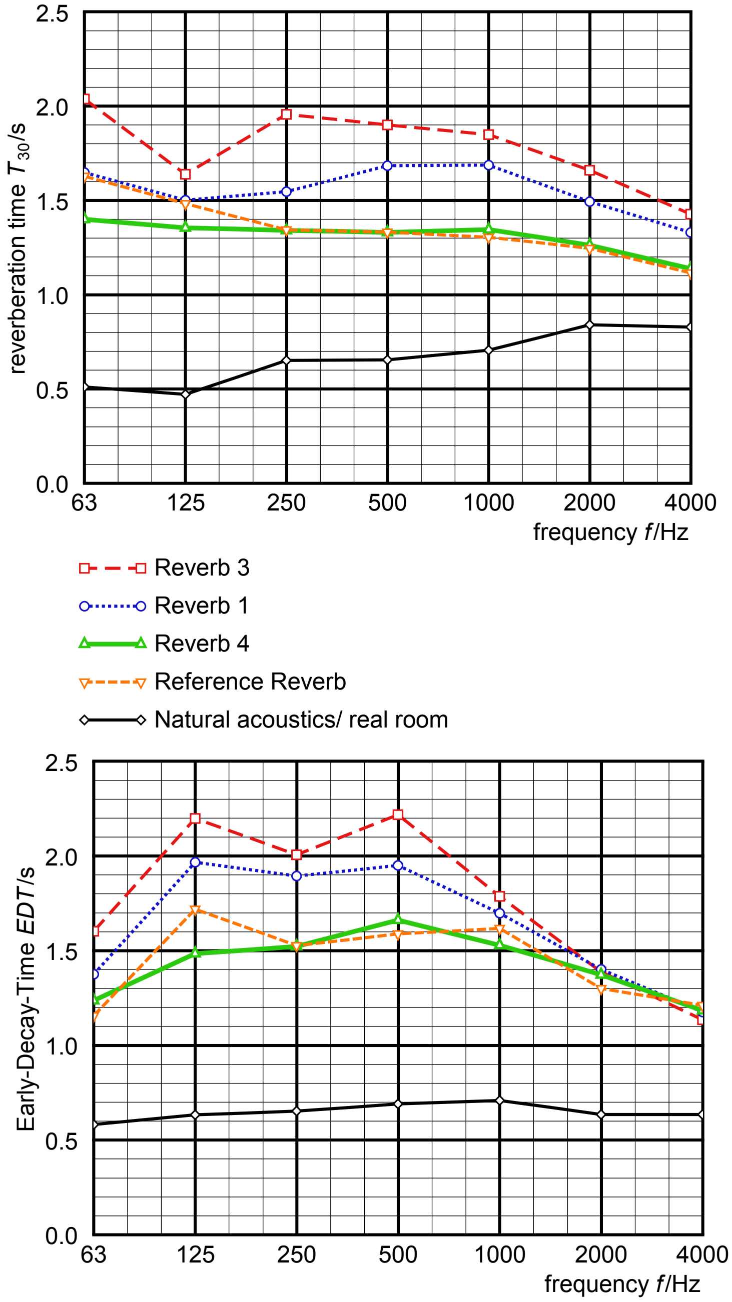


Figure 3.18. Reverberation time T_{30} (top) and early decay time EDT for the different acoustic situations as set by the participants for equal reverberance.

Table 3.3. Conventional and extended ISO-Impulse Response Parameters (Averaged 500/1000 Hz , rounded to 0.05). G values are normalized to the condition of the real room/ natural acoustics. The grey area shows the stimuli values after being matched to equal reverberance.

	T_{30} [s]	EDT [s]	G_{inf} [dB]	G_{80} [dB]	G_{late} [dB]	C_{80} [dB]	C_{50} [dB]	C_5 [dB]
Real room	0.70	0.70	0.0	0.5	-0.6	7.8	3.8	-2.8
Reference	1.30	1.60	6.9	4.5	2.0	-0.9	-4.6	-12.5
Reverb 1	1.70	1.80	5.2	3.6	1.6	-1.0	-4.1	-9.9
Reverb 3	1.85	2.00	4.7	3.4	1.4	-1.0	-3.4	-9.3
Reverb 4	1.35	1.60	6.8	4.7	2.4	-2.1	-5.0	-11.5
Std. Dev.	0.24	0.17	1.0	0.9	1.1	0.5	0.6	1.3

Table 3.4. Conventional ISO- and combined Impulse Response Parameters (Averaged 500/1000 Hz, rounded to 0.05 for decay times).

	T_{30} [s]	EDT [s]	T_L, [76] [s]	EDT_L, [76] [s]	L_{Aeq} [dB]
Real room	0.70	0.70	0.75	0.75	51.9
Reference	1.30	1.60	1.25	1.45	63.9
Reverb 1	1.70	1.80	1.45	1.60	62.9
Reverb 3	1.85	2.00	1.65	1.70	63.1
Reverb 4	1.35	1.60	1.25	1.40	64.8
Std. Dev.	0.24	0.17	0.16	0.11	0.75

attribute. All other energy parameters have deviations from the mean of around 1 dB, so in the order of one JND for level parameters. Early Strength G_{80} performed well, which is a little surprising, as it is not usually thought of for reverberance. This fact might be due to a level matching between stimuli which depends on the earlier, stronger part of the IR. G_{late} was expected to perform better as it measures the energy of the late part of the IR. Note that these are octave band average 500/1000 Hz as reported according to standard in current practice. Broadband averages (125-4000 Hz) performed slightly better for all parameters, shown in the Appendix, p. 147, Tab. 6.1.

For calculating absolute values Lee et al. [76] suggested a listening level correction for a decay time as shown in Eq. 3.1 based on experimental data where L is the listening level (not the reverberation level), measured as L_{Aeq}:

$$EDT_L = EDT^{\frac{L}{80}} \tag{3.1}$$

The exponent compresses or expands the original reverberation time value. For the present experiment, this is calculated in Table 3.4. It can be seen that the variation does decrease compared to the ISO parameters but is still larger than the just noticeable difference.

As previous results suggest that reverberation time can be balanced with reverberation level it was attempted to combine a decay parameter with an energy parameter, e.g. T_{30} or EDT together with strength G, which is also more general than listening level.

Cremer and Müller suggested a relationship between changes in reverberation time and room volume as presented in Eq. 3.2 [13, p. 493]. Combining this suggestion with the relationship between sound strength G and room volume/reverberation time (Eq. 3.3), a change of reverberation time T from a given T_0 is assumed to be equally perceived as a change in Strength G (Eq. 3.4).

$$lg(\frac{T}{T_0}) = \frac{1}{1+\theta} lg \frac{V}{V_0} \tag{3.2}$$

$$G \sim -10 lg \frac{V}{T} \tag{3.3}$$

$$\frac{T}{T_0} = 10^{\frac{-\Delta G}{10\theta}} \tag{3.4}$$

θ is the quotient of the just noticeable differences (JNDs) for level parameters, divided by the JND for reverberation time. When applying the data from the present experiment the quotient is θ=1.45, calculated from the strength and decay time differences of equally perceived stimuli (grey, Tab. 3.3). Applied to Eq. 3.4 it can be seen in Figure 3.19 that a reverberation level change of 1 dB is equivalent to a reverberation time change in T_{30} of 15%. Or, in other words, if reverberation is louder the decay can be shorter. For EDT, not shown, the change equals 10% for the present data (θ=2.4)

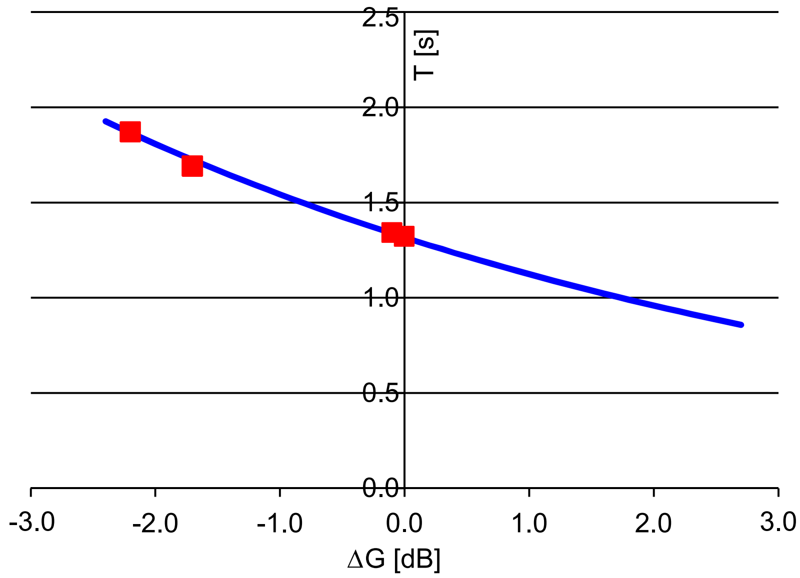


Figure 3.19. Relationship between change in strength and reverberation time for maintaining equal reverberance, calculated using Equation 3.4. Red points are data from Table 3.3.

3.4.3 Discussion

The curve in Fig. 3.19 describing the relationship between changes of reverberation level and reverberation time is steeper in this study than the relationship suggested by Cremer et al. [13]. There the values for the quotient were given as 1 dB and 4% respectively, based on experimental data from noise stimuli. JNDs are known to differ depending on the stimulus set. It can be concluded that values found here are only valid for the given set but the qualitative conclusion is the same.

Regarding the test methodology, participants were asked to do the second try in a different order but the reverberation engines stayed in the same spot and were not redistributed as should be done for a randomized presentation. The possibility of switching instantaneously seemed to be a very effective comparison method (single click on *solo* button). Even though there was no level reading, there is a chance that orientation for the repetition came from the fader position on its scale. Yet, the fader covers a range of -60 to 0 dBFS and participants would have had to memorize the position very exactly before it was reset (0.8 dBFS accuracy was found above). A lot of the participants did not seem to look at the screen much but “into the distance”. This could be avoided by providing endless faders but this was not easily possible to implement in the setup.

In general, the task was well understood and conducted. When asked for remarks, some of the participants gave some feedback afterwards, the rest felt no particular challenge or difficulties. Two participants said that the matching task was difficult. One of which remarked the duality of reverberation decay or strength and uncertainties regarding which to use as a matching criterion. As this difficulty was said to be part of the test, this listener was confirmed in his/her procedure and internal criterion during the preceding test. Another participant mentioned this duality as well and that the reverberation would change with the progress of the signal. One person complained about the length of the stimulus being unsuitable. In the beginning of the experiment, this participant listened mainly for the decay of the phrase. Since this would be audible only every 6-7 seconds, it appeared to be difficult.

3.4.4 Conclusions

A reverberance matching experiment was conducted in a room enhancement environment. Stimuli with up to 20% shorter reverberation times were rated equally reverberant when reverberation level increased. The study demonstrates clearly that perception of reverberation highly depends on the reverberation level. Only assessing decay time is insufficient and must be combined with an energy parameter such as strength G.

3.5 Loudness-based reverberation analysis

An alternative approach, implemented by above mentioned authors Lee et al. [75], is to input impulse responses into a loudness model and derive reverberance estimates afterwards. The approach would thus include the duality between reverberation level and reverberation time since level dependent properties are accounted for in the hearing model and also other possibly relevant factors such as masking are reproduced. The findings were evaluated for rather large parameter changes and have not been re-validated or tested in practice, which would include checking other models as a step towards a standardized procedure (see also [88]).

3.5.1 Setup

Loudness analysis was first conducted with a Loudness-Decay-Time model by Lee et al. using loudness models from Moore and Glasberg, both included in the Matlab toolbox Aarae [52]. The Matlab script outputs a reverberation time T_{N} and an early decay time $\mathrm{EDT_N}$. Based on the original ISO measures the evaluation ranges are suggested to be a halving of the loudness for EDT (10 dB) and appropriately for T_{30}. The choice of other ranges were reported to not have increased performance [75].

As expected, acoustic impulses played back by the speaker and recorded on-site, would not offer usable information as the level was not high enough to reach a reasonable signal-to-noise ratio. Thus, impulse responses were utilized. The maximum of the IR was set equal a level of L_{AFmax} = 79 dB. A recording of the acoustic situation was conducted at the listener position with an artificial reference head (Section 2.3.1.1), at the median gain that was set for each reverb. Sound pressure levels and loudness values for the audio recordings were derived using PsySound3.

3.5.2 Results

3.5.2.1 Loudness-based parameters

Table 3.5 compares ISO- and loudness parameters. Very good performance is offered by the loudness-based IR parameter substituting early decay time: $\mathrm{EDT_N}$ was matched closer than the JND of reverberation times (0.05 s). T_{N} performs, at most, slightly better than the conventional parameter. When analyzing stimulus sound pressure levels, it can be observed that the real acoustics, which served as a starting point for every reverb-adjustment, were 12 dB quieter than the reference condition. It is therefore quite possible that a level or loudness adjustment was done at first to a certain point. From this adjustment, other factors have continued as a decision criterion as the final sound pressure values differ by roughly 2 dB (similar for loudness).

Table 3.5. Conventional ISO parameters (averaged 500/1000 Hz), loudness-based impulse response measures (middle), rounded to 0.05. Stimulus recording parameters on the right.

	T_{30} [s]	EDT [s]	T_{N} [s]	$\mathrm{EDT_N}$ [s]	L_{Aeq} [dB]	N [sone]
Real room	0.70	0.70	0.90	1.15	51.9	7.5
Reference	1.30	1.60	1.40	1.85	63.9	13.8
Reverb 1	1.70	1.80	1.70	1.90	62.9	13.7
Reverb 3	1.85	2.00	1.90	1.90	63.1	12.1
Reverb 4	1.35	1.60	1.40	1.85	64.8	14.2
Std. Dev.	0.24	0.17	0.21	0.02	0.75	0.80

Fig. 3.20 shows the loudness processed impulse responses from the model output over time. It can be seen that the initial decay seems to have been matched (which is where $\mathrm{EDT_N}$ is measured) leaving different length of late decay.

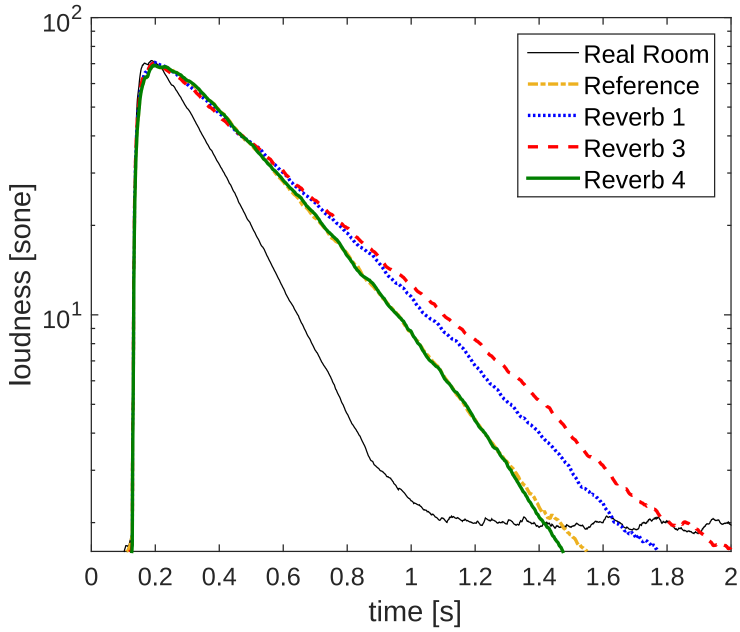


Figure 3.20. Loudness processed impulse responses of the different acoustic situations as set by the participants.

3.5.2.2 Level dependency

In Fig. 3.21 the model estimates for the five stimuli are given for five different levels. The level of the impulse response is varied. The maximum value (L_{AFmax}) is changed in steps of 10 dB between 60 and 100 dB. As expected, with increasing level, reverberation times are longer. The difference or spread between values depends on the stimulus (ca. Δ0.25 s for “Real room” to Δ0.7 s for “Reverb 3”.

3.5.2.3 Comparison between models TVL, DLM and ISO532-1:2016

Lee et al. mainly used the Time Varying Loudness Model (TVL) and argued that the Dynamic Loudness Model (DLM) and TVL yielded similar loudness decay functions with almost identical slopes [75]. However, they did not discuss the model output

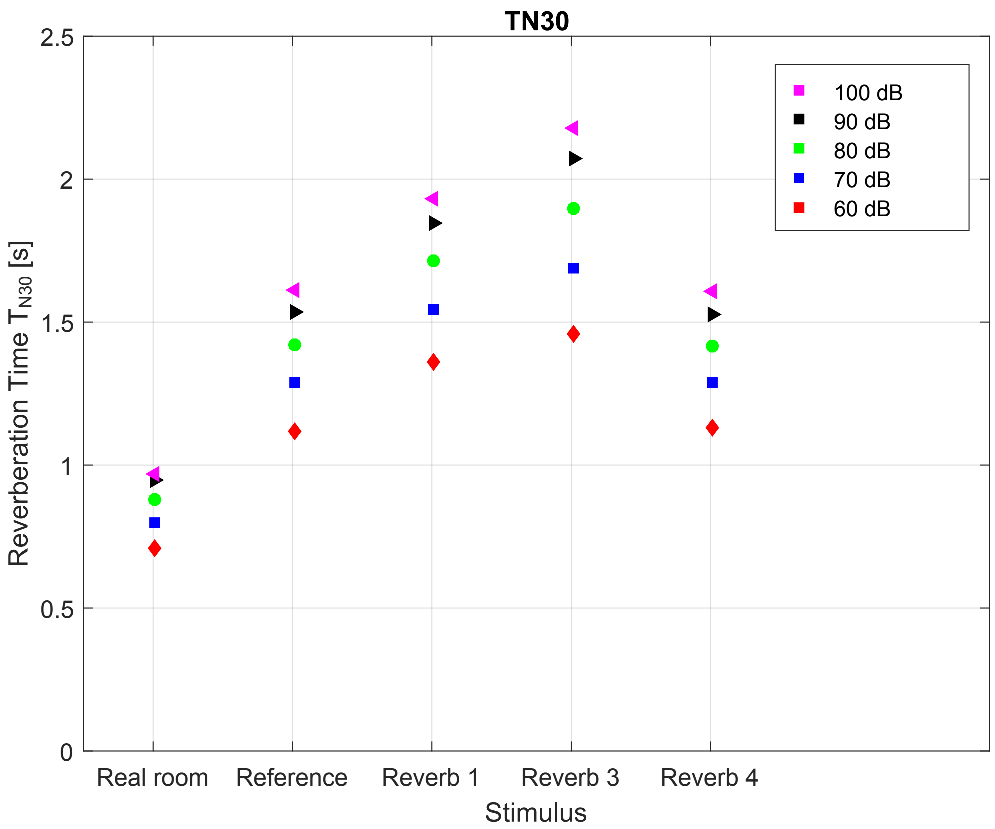


Figure 3.21. Loudness-based Reverberation Time (T_{N30}) for different input levels in the reverberance matching experiment.

estimates. If loudness decay analysis is to be used more widely it is important to know if and how the loudness model outputs differ as this could change the decay time values or exhibit other inconsistent behavior. Thus, a comparison between DLM, TVL and ISO532-1:2016 is done.

The implementation of DLM used is the version as included in the Matlab Toolbox PsySound 3 [89]. The calculation of the ISO532-1:2016 model estimates is done with a C# implementation based on and validated using the C-code available with the standard [90]. It was integrated into the loudness decay analysis as part of a master thesis supervised by the author [91]. The loudness decay analysis was tested to give identical values as the Matlab-implementation, so the models themselves are the only variable. The ISO model is based on a Zwicker loudness model and has been revised and standardized to include instationary loudness. The models have been calibrated to output 1 sone at 1 kHz for a sound pressure level of 40 dB. Subsequently, the loudness-based early decay time ($\mathrm{EDT_N}$) and loudness-based reverberation time (T_{N30}) are compared. As above, the level of the impulse response is varied. The maximum value (L_{AFmax}) is changed in steps of 10 dB from 60 to 100 dB.

T_{N30}, the loudness-based equivalent to the reverberation time T_{30}, is shown in Fig. 3.22 for the three models. The 80 dB values for TVL (solid green) were previously given in table 3.5. The change in decay time with different levels is apparent for all models. Overall, the two models DLM and TVL are more similar with the best accordance for levels around 70 dB (blue). For lower levels TVL estimates are shorter and for higher levels longer than DLM. This finding equates to a larger level-dependent spread. Here, the regular reverberation time T_{30} (marked with x) lies never above an estimate equivalent to 80 dB. ISO532-1 estimates are overall noticeably shorter than the other two models.

In the experiment, reverberance was equalized between stimuli “Reference” and “Reverb” 1/3/4. Thus a difference of 0 s between the estimates of those stimuli would be expected if equal reverberance was predicted properly. This is not the case, none of the estimates are similar among the four matched stimuli.

Results for the $\mathrm{EDT_N}$ are shown in Fig. 3.23. Again, it can be seen that, as expected, models DLM and TVL estimate longer decay times with an increasing level. The increase from level 60 dB to 90 dB leads to 0.1-0.3 s longer $\mathrm{EDT_N}$ values for the models DLM and TVL. Secondly, the DLM-model estimate (empty markers) is always somewhat shorter than TVL, on average by 0.15 s, depending on the stimulus. For stimulus “Reverb 3” the estimates are almost equal, this IR has a long decay (see Fig. 3.20. For the TVL, the highest input levels (100 dB, magenta) yields shorter decay times than with the 90 dB for all stimuli. This somewhat odd behavior likely has to do with the amplified noise floor and will be described below. ISO532-1 model exhibits rather inconsistent behavior. Estimates are sometimes longer (Stimuli 1, 3, 4) or shorter than the other two models (Stimuli 2, 5). Also, for stimulus “Reverb 1” and “Reverb 4” the level change has no effect.

The target of predicting the same reverberance is met best for early decay time $\mathrm{EDT_N}$ with TVL model at level 70 dB (filled blue).

3.5.3 Previous experiments

Magnitude estimation experiments from previous sections are reanalyzed with DLM and TVL:

Figure 3.24 shows loudness-based early decay time $\mathrm{EDT_N}$ for Experiment 1 from Section 3.3. As before, the decay times increase with level. The highest level condition in TVL (filled magenta, 100 dB) however is shortened. The TVL model always estimates longer $\mathrm{EDT_N}$ values than DLM, as previously, by around 0.15-0.2 s for lower levels to 0.25-0.3 s for the high levels. The offset between the models is thus not constant but also depends on the stimulus and the level.

Loudness-based Reverberation Time T_{N30} is given in Fig. 3.25. The values also show similar changes as before: T_{N30} estimates of the two models are overall relatively similar, TVL predicts shorter T_{N30} at 60 dB and longer T_{N30} at higher levels 80-100 dB. Again, at around 70 dB the models are the closest to each other. It is noteworthy

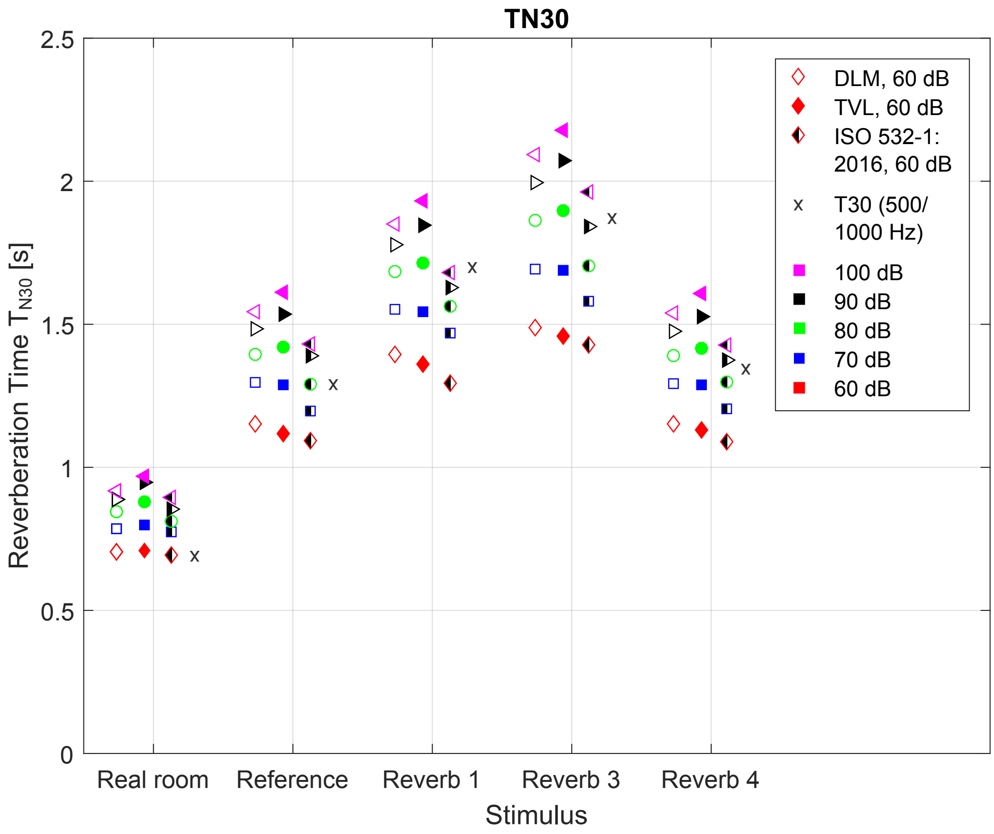


Figure 3.22. Loudness-based Reverberation Time (T_{N30}) for Loudness models DLM (empty markers), TVL (filled), ISO532-1 (semi-filled) and regular T_{30} (x) for different input levels.

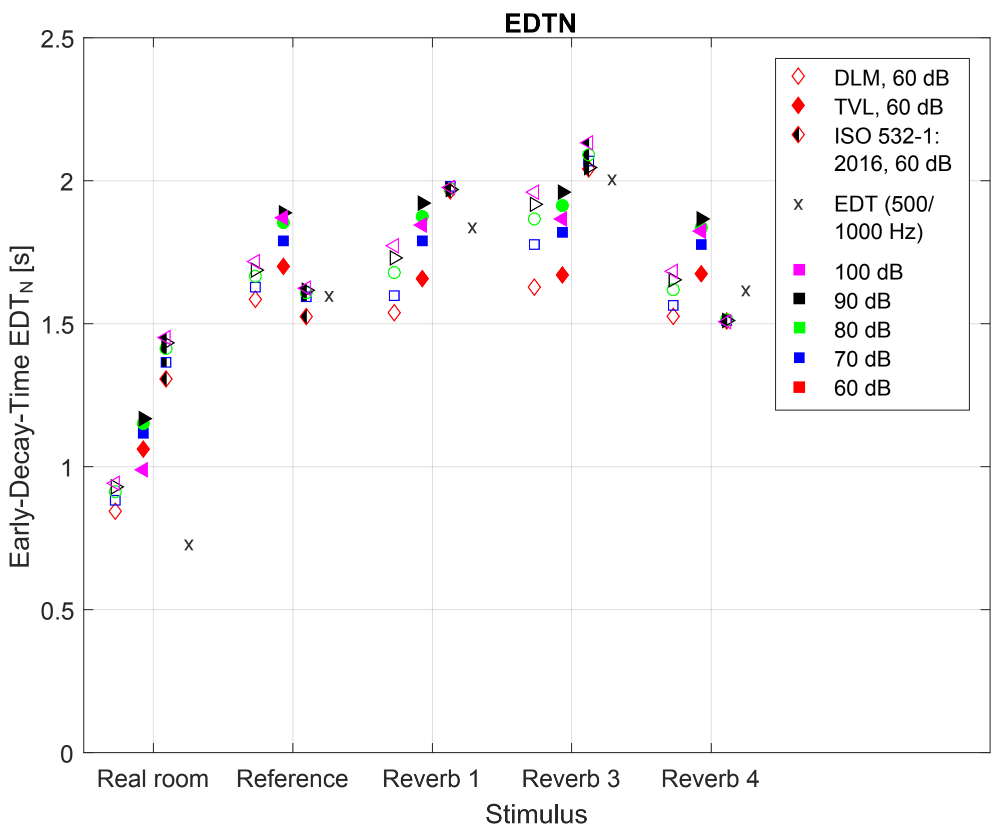


Figure 3.23. Loudness-based early decay time (EDT_N) for Loudness models DLM (empty markers), TVL (filled), ISO532-1 (semi-filled) and regular EDT (x) for different input levels.

that the 30 dB-level increase leads to an increase in reverberation time of up to Δ2.5 seconds with one acoustic condition (e.g. "HighRT,HiGain").

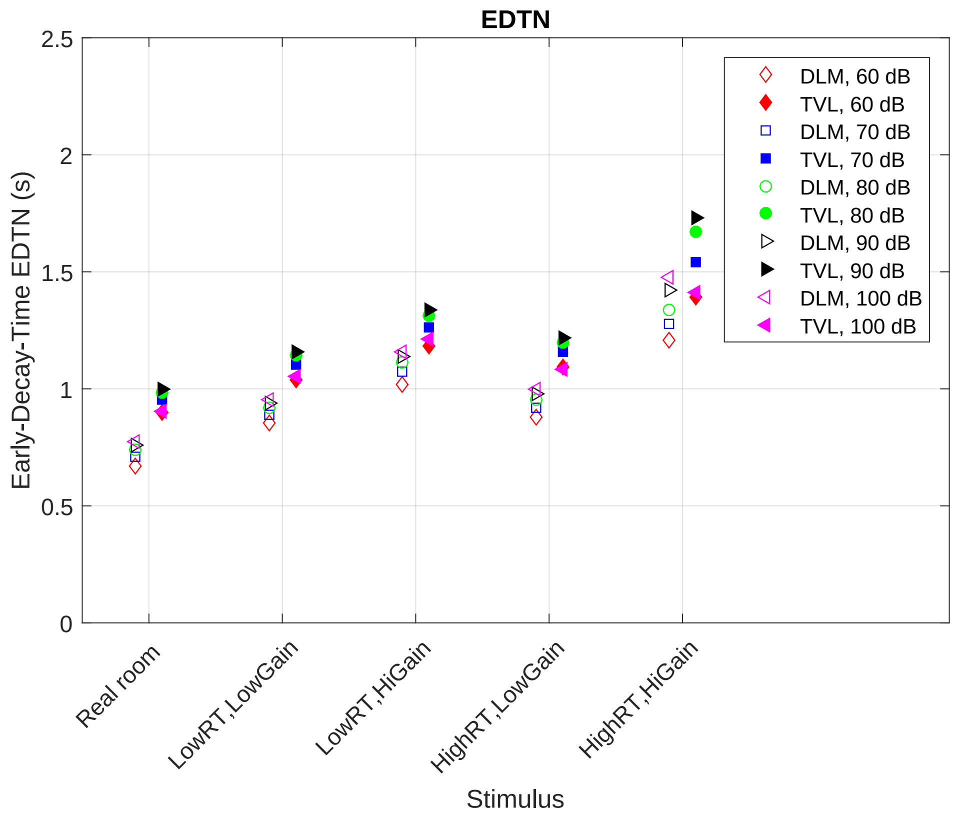


Figure 3.24. Loudness-based early decay time (EDT_N) for Loudness models DLM (left) and TVL (right, filled) for different input levels for **Experiment 1 in Section 3.3**

The results for Experiment 2 are similar (Appendix, page 147, Fig. 6.1). However, the higher level conditions output very large values: T_{N30} of up to ca. 25 seconds were given. The TVL model predicts these high values for almost every stimulus, but also the DLM model in one case. Analyzing the loudness decay for one of these situations (TVL at 100 dB for stimulus "HighRt, LowGain", Fig. 3.26) it can be observed that the noise floor of the impulse response has been increased so much that it is in the dynamic range of the regression fit.

3.5.4 Discussion

3.5.4.1 Possible reasons for the differences between models

Even though the overall estimates of the models are similar there are substantial differences between the predictors from the three models, especially in the parameter related to early decay.

The peak sone values are different: TVL gives higher values for impulse responses. This has been reported for ramped stimuli (Table 1 and Fig. 5 in [47]). A reason for this could be the different middle-ear filters. When comparing the frequency response

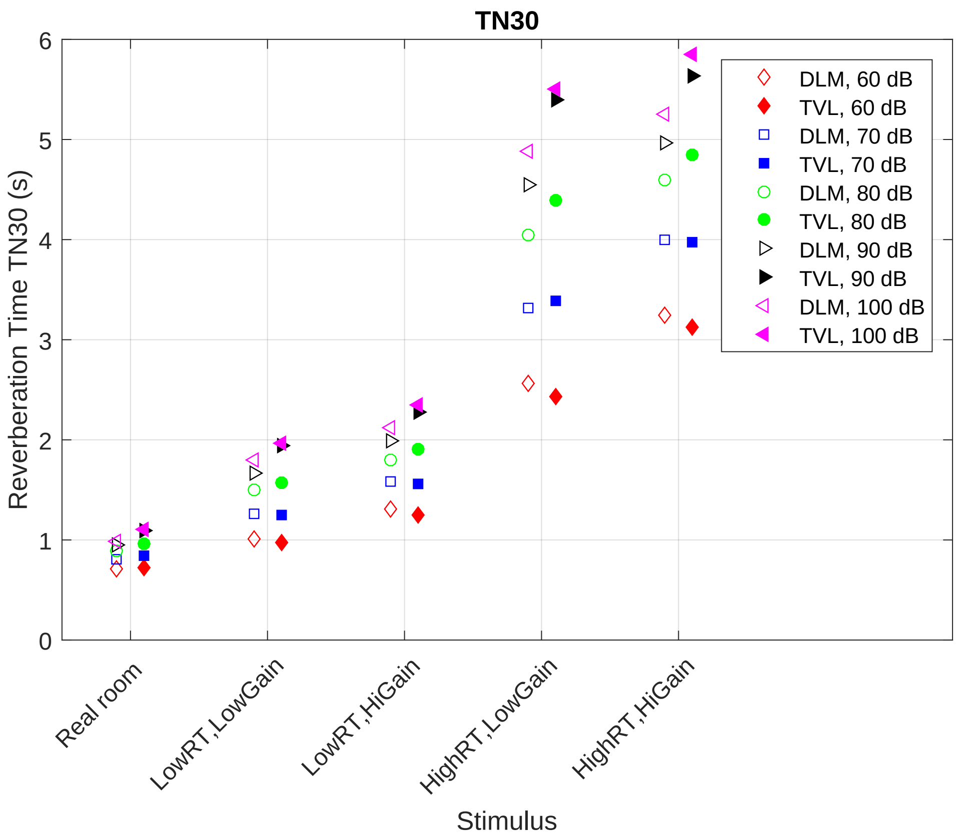


Figure 3.25. Loudness-based Reverberation Time (T_{N30}) for Loudness models DLM (left) and TVL (right, filled) for different input levels **Experiment 1 in Section 3.3**

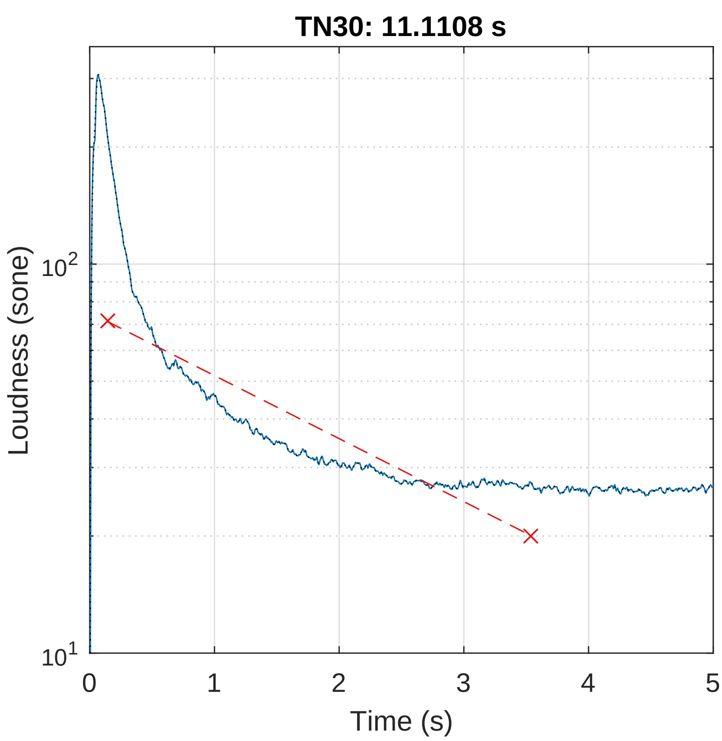


Figure 3.26. Plot of the loudness decay with regression for Reverberation Time (T_{N30}) for Loudness model TVL at level 100 dB (Experiment 2 in Section 3.3, stimulus “HighRt, LowGain”).

of the filter, TVL has about 3 sone higher amplitude at the most sensitive area (around 1-2 kHz), where the initial impulse also is likely to have the highest energy. Particularly the changes in $\mathrm{EDT_N}$, which is calculated from the early decay, suggest that there are differences in the loudness and time-function. This information could be due to the usage of different time integration constants in both models. Again, Rennies et al. reported a steeper decay for DLM (Fig. 2 in [47]). The model ISO 352-1:2016, which is based on Zwicker model is overall different as it has a much slower initial response to the impulse. Thus, the peak from the early energy is not followed as quickly. This effect has been investigated in more detail by Englberger [91, pp. 58-63].

3.5.4.2 Further considerations

The present loudness models are, as is known, not quite optimal at predicting loudness for impulsive signals. Also, subjective data is compared to estimates derived from an impulse (as in ISO 3382). Yet, in the study by Lee et al. and also partly in our current study, loudness decay of impulse responses performed well. This correlation is possibly because an IR, despite its impulsive character is after all, still a time signal of a certain duration.

As for the measurements, $\mathrm{EDT_N}$ requires listening level information which is often not given in practice. It should thus be considered to calculate the SPL from one or two typical sound sources with the given strength G from the room. Also, it should probably be calibrated to the direct sound and not the peak.

3.5.5 Conclusion

Impulse responses from the equal reverberance test have been input into three loudness models: DLM, TVL and ISO532-1:2016 delivering loudness reverberation times $\mathrm{EDT_N}$ and T_{N30}. TVL predicted a suitable $\mathrm{EDT_N}$, outperforming the conventional parameter.

TVL estimates are generally a bit longer than DLM, ISO estimates are overall shorter and inconclusive for the early decay as the ISO-model seems too slow to follow those changes as measured by $\mathrm{EDT_N}$. For TVL, any impulse response with L_{AFmax}-level of more than 70-80 dB predicts values that are too high, as a result of the amplified impulse response background noise. Model differences are likely due to filter gains and time constants complicating an independent analysis.

3.6 Reverberation gain in electronically enhanced rooms

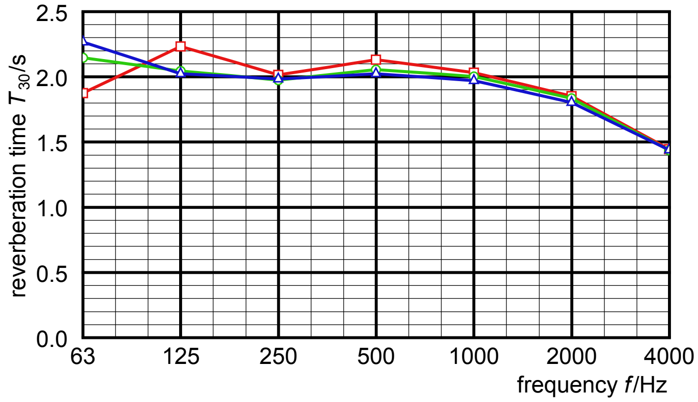
During the initial setup measurements for the previous experiments it became clear that changes in reverberation gain would always alter the reverberation time in the room (see Fig. 3.27). Thus, the slope of the decay changes as well. The ratio EDT/T_{30}

becomes larger than 1 since EDT increases (see Fig. 3.28). In the natural acoustics the room has mostly shorter EDT than T_{30} (black). Even though the changes are not excessive as such, they do exceed the JND of 5% for T_{30}.

It could be argued that neither of the two setups in this thesis, the Konzerthaus and the lecture hall, are a proper enhancement as both spaces are comparably small and a line input was used. Hence, a permanent room enhancement installation with input from microphones instead of an audio file was measured. It is a rather large, and wide, but positively rated opera venue (ca. 1400 seats) where for some selected romantic operas or concerts additional reverberation is desired, a typical scenario. A group of sound engineers was visiting the venue and was given a demonstration of the system. For the occasion, an opera singer on a professional level, accompanied by a pianist, performed two Verdi and one Mozart arias repeatedly. The system was set to an appropriate and an excessive preset by a system engineer and the author, confirmed informally by the attending sound engineers [6]. The corresponding decay times for different evaluation ranges are shown in Fig. 3.29. The reverberation times change noticeably with increasing energy from the enhancement. It is apparent that the decay slopes must be changing accordingly. For the middle, "appropriate" setting (green), T_{10} to T_{30} are fairly similar, the decay slope is only changing once. The relative reduction in the lowest frequency band is due to both aesthetic (undesirable content) and technical reasons (lower acoustic output from the speakers).

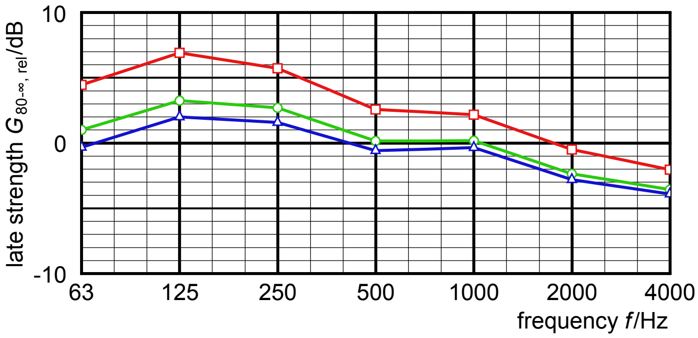
Lastly, a concert hall without electronic enhancement but instead coupled reverberation volumes is shown. As discussed in Section 3.1.4, real halls might also reveal non-linear level decays. Data from the recently opened concert hall Philharmonie de Paris, with its reverberation volumes, is compared to a conventional rectangular hall, the main hall in Tonhalle Zurich. Both halls are overall well received, although Paris was only recently opened at the time of writing. Two receivers in each hall are analyzed, both at similar distances and positions from the conductor (ca. 15 m). The data for the audience position in Paris was kindly made available by Magne Skalevik. Both halls were essentially unoccupied, hence presumably more diffuse, which theoretically results in a linear level decay. It must be noted that Philharmonie de Paris has more absorption in comparison, due to the heavily upholstered seating. An impulse from a exploded balloon was recorded in Paris, in Zurich the impulse response was generated from a swept-sine measurement. Orchestra chairs were on stage in both venues. It can be seen in Fig. 3.31 that there is a rather large delta of 0.9 s at 500/1000 Hz in Paris, but very little variation in Zurich. According to a scale model study the coupled volumes contribute to T_{30} in the inner volume by 0.3 s [82]. Occupancy in Paris reduces T_{30} to a hall average of 2.5 s according to a presentation by the acoustical

[6] The regular configuration could be somewhat different or change depending on the production.



□—□	1.90	2.25	2.00	2.15	2.05	1.85	1.45
○—○	2.15	2.05	1.95	2.05	2.00	1.85	1.45
△—△	2.25	2.00	2.00	2.00	1.95	1.80	1.45

□—□ Fixed internal RT, Gain -10 dBFS
○—○ Fixed internal RT, Gain -15 dBFS
△—△ Fixed internal RT, Gain -17.5 dBFS



□—□	4.4	6.9	5.7	2.6	2.2	-0.5	-2.1
○—○	1.0	3.2	2.7	0.2	0.2	-2.4	-3.6
△—△	-0.3	2.0	1.6	-0.6	-0.4	-2.8	-3.9

Figure 3.27. Measured reverberation times (top) and -levels (bottom) in Konzerthaus Detmold for different gains of the enhancement system set to the same internal decay time.

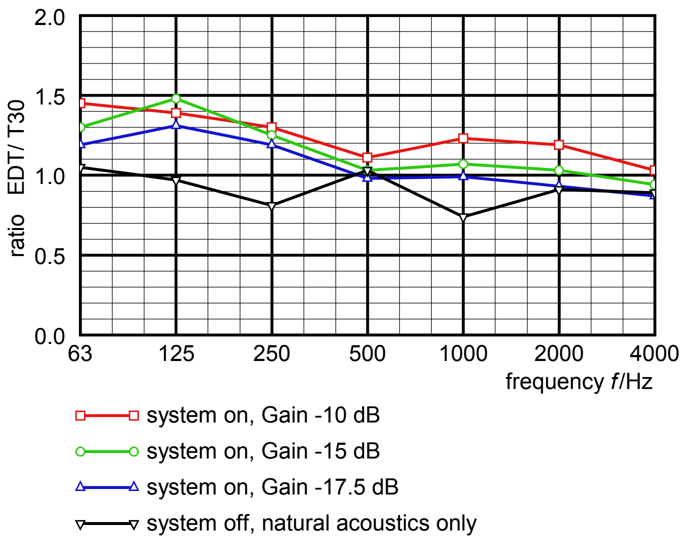


Figure 3.28. Ratio of EDT to T_{30} for the setup in Konzerthaus Detmold.

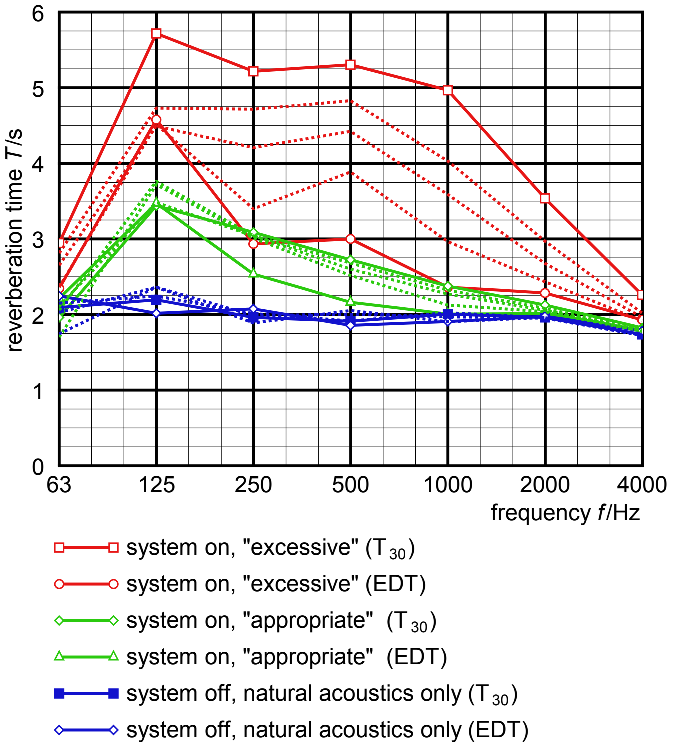
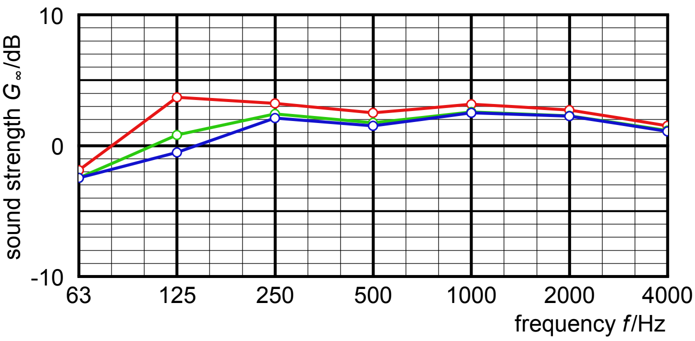


Figure 3.29. Reverberation times for different evaluation ranges in an opera venue with enhancement system. EDT and T_{30} are labeled and in solid lines, dotted in between are T_{10}, T_{15} and T_{20}.



	-1.9	3.7	3.2	2.5	3.2	2.7	1.5
	-2.5	0.8	2.4	1.7	2.6	2.3	1.1
	-2.5	-0.5	2.1	1.5	2.5	2.2	1.1

system on, setting "excessive"
system on, setting "appropriate"
system off, natural acoustics only

Figure 3.30. Strength values G for three acoustic conditions in the opera venue.

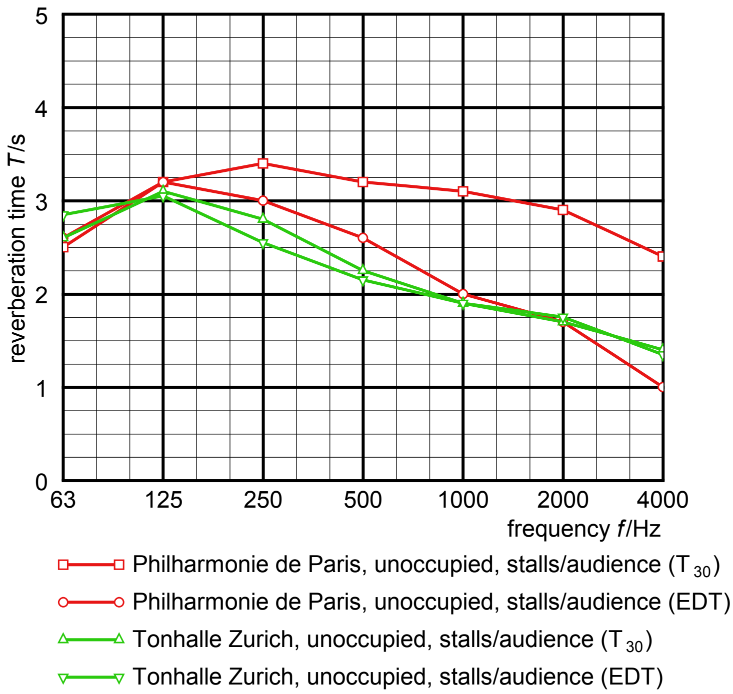


Figure 3.31. Reverberation times for evaluation ranges T_{30} and EDT in Philharmonie de Paris and Tonhalle Zurich.

consultant Chris Day at the IOA Paris 2015 (values not published in writing), the early decay time EDT is unknown.

3.6.1 Discussion

The somewhat unexpected non-linearity of level decay is, at closer inspection, explained by the change in the diffusion or equivalent absorption area when adding electronic reverberation. A similar effect appears when absorption is distributed unequally in a room. As this affects the linearity of the decay, the evaluation for decay range parameters such as EDT and T_{30} generate different values.

Reported decay time values from real halls had on average 6% shorter EDT than T_{30}, ranging from -21% and +6% [83]. This range is about that found here for the non-excessive enhancement settings. Some non-linearity might actually be liked [80, pp. 126-129]. Those particular listening tests were done with decay times of T_1 = 1.4 s and the coupled volume T_2 = 2.2 seconds, a ratio of 0.65. Early decay time can be heavily influenced by early reflections. For the experiments shown here, the early strength (G_{80}) changed around 0-0.7 dB at 250 Hz between stimuli compared to late strength (G_{late}) with 2-14 dB. Thus, early energy is not seen as the main influence.

It could be argued that too massive a change of the decay slope is perceived as unnatural. This idea might be true, though the author assumes that the overall reverberation gain and the resulting reverberance are more crucial and the non-linearity is a side effect. In the real enhanced space this cannot be tested as the natural acoustics are fixed in level. However, a successive laboratory test where the exact sound field is convolved and the reverberant part presented at different levels with the same non-linear decay would likely confirm this.

3.6.2 Conclusion

Multi-slopes or non-linearity in the level decays seem to be inherent to the technique and utilization of room enhancement and is likely not a sign of bad quality. The phenomenon also appears to a lesser degree in real venues where it depends on the design: Coupled-volume halls exhibit a non-linear level decay.

3.7 Discussion and conclusion

In this chapter, the state of research and further developments in perception and measurement of reverberation were investigated. In the framework of semi-virtual sound fields, a room enhanced lecture hall and a medium-sized concert hall could be successfully used for investigations in this combined laboratory and in situ environment.

Firstly, it was shown that an enhancement system is for the given situations indeed judged to make the reverberation more appropriate in a concert hall, both by normal and expert listeners, the latter being more precise in their answers: A slight gain in the room enhancement system was subsequently preferred, resulting in an increase of the overall reverberation energy of around 1 dB in the concert hall. The change appears small considering just noticeable differences of 1 dB. Reverberation times were raised by approx. 0.5 s. On the other end, too much artificial reverberation was found to be inappropriate as well, possibly even more so than too little reverberation as was shown in related research for automated reverberation mixing [87]. This conclusion suggests that only a well set enhancement is judged as appropriate which would explain the somewhat controversial experiences from musicians, for example.

Independently varying the resulting reverberation time and level did not appear possible, as is discussed in section 3.6. Thus, it is unlikely that the system engineer indeed controls decay time or level independently, as was claimed to be the case.

Therefore, it also remained unclear whether the lengthened decay or the higher reverberation level was more decisive. Two experiments with varying reverberation time and level (magnitude estimation) and a reverberance matching experiment (magnitude adjustment) were conducted. It was seen that the reverberation level played a more important role in predicting reverberance.

From the standardized ISO-3382:1 parameters the most consistent performance was given by energy parameters such as Strength G which is currently affiliated more with perceived loudness. The recommended parameters from ISO-3382:1, Reverberation Time T_{30} and early decay time EDT, did not show as consistent performance. However, it would not be reasonable to claim that strength parameters G_{inf} or a time-windowed version such as G_{late} are always superior as it strongly depends on the set of stimuli. Nevertheless, in agreement with the current state of research, additional evidence was presented that overall room gain or reverberation level as measured by G_{inf} and G_{late} is very important to consider.

As one recent development, loudness-based impulse response analysis was investigated: Using a loudness model with level-calibrated impulse responses the approach presents a intermediate of the purely physical evaluation (ISO 3382) and more strongly perceptually motivated analysis (e.g. [79]) [7]. Subsequently, good performance was observed for the reverberance matching experiment. Meanwhile, similar positive results were reported from other researchers for different topics such as reverberance in audio mixing [93], speech intelligibility ([94]), echo analysis [72] (supervised by this author). However, it was observed that that there is an influence of the type of loudness model in use: Dynamic Loudness Model (DLM) and Time Varying Loudness (TVL) as openly available and well established models showed strong level-dependent differences, especially for measures connected to the early part of the impulse response (EDT_N). As neither of the two models/codes is currently standardized, the draft of the recent loudness standard ISO 532-1:2016 [90], closely related to DIN45631/A1:2010 [95], was investigated (more closely in a master thesis supervised by this author [91]) where even stronger model influences were exhibited. This leads to quantitative differences in loudness decay times T_{N} or EDT_N depending on the model in use which impedes practical investigations and comparisons. Considering the additional complexity and degrees of freedom when including the loudness models this alternative approach might be argued to be viable judging from the given set of stimuli.

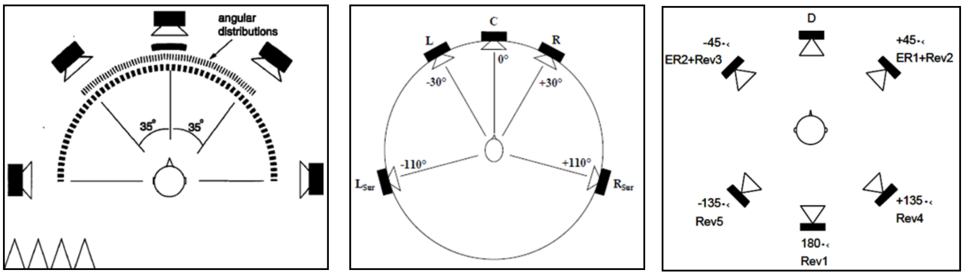
[7] The author co-wrote a study [92] testing estimation of acoustic percepts from recordings. It was found to be functioning but at a comparably early development stage, e.g. level dependency was too strong, the original model fit was not valid for presented stimuli etc.

4. Spatial distribution of reverberation and envelopment

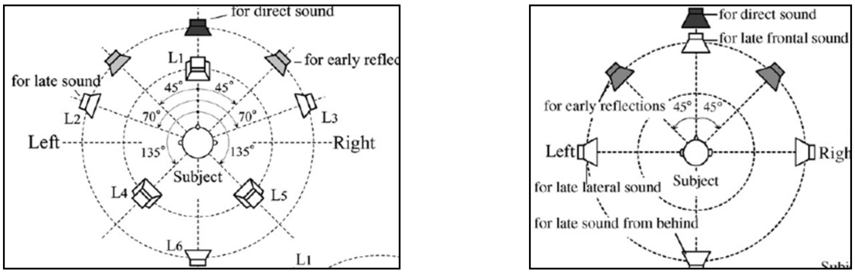
4.1 Introduction

In Section 1.1.2 it was discussed that auditory perception in concert halls can be described thoroughly with a number of attributes. The previous chapter dealt with the attribute *reverberance*, related to the perception of level and duration of reverberation. An important spatial attribute related to late energy is *listener envelopment* (LEV), defined as being surrounded by sound. LEV is important for the concert atmosphere (attribute *immersion*), and is a quality factor of a good concert hall [19, p. 30] (see also [96]).

The concept was put forth by Morimoto et al. and it was argued that the ratio between frontal and rear reverberation would be a measure of LEV [97]. This concept was rendered more precisely and investigated by Bradley et al. [98], where it was shown that LEV is primarily affected by the level and direction of late energy in the room, as opposed to the early reflections influencing the perception of the source width. Here and in a successive study [99], it was observed that the late energy from lateral directions is most influential. Evjen et al. [100] extended the procedure by including 8 speakers in total. These were alternated to different elevated and backward directions, thus there was only a spatially somewhat incomplete sound field. It was observed that late lateral level still offered the best prediction as a measure, thus confirming previous findings by Bradley. Front and rear energy were found to have a similar, reduced influence on LEV. Energy from above showed small effects and was interpreted as an artifact. Around the same time, Morimoto emphasized that late energy from behind did increase LEV and that the front-to-back ratio would be decisive [101] which is in contradiction to Evjen and Bradley. Another study found again a stronger relevance of lateral late energy but contribution from the front as well. Separate comparison of directions was found to be misleading and that the overall “spatial balance” must be adequate [102].



(a) Bradley et al. [98] **(b)** Soulodre et al. [103] **(c)** Morimoto et al. [108]



(d) Furuya et al., Experiment 1 [105] **(e)** Furuya et al., Experiment 2 [105]

Figure 4.1. Typical laboratory setups used for investigating listener envelopment.

The measures established until then, late lateral sound level L_J (associated with LEV in ISO 3382 [6]) and late interaural cross-correlation coefficient IACC$_{late}$ (see next section) were revised by Soulodre et al. [103], [104]. By introducing different integration times per frequency band for omni-directional strength G, to account for perceptual properties of the hearing system, and connecting this level information with IACC or LF for the spatial component, envelopment could be predicted well for a surround setup. Furuya et al. [105] experimented more with different directions of the late reverberation, finding an importance of approximately 60 and 35% for late rear and late ceiling reverberation compared to late lateral energy. Beranek [106] noted the results from these studies and proposed a practically oriented approach to predict listener envelopment: combining G_late and IACC, Beranek calculated values for his database, yet without showing correlations. Kahle et al. [107] pointed out the importance of the directional late energy of the sound field when observing, for example, the ceiling of a real concert hall absorb too much reflection energy. This absorption could be due to a double ceiling, technical installations or possibly Schroeder diffusers' residue absorption. By installing reverberation enhancement systems this lack of late reverberation from certain room directions can be compensated for.

With the foregoing research, there is an unresolved discussion whether lateral or omni-directional late energy is more accurate for predictions of envelopment and how the different directions contribute, so stated by Beranek in 2010 [106]. The ISO-standard suggests only late lateral energy as a predictor, measured with a figure-of-eight microphone.

Also, most of the above mentioned studies regarding the direction of late energy and LEV were laboratory studies with fully synthesized stimuli, as only in this environment the necessary degrees of freedom were possible to realize – certainly also due to technical limitations of the time. This situation resulted in rather simplified reproductions of the acoustic situations compared to the complexity of the sound field in a real venue, specifically accurate reproduction of the direct sounds and reflection paths of multiple sources or by oversimplifying both early reflections and the multitude of late reflections when "panning" into a few loudspeakers (Fig. 4.1). Lastly, the loudspeaker distribution was often not uniform, namely no rear or ceiling loudspeakers (see Fig. 4.1a). Partly, the stimuli were set to realistic room acoustic values. Few of the above mentioned studies provided sufficient consideration and discussion on how the stimuli were calibrated, which seems important when comparing reverberation directions.

The virtual acoustic group in Aalto started conducting experiments with an elaborate re-synthesis. Several reputable, real concert halls were measured with a multi-channel loudspeaker orchestra [24] and recreated in a 24-loudspeaker laboratory set-up. By panning the measured hall impulse responses at each sampling instant to the speakers closest to the actual directions of arrival [25], and convolving a multi-ch anechoic orchestra with the resulting multi-ch room impulse responses, a noticeable step toward realism has been achieved. However, no manipulations to the direction of the late sound field had been investigated up to now.

The purpose of the following experiments is to investigate the influence of different late energy directions for more realistic sound fields in order to validate the results obtained using fully synthesized sound fields. This research is done by presenting sound fields including all dimensions and following an increasing scale from lecture hall over chamber music venue to two different large concert halls, see Fig. 4.2. The goal is, furthermore, to investigate the standardized measures and discuss implications for room acoustic and enhancement system design.

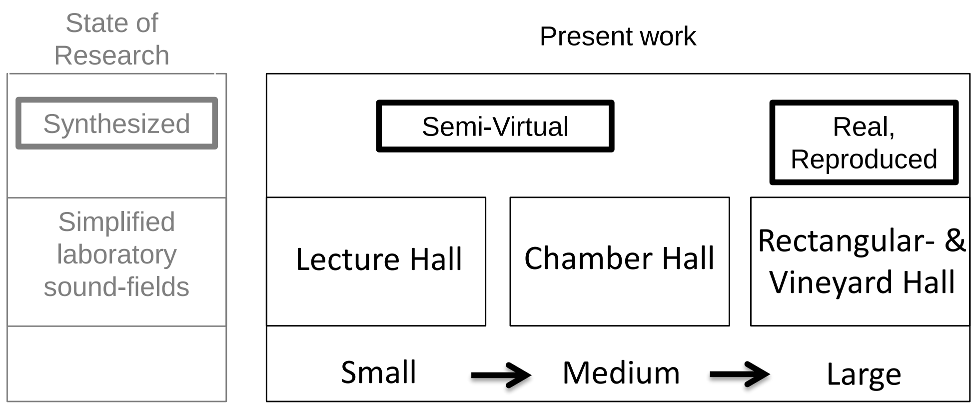


Figure 4.2. Schematic overview of experiments regarding spatial distribution and listener envelopment in the present thesis.

4.2 Small venue: lecture hall with semi-virtual acoustics

4.2.1 Setup

The basic setup was similar to the previous experiment in the lecture hall, see Fig. 3.16 on page 49. A single directional loudspeaker served as the source for the direct sound (Genelec 8030A). However, in this experiment a clapping sound was used as a stimulus since the pre-tests showed that this impulsive signal provides easier cues for detection of directions than most music stimuli while still being a realistic signal. The clapping sequence was a four seconds long anechoic recording of five similar claps, presented as a continuous playback loop, thus providing some reverberation tail without a too big gap between repetitions (5.5 s with silence). As before, the line signal was used as a signal for the Vivace reverberation processor. The 52 loudspeakers were divided into 4 groups corresponding to room directions “Front”, “Side”, “Ceiling” and “Rear”. Figure 4.3 shows the distribution, the decision for distributing the loudspeakers to the respective directional groups is somewhat arbitrary and was done according to common sense as well as to have a similar amount of loudspeakers per group. A setting “All” and an attenuated “All-6dB” setting with the same (monaural) energy as the other directional stimuli were used as anchors.

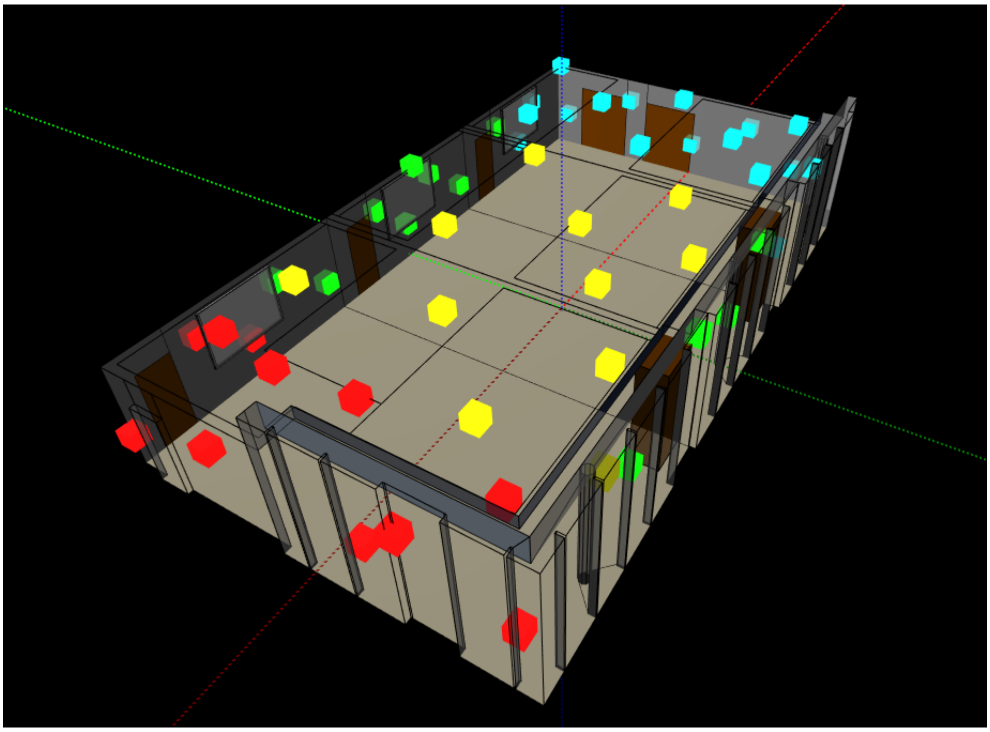


Figure 4.3. 3D-Model Lecture Hall with loudspeakers and directional late energy groups (colored): “front” speakers (blue), “rear” speakers (red), “side” speakers (green) and “ceiling” speakers (yellow).

The loudspeakers were first matched individually in level at the main listening position with pink noise and delayed to arrive at similar points of time. Then, levels were matched among the four loudspeaker groups with white noise. This setting was later further adjusted with the music stimulus playing as there were differences between the perceived stimulus reverberation loudness. The Vivace-Preset “Chamber Hall-1-7” was used, with the first 80 ms of the impulse response muted. Only the late energy was changed, direct sound and early reflections stayed constant. One of the four reverberation engines in Vivace was routed to each loudspeaker group. The speaker groups were then further equalized by performing impulse response measurements and matching the ratio of early to late energy (omnidirectional clarity C_{80}), see Fig. 4.4. Unfortunately, it was not possible to fully match both C_{80} and listening levels as they were affecting each other. The resulting listening levels are shown in Table 4.1, measured with a Class 1 SPL meter. Between directional groups there are small level deviations of around 0.5 dB. The background noise was measured at L_{Aeq} = 24 dB.

Table 4.1. Stimulus sound pressure levels

Condition\ Level	L_{Aeq} [dB]	L_{AFmax} [dB]
front late reverb	60.3	67.9
side late reverb	60.5	68.1
ceiling late reverb	60.9	68.9
rear late reverb	60.9	68.1
all late reverb	65.6	72.8
off/natural acoustics	56.6	67.1

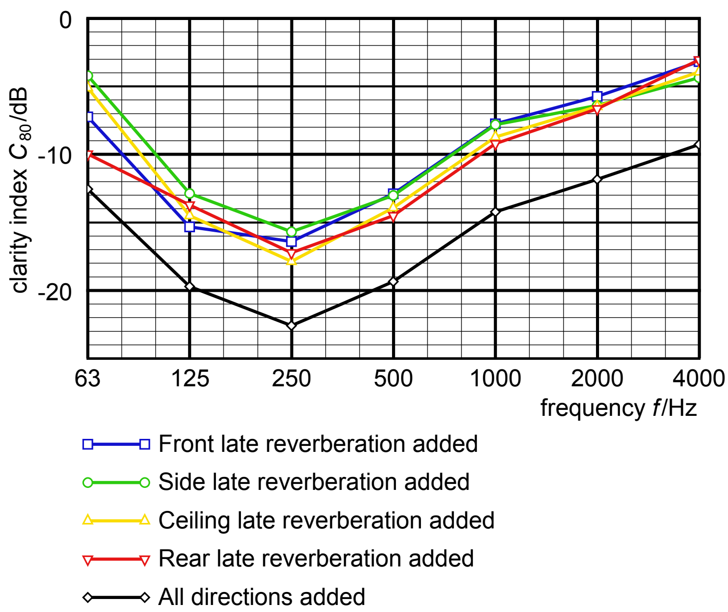


Figure 4.4. Measured Clarity C_{80} for different late energy directions/stimuli.

For deriving the directional parameters, measurements were conducted with a microphone array (see Section 2.3.3). Visualizing the different reverberation directions (see Fig. 4.5), one can see that the direct sound (red, from the right) and strong earlier reflections (green) stay constant, whereas the late energy changes the direction (light blue). The plot for the ceiling reverberation (4.5e) reveals some unwanted imbalance to the right of the listener.

The listening test was done with a paired comparison design (see Section 2.4), testing each of the situations against each other and choosing the dominant option regarding listener envelopment (LEV), apparent source width (ASW) and clarity of sound. Envelopment was defined as the feeling of being surrounded by sound, ASW as perceived size or width of the sound source, and clarity as the ease of perceiving and clarity of the sound. The participants entered judgments on a user interface on a laptop.

The experiment was an individual test with one person at a time. 20 Employers (n=20) of Müller-BBM took part in the test of which half could be considered expert listeners. Most of the participants had some background in acoustics but varying levels of listening experience/training, the whole group could thus be called an advanced (not expert) listening panel. The average age was 42 years (median 40 y, 17 male and 3 female) with a duration of 11 min on average (median 10 min), shortest completion of 7 min and longest of 22 minutes.

4.2.2 Results

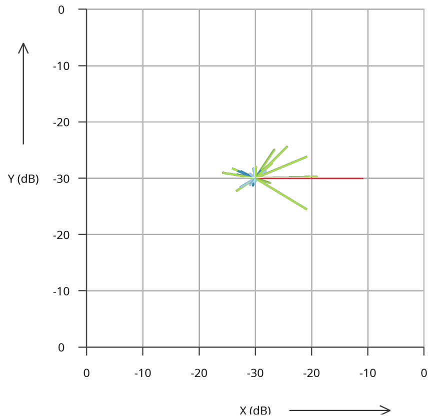
4.2.2.1 Listening Test: Envelopment

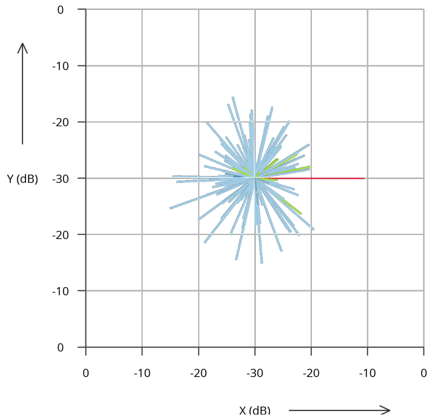
Fig. 4.6 shows the results for listener envelopment. It becomes clear that late reverberation from the side offers more envelopment than the other directions. In decreasing order from normalized uniform distribution (“All-6dB”) to rear, ceiling and front reverberation.

It can be concluded, that the overall reverberation level dominates the direction of late reverberation for LEV. In this experiment, side reverberation is the most effective, followed by rear, ceiling, front. Equal late energy from all directions (“All-6dB”) is less enveloping than from the side only. However, other directions do also contribute to LEV and of those mostly the late energy from behind.

4.2.2.2 Measurements

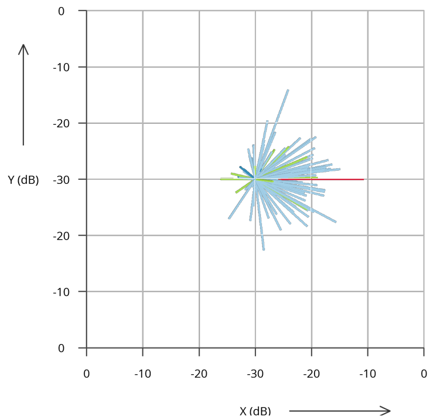
The previous section introduced interaural cross-correlation coefficient (IACC) as one measure of spaciousness. To demonstrate the effect of the different directional reverberation, IACC over time is calculated in steps of 25 ms and plotted over the duration of 2 seconds. Fig. 4.7 shows this for the stimuli “front”, “side” and “all

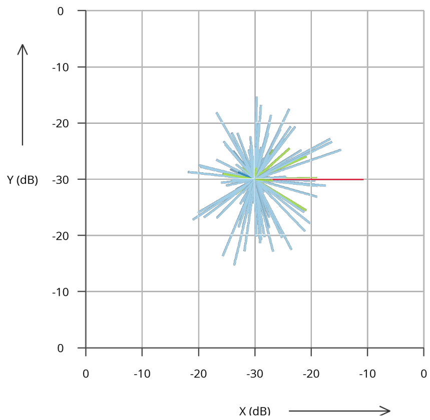




(a) No late reverberation added

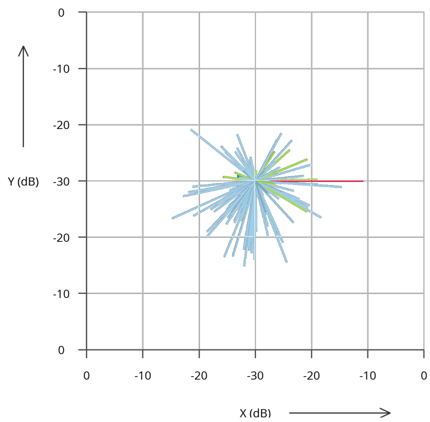
(b) Late reverberation added from all directions

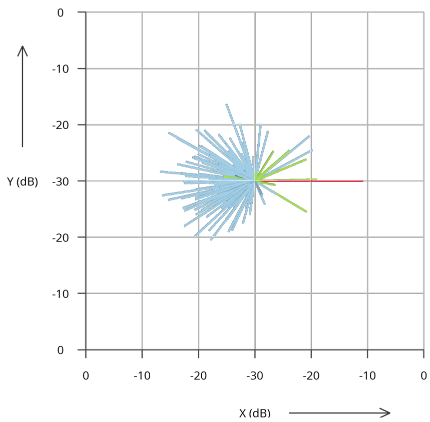




(c) Frontal late reverberation added

(d) Side late reverberation added





(e) Ceiling late reverberation added

(f) Rear late reverberation added

Figure 4.5. Ground view/lateral plane spatial plots. The direct sound arrives from the front (red color, shown to the right in each figure). b) to f) are stimulus sound fields with different directions of late energy (resolution 4 ms). Energy after 80 ms is shown in light blue.

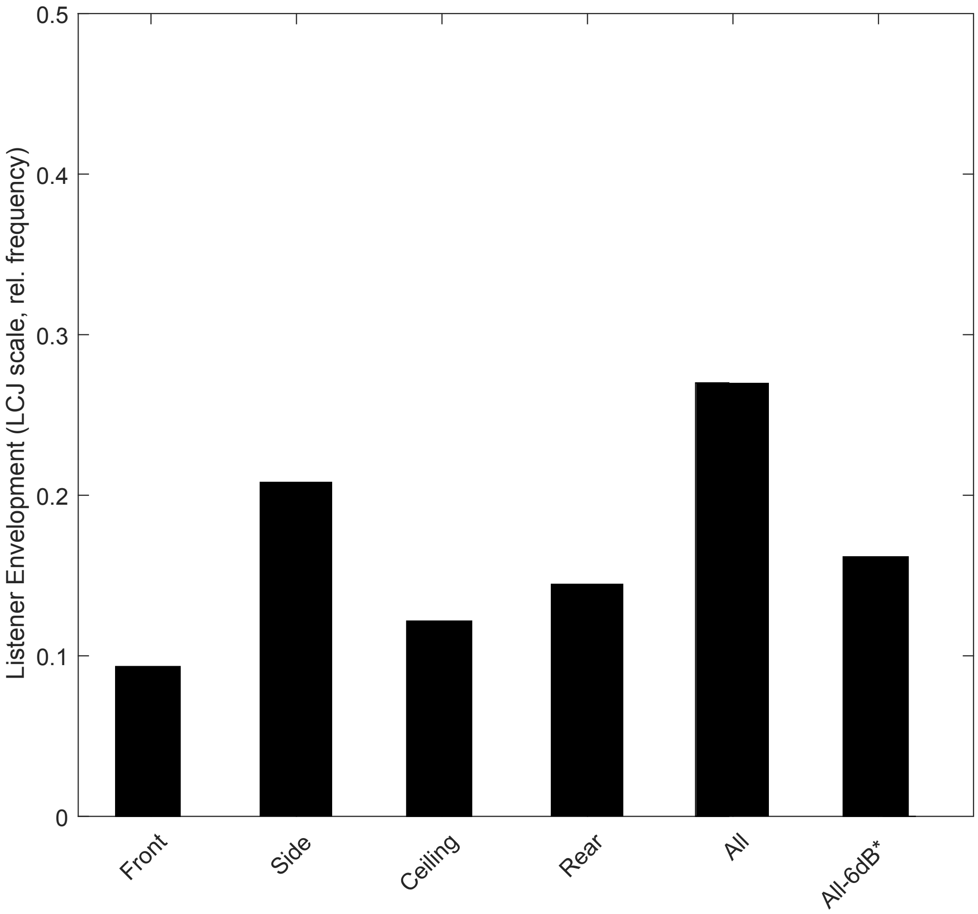


Figure 4.6. Envelopment (LEV) for different directions of added late energy.

directions". Until ca. 100 ms IACC stays constant as the early energy is not manipulated, but changes afterwards. Then the frontal reverberation increases the most, as expected, since the incident sound is fairly symmetrical for both ears. The side speakers are apparently the most dissimilar and yield the lowest IACC. In between, the green line is the mixture of all directions ("All"). This result is according to expectation, however, most envelopment was offered by the condition "All" due to the increased reverberation level. Thus IACC does not predict envelopment well, more discussion on this is given in the following Section 4.3.

The second approach is to analyze energy parameters. Lateral late strength (L_J) measures the energy after 80 ms in a figure-of-eight direction pattern. Lateral late energy is averaged between 125-1000 Hz and calculated from the directional impulse response. Late Strength G_late measures the omnidirectional/monaural energy after 80 ms. Values for both parameters are shown in Fig. 4.8. It can be seen that the trend is predicted similarly and in accordance with the listening test results for envelopment. On closer inspection, "side" energy is more accurately rated with late lateral level L_J. When correlating the listening results with the measurement parameters coefficients of 0.89 for L_J, 0.82 for G_late and -0.80 for IACC were computed.

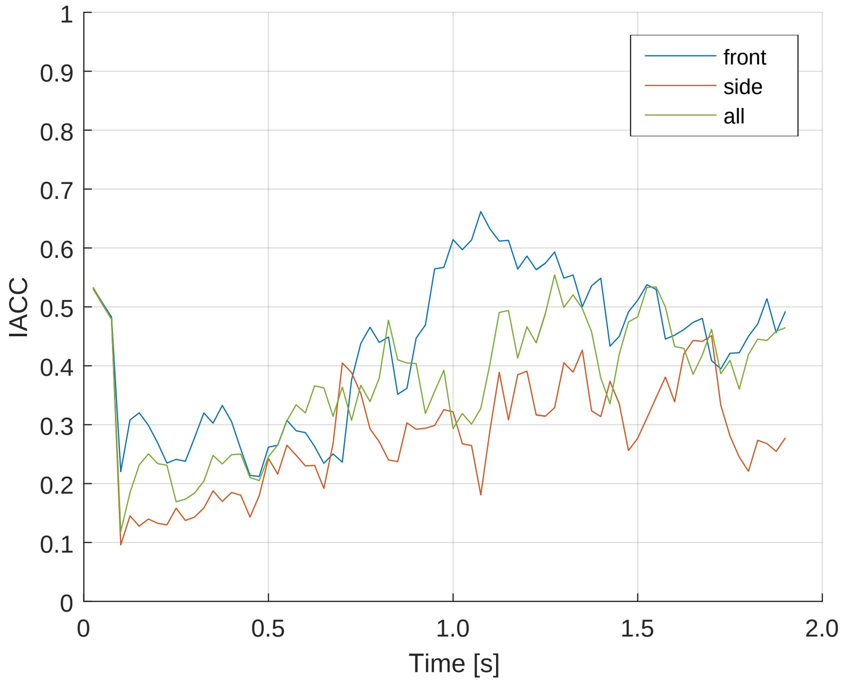


Figure 4.7. IACC over time calculated in 25 ms windows, unfiltered. Artificial late reverberation was added from the front speakers (blue), all speakers (green) and side speakers (red).

4.2.2.3 *Apparent Source Width and Clarity*

Even though ASW is a perceptual descriptor of the size or width of a sound source, and usually thought to be affiliated with early reflections (before 80 ms, which were not varied here), it was asked for as well. Results are shown in the Appendix in Figure 6.4, page 149. The overall raised late level ("All") increases ASW. A dependency of ASW on the overall late energy is observed, which is somewhat surprising and probably driven by the loudness gain. The direction does not seem to matter too much, but a trend can be observed that front and rear reverberation raise ASW more whereas late energy from the side gives the smallest source width.

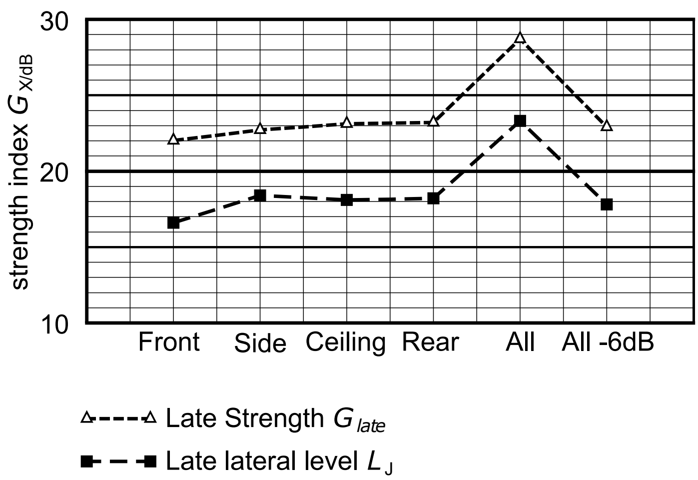


Figure 4.8. Late sound strength and late lateral level for the different late energy directions. Strength reference calculated from room volume and measured reverberation time.

Lastly, a hypothesis based on practical experiences was to be investigated, that is, if clarity would be decreased when the late reverberation comes from certain directions such as frontal. The hypothesis could not be confirmed in the experiment (Fig. 6.5 on p. 149 in the Appendix). No noticeable differences between individual directions regarding clarity appear, particularly, the direction of reverberation does not seem to play a role for perception of clarity. This finding is interesting as it was reported that late energy from the front could be more detrimental (i.e. Kahle et al. [107]). However, it should be considered that the stimulus was only a clapping noise and some participants even mentioned that they had trouble to associate this sound with *clarity* as such. Also, in the beginning it was asked for a preferred stimulus. Without an aesthetic context the question for preference did not appear meaningful and was therefore discarded.

4.2.3 Discussion

The results are mostly in agreement with current literature for envelopment, side reverberation seems to be more effective and late lateral level a good correlate. Yet, there is a contribution of rear reverberation which was not clear from all previous studies. Also, most earlier studies did not equalize the energy between directions.

Some comments were collected informally after the test: two participants actually noticed the unintended slight shift in direction for the ceiling stimulus. Only some participants noticed during the test that the direction of a sound field component was changed. When revealed, participants turned their heads and noticed quickly. The difficulty of the task was judged between overall rather easy and rather hard depending on the pair. The sound was rated between good to rather artificial, the gap between early reflections and the late sound from the electronic reverberation was mentioned a few times. Some participants had difficulty with the attributes and terminology for the given stimuli or task.

4.2.4 Conclusion

Overall reverberation level dominates LEV, hence spatial distribution of the reverberation is overall less important. Between directions, side reverberation is most effective, but rear reverberation contributes as well. Late lateral level outperforms IACC for predicting LEV. Interestingly, the source width is broadened with excessive late energy. Perceived clarity of sound was not affected by spatial distribution of reverberation here.

4.3 Further observations for interaural cross-correlation over time

The parameter interaural cross-correlation (IACC) measures the similarity or coherence between the two ears with a maximum of "1" (similar) and "0" (dissimilar). The maximum of the interaural cross-correlation function is calculated for each millisecond and then averaged along a sliding window of 10 ms over one second [19, p. 524]. In accordance to the energy criteria such as clarity C_{80} time windows of 0 to 80 ms and 80 ms to infinity are averaged to single values for IACC$_{early}$ and IACC$_{late}$ respectively. IACC is also calculated in octave bands and then averaged over 2-3 frequency bands, from 500 to 1000/2000 kHz.

4.3.1 Konzerthaus Detmold

In the following, IACC over time is analyzed for the experiment in Konzerthaus Detmold from Section 3.2 where artificial (late) reverberation was added, starting from the natural acoustics (c1) in six steps (c2-c7), where c3/4 were the most preferred. For every condition, IACC is calculated from the binaural impulse response (measured with Neumann KU100) for a window size of 10 ms and shown unfiltered in Fig. 4.9.

It is interesting to note that there is a rather steep drop only after the first 100 ms, the earlier reflections are still fairly coherent. The early progress is hardly influenced as the artificial late energy does not increase the energy before 150 ms much. Also, the more noticeable changes happen after one second. In the unfiltered/broadband analysis IACC increases with additional reverberation. In the lower frequency bands (125/250 Hz, not shown here) IACC is, as expected, close to 1 and drops with increasing reverberation. Only little is seen in the other bands, even though those are normally used for single value calculation.

4.3.2 Berlin Konzerthaus and Philharmonie

A second example with a different comparison is investigated, Berlin Konzerthaus (BK) and Philharmonie (BP). One seat in each of these two halls is tested and analyzed in a subsequent section regarding envelopment (Section 4.5, page 90 ff.). For orientation: BK has a longer reverberation time of 2.1 s (BP 1.9 seconds) and about 2.5 dB more late lateral energy. The binaural recordings were done at the same distance in both halls with one directional speaker at the soloist position on stage (data from Lokki et al. [12]). The early reflections are quite different between the two halls. Fig. 4.10 accordingly shows that IACC$_{early}$ is lower in BK than in BP, which is in good agreement with measurements from Beranek [19, p. 509]. Late IACC is given there with ca. 0.13 (125-500 Hz) for hall averages in both venues [19, p. 526]. In the unfiltered condition

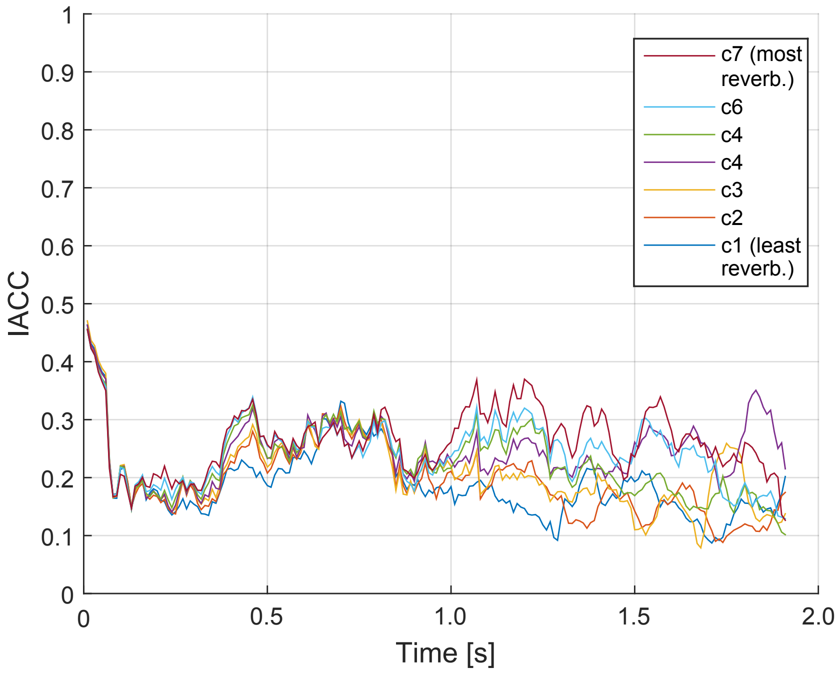


Figure 4.9. IACC over time, unfiltered, for increasing amount of reverberation ("c1" with least reverberation to "c7" with most) in Konzerthaus Detmold, energy before 80 ms stays constant (setup see Section 3.2).

4.10, the hall with *more* late reverberation (Berlin Konzerthaus) again *increases* in IACC after roughly 0.5-1 s.This result is likewise when averaging over all 24 source loudspeakers on stage (see Fig. 6.3 on page 148 in the Appendix).

4.3.3 Discussion

From the results it does not appear to be the case that low values of late IACC are necessarily a criterion for reverberation. This finding is in contradiction with the general opinion that low late IACC values are preferred. It is expected that IACC decreases with greater amount of late, de-correlated reverberation, which is considered desirable for the listening in concert halls. Ando described the connection between spatial impression and cross-correlation [109].

However, according to the ISO standard 3382 [6] there is currently no consensual agreement and understanding on its relevance. It might have to do with the fact that most observations were made in fairly idealized sound fields. Beranek analyzed correlations between IACC and the questionnaire ranking list. No connection was found between envelopment and IACC$_{late}$ for the hall averages of venues ("...as a measure of sound diffusion [...] it appears to have little value", [19, p. 525]). The observations from the present experiments are thus in line with the assessment of IACC in ISO 3382 ("not been accepted uniformly", [6, p. 28] and Beranek's judgment regarding IACC$_{late}$. Also, most differences seemed to arise after 0.5-1 s whereas IACC$_{late}$ was suggested to be analyzed only up to 1 second.

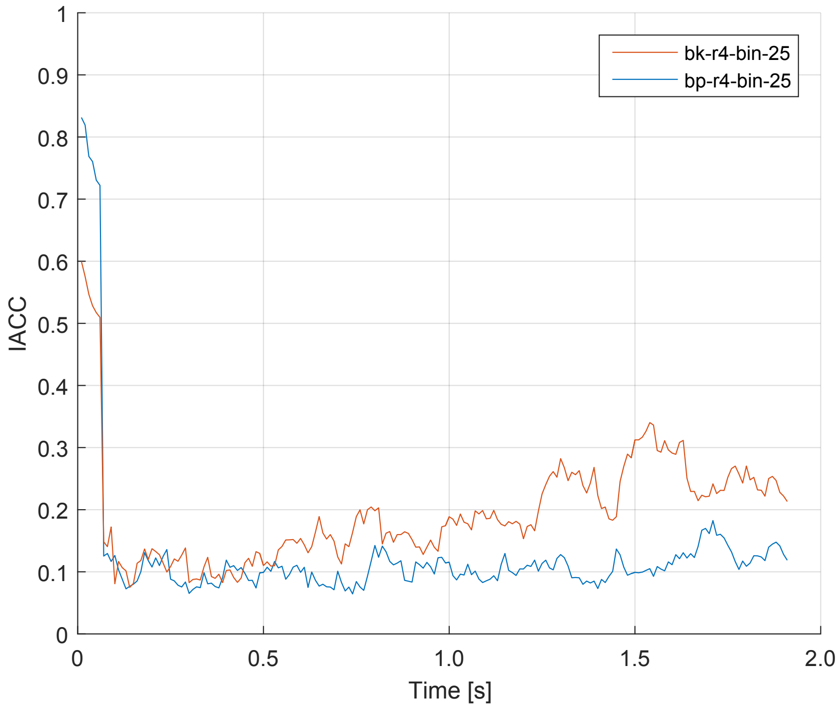


Figure 4.10. IACC over time, unfiltered, for one source/receiver combination in Berlin Philharmonie/Konzerthaus (see Section 4.5).

On the other hand, IACC$_{early}$ seemed to correlate well with overall quality leading to a "binaural quality index" (BQI = 1-IACC). Another recent study showed IACC/BQI-values from actual halls and compared to listening test results, with somewhat contradicting results such as higher preference with increasing IACC [110]. An important general observation was the connection between sound-level, IACC and envelopment [111]. If the late reverberation is not heard enough because either the signal is too quiet (or contains too little late energy), IACC does not matter. This conclusion is quite logical and might well explain the previous findings. Lastly, Klockgether et al. recently found that due to the low sensitivity of listeners "small differences in IACC in real rooms will not be distinguishable" [112].

4.3.4 Conclusion

The usage of the parameter interaural correlation-coefficient was discussed along with results for IACC over time of some experiments in this thesis. The parameter, affiliated with spatial attributes such as envelopment, seems to be inconclusive and less efficient as the level influence is not accounted for.

4.4 Medium-sized venue: chamber hall with semi-virtual acoustics

The same question is investigated in a more musically relevant context. In this experiment a music stimulus is used and the late sound field in a chamber hall is altered with electronic room enhancement distributing reverberation to four room planes.

4.4.1 Setup

The basic setup in the Konzerthaus was described in Section 2.2.1.2, page 18. The anechoic audio stimulus here was an anechoic saxophone recording with a dynamically relatively steady line of 37 sec. A directional studio loudspeaker (Neumann KH120A) played the direct sound signal and was placed in the center of the stage (see Fig. 4.11, height: 1.5 m). The receiver was located at a distance of 13.5 meters in the second row of the permanent seating area (seat no. 163). The distance to the ceiling is 7 meters. The sound pressure level without additional reverberation was set by ear to provide a realistic listening impression (L_{Aeq} = 55 dB, L_{AFmax} = 67 dB). During or after the test none of the audio engineer participants mentioned this being too loud, too quiet or unrealistic.

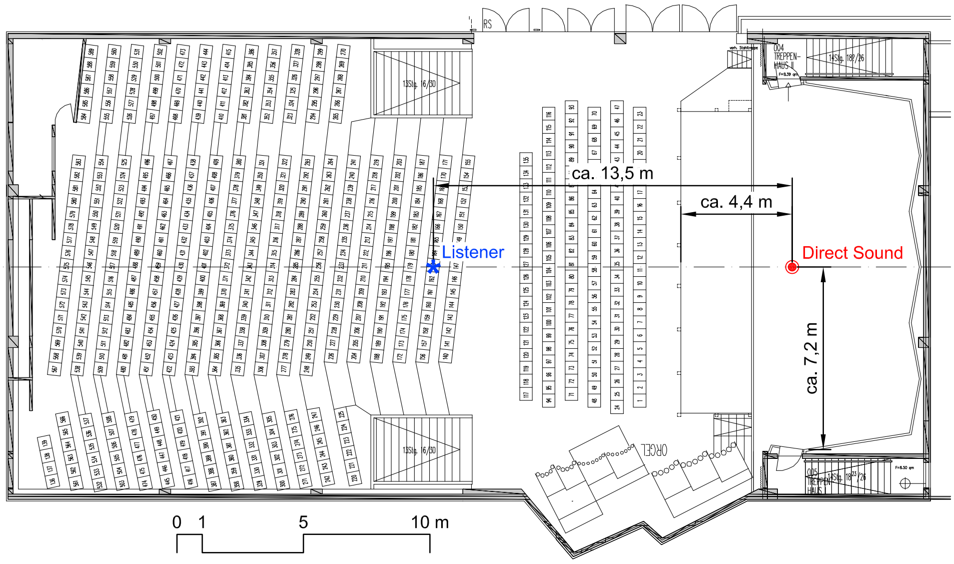


Figure 4.11. Konzerthaus Detmold with the receiver (left) and the direct sound loudspeaker on stage (right).

As before, the Vivace reverberation engines were distributed to four speaker groups defined as front, rear, side and ceiling (see Fig. 4.12). It was intended again that the energy from the different systems is equalized. Calibration was started from the basic system setup where all loudspeakers were brought to the same level and delayed for the closest audience group. Then the levels between the four loudspeaker groups at the chosen receiver position were adjusted with pink noise. This was further verified by

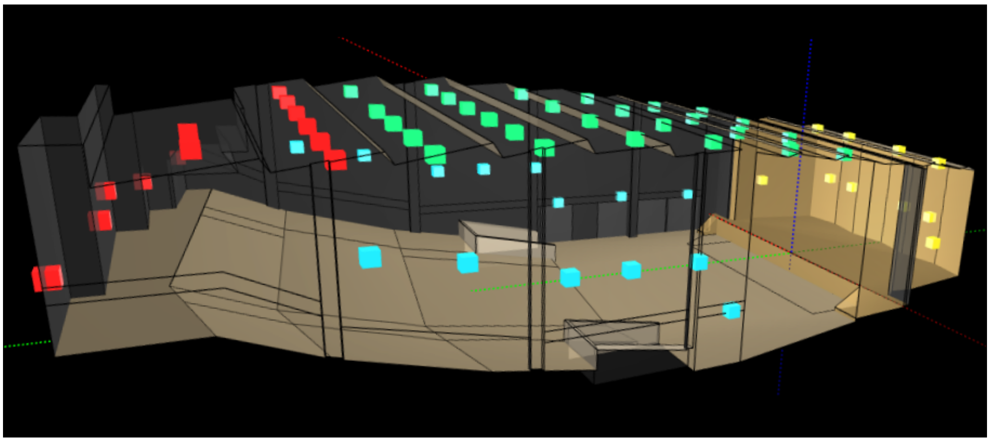


Figure 4.12. 3D-Model of the Konzerthaus with loudspeakers and directional late reverb groups (colored): "front" speakers (yellow), "rear" speakers (red), "side" speakers (blue) and "ceiling" speakers (green). The stage is on the right, the speaker gap on the side wall is due to the organ.

matching the early-to-late energy ratio (clarity C_{80}) as close as possible between the four groups by further manipulating the level and spectrum of each reverberation. Only late reverberation, i.e. energy after 80 ms was added.

Playback was done with a RME Madiface XT with analog outputs to the direct sound speaker and digital Madi output through Vivace to the reverberation loudspeakers. After the test design was finalized, the 64 channels of vivace output were recorded digitally for each stimulus and stored as 64 ch - wave files. An existing MaxMSP paired comparison patch (Aalto Virtual Acoustics Group) was reprogrammed to accommodate 64 ch-files/outputs and allow for smooth playback and switching between those stimuli.

4.4.1.1 Experiment

A paired-comparison design was used, the graphical user interface was shown in the methods Section (Fig. 2.10). The participant rated one pair of stimuli at a time, deciding which one offers more envelopment. When a decision is made the stimulus must be chosen to proceed, however the judgment "the same/no difference" could also be given and noted on a separate paper. The audio files were playing on an endless loop and could be switched seamlessly facilitating the A/B comparison. The order was randomized for every participant, each stimulus was rated twice resulting in 10 x 2 = 20 comparisons overall.

A short training session of 5 min was done with every participant to ensure proper understanding, followed by a short discussion for open questions. The stimuli for training also appeared in the test but in a different order. 12 participants took part in the listening test (10 male, 2 female, average age 27 y), 8 of which were audio engineering students at least 2 years into their studies and could thus be considered expert listeners. The duration was on average 11.5 min. The listening level did not exceed L_{Aeq} = 68 dB (or L_{CFmax} = 75 dB), background noise level was measured at L_{Aeq} = 20.5 dB.

4.4.2 Results

4.4.2.1 Listening Test

The results can be seen in Fig. 4.13. The probability for a stimulus to provide *maximum* envelopment from the given group of stimuli is shown. The condition "All" refers to all four directions being played back at the same time and is thus, as expected, rated the highest ($p<0.005$ after Bonferroni correction). For the individual groups "side" reverberation is again the most effective, not significantly different however from "rear" reverberation ($p=0.43$), followed by "ceiling" and "front" late energy. When comparing the individual magnitudes, reverberation from behind is approximately 80% as effective as the "side reverberation" in terms of LEV, "ceiling" ca. 35%.

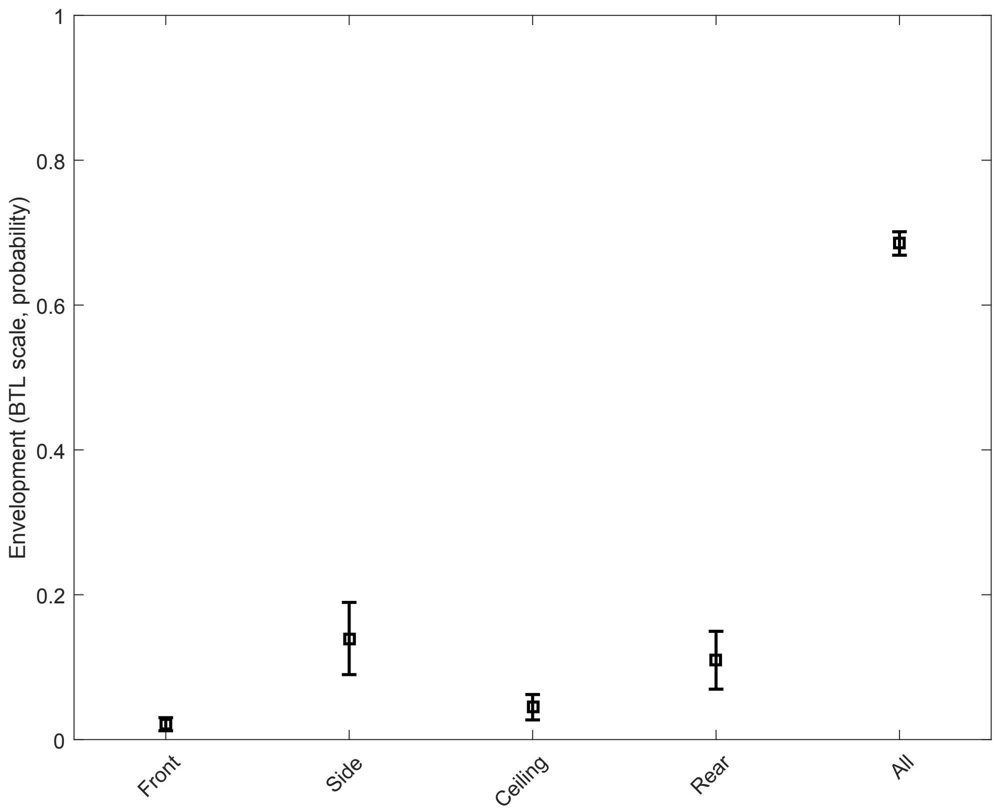


Figure 4.13. Envelopment for different directions of added artificial late reverberation. Saxophone stimulus with n=12 participants, error bars represent ±1 SE.

4.4.2.2 Measurements

In the setup process, the loudspeaker groups were first matched using pink noise and afterwards with IR measurements to equalize the ratio of early-to-late energy. This adjustment of clarity C_{80} over frequency is shown in Fig. 4.14. Similarly as in the previous experiment the matching worked well overall, only the energy below 63 Hz of the side group stands out. Likewise, levels of the stimuli conducted and shown in Table 4.2 are fairly similar, with some deviations, "rear" is 1 dB louder than desired.

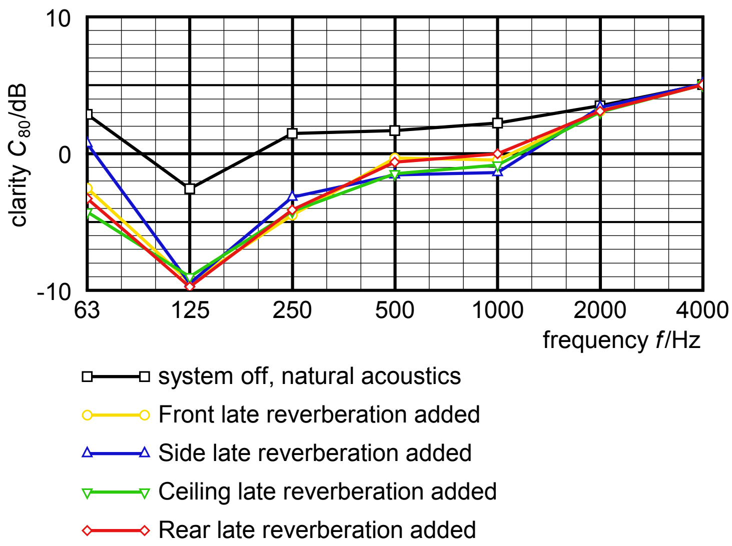


Figure 4.14. Measured Clarity C_{80} for the different late energy directions/stimuli.

Table 4.2. Level measurements for the five conditions.

Condition/Level	L_{Aeq} [dB]	L_{CFmax} [dB]
front only	59.4	73.3
side only	59.6	72.6
ceiling only	59.8	73.0
rear only	60.7	73.8
all	62.3	75.0

Directional measurements were done with a 6-ch-microphone probe and evaluated with the Spatial Decomposition Method (SDM, see Section 2.3.3), with modifications by S. Amengual and the author. This alternative visualization method is used here for clearer level analysis and will also be used in the next section. The visualization is limited to the energy after 80 ms (Fig. 4.15). The natural late reverberation (black) is overlaid with the artificial reverberation (in red), averaged over angles of 15 degrees without additional spatial weighting. It shows nicely how the late energy is deformed in one or the other direction. However, the additional energy is not sharply separated because of the spatial spread of each loudspeaker group, and the created higher order reflections from the artificial reverberation. This factor is the most pronounced for the frontal condition (bottom, right) were the added energy from the front is likely also reflected from the back wall. It can also be observed that the original late reverberation (black) is by default the strongest from the front, and weaker for the side and back of the hall.

Late lateral Level L_{J} was computed from the measurements (Fig. 4.16). It can be seen that the figure of eight directivity with cancellation to the back underestimates the contribution of rear reverberation when compared to perceptual results.

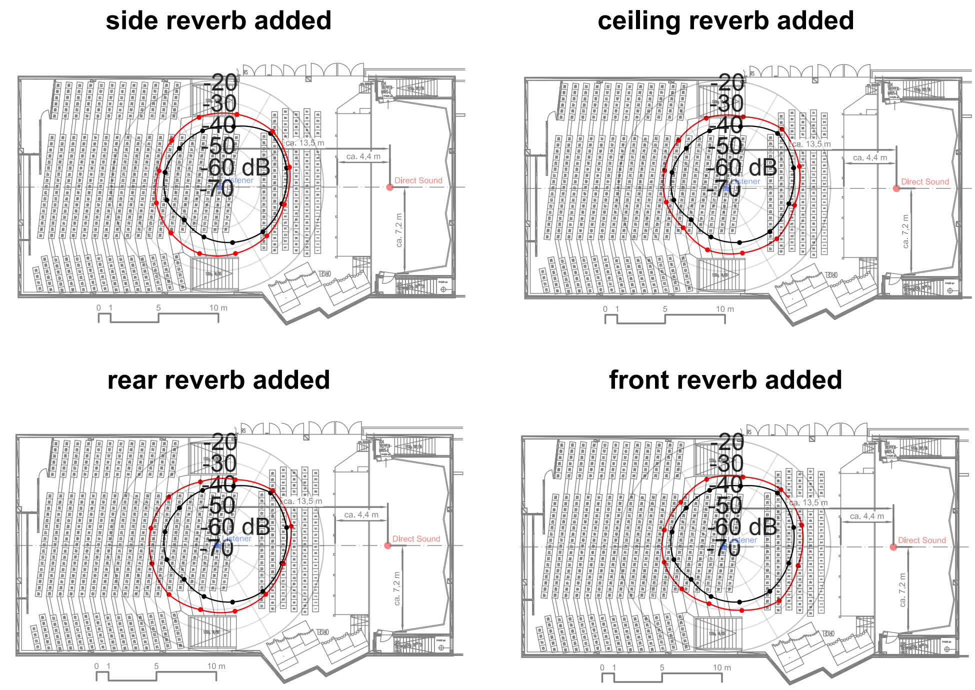


Figure 4.15. Ground view / lateral plane of Konzerthaus Detmold with the configurations for late reverberation (after 80 ms). Black denotes the natural late energy, the stimuli are shown in red with added reverberation.

4.4.3 Discussion

After the experiment 10 of the 12 participants commented with a general opinion on the test, difficulty of the task and the sound. No special difficulties were reported and the A/B-comparison was found to facilitate the task. Two of the less experienced participants would have preferred a shorter stimulus or longer notes for facilitating the comparison. One would have preferred the stimulus to start from 0 seconds instead of looped parallel audio. If described, the sound field was judged as nice and realistic (4 times), except when too much reverberation was present, (likely the setting “All”, described as “spilling over”, mentioned twice). The room volume and room “information” was said to increase without introducing distance to the source. The stimuli range was judged to be somewhat unequal by the four less experienced participants. Some pairs were quite different and thus hard to decide, others not much, whereas the audio engineering students did not report difficulties. Spectral effects or differences in frequency were mentioned by 5 participants. One participant remarked that there was coloration due to too much low-mid energy leading to a “narrow” sound. Another person found the low frequencies to help envelopment. Loudness differences were only mentioned twice and only for the setting with “all directions” suggesting that loudness might not have been a decisive criterion.

A reasonable number of 16 % of the comparisons were judged as “no difference/the same”. Moreover, more envelopment does not necessarily seem to be better. Even

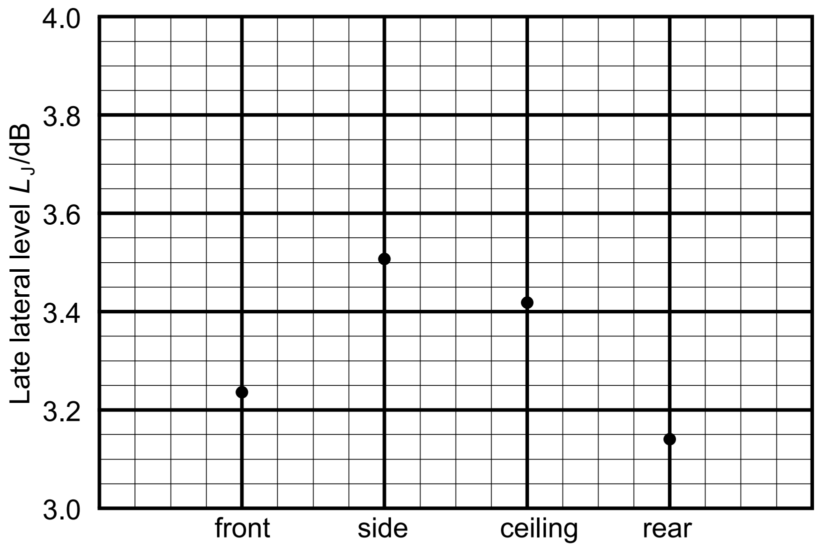


Figure 4.16. Late lateral level (L_J, 125-4000 Hz) for the modified directional sound fields, e.g. condition "front" with electronically added late reverberation from the front. Strength reference calculated from room volume and measured reverberation time.

though listener envelopment was defined orally before the test (feeling of surrounding reverberation etc.), around 7 participants mentioned ambiguities: when there was more reverberation audible for example, likely the "all directions" stimulus, the sound was not liked more because the naturalness, transparency and *immersion* suffered. After further discussion the LEV or surrounding property was emphasized to be decisive for the judgment and was clear from then on. However, it was noted down several times that the less enveloping stimulus was preferred for instance for more clarity or naturalness.

As in the previous test the different directions of late energy as such were not detected too much. 5-6 participants pointed out rear reverberation to be different and mostly liking it as it would "leave room for the source". Other directions were not named specifically, except two participants noticed imbalanced side reverberation (more left than right) possibly due to the lack of speakers on the right side with the organ.

4.4.4 Conclusion

Different directions of added artificial reverberation were compared for a musical stimulus in a chamber hall. The reverberation level was found to be overall dominant. Results for the first stimulus suggest again that lateral late energy is most effective for LEV and reverberation from the back of the hall is less, but also, important.

4.5 Large venues and comparison between rectangular and vineyard design: reproduced Berlin halls

The previous two experiments manipulated the sound field inside a real environment. This experiment judges LEV from the real sound field for two well-known concert halls of different design. The late part of the sound field of two measured concert halls Berlin Konzerthaus and Berlin Philharmonie, auralized with a state-of-the-art reproduction method, is altered virtually regarding the direction and level: reverberation is excluded instead of increased for certain room directions (see also [96]).

4.5.1 Setup

Two concert halls of international reputation have been chosen for comparison in this study: The Berlin Konzerthaus (or Schauspielhaus am Gendarmenmarkt, abbreviated "BK") and the Berlin Philharmonie ("BP"). Even though only two halls, they are both critically acclaimed (Top 15 in an often referenced rank list [19]) and present two very distinct types of concert halls: a classical shoebox and a vineyard-style surround hall. Due to the architecture they differ noticeably in numbers: the Konzerthaus seats 1,600 people whereas Philharmonie can offer space for up to 2,200 people. The cubic volume of BK is given with an average of 15,000 m^3 whereas the Philharmonie's volume is 21,000 m^3. However, both halls are being used for large orchestra concerts with the main classical repertoire and therefore can be judged equally.

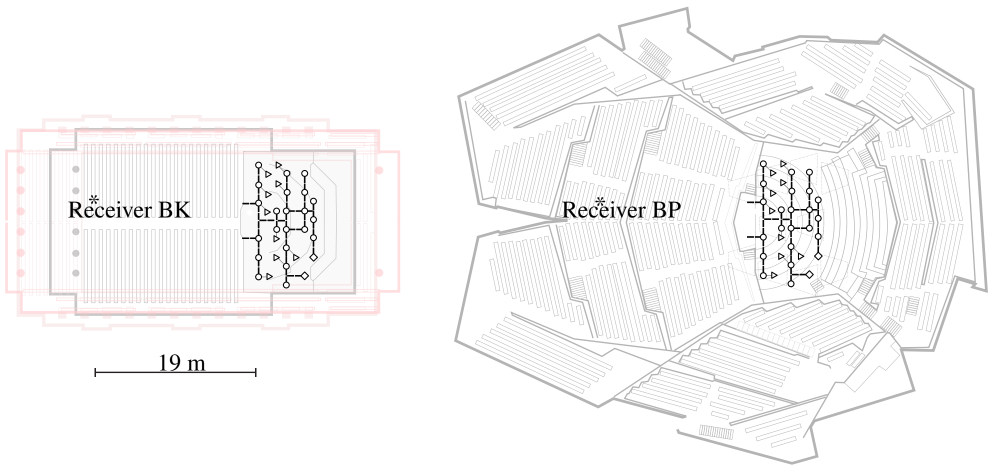


Figure 4.17. Groundplan of the Berlin Konzerthaus (left) and Philharmonie (right) with loudspeaker orchestra on stage.

The auralisation and measurement material stems from the measurement tour of the Aalto Virtual Acoustics Group through various European concert halls in 2012 which resulted in a number of publications, a recent overview can be found in [113]. The sound source employed during this tour is the so-called loudspeaker orchestra, an arrangement of 34 directional loudspeakers that are placed in a way to represent the real orchestra, see [24]. The loudspeakers on stage are arranged identically in every hall, likewise with the receiver distances: a constant source and receiver situation where only the acoustics of the particular hall is differing. Impulse responses from every loudspeaker on stage were then recorded with a directional measurement system (a six-channel G.R.A.S. Intensity-probe VI50) for several receiver positions.

In each of the two halls, one seat was chosen that is located at a fixed distance from the first row of instruments by about 19 m and along the middle aisle to the left. The seating position can be seen in Fig. 4.17. The seat was also chosen for the possibility of comparing to previous experiments conducted by Haapaniemi et al. [66] as well as a representative distance well in the reverberant sound field.

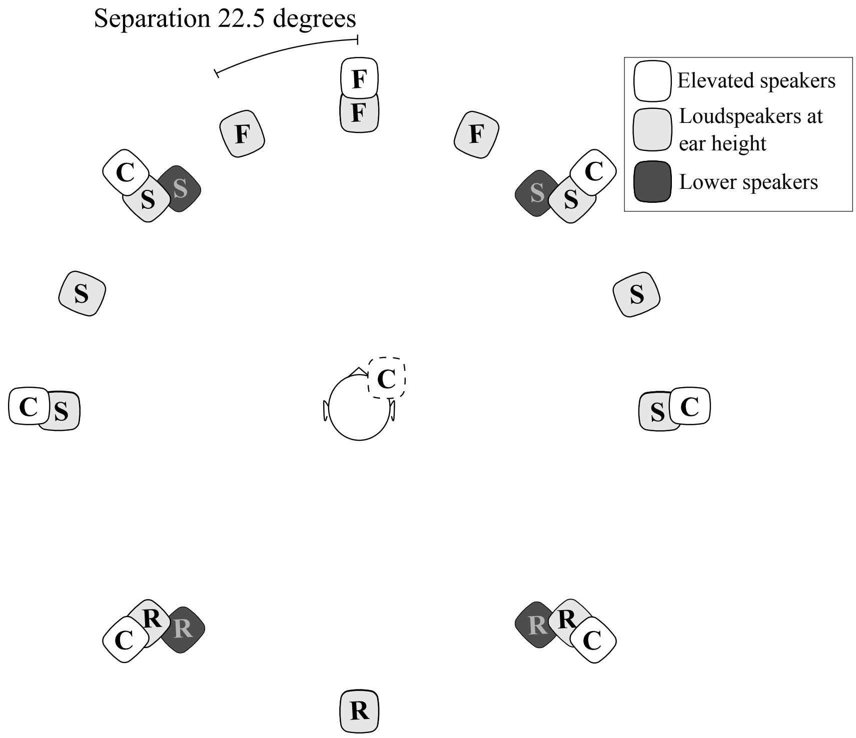


Figure 4.18. Reproduction setup with 24 speakers at three heights, the late reverberation of the IRs is distributed to speaker groups **Front (F), Side (S), Ceiling (C), Rear (R)**.

In the laboratory, the measured impulse responses are convolved with anechoic orchestra samples and played back through a 24-ch-loudspeaker reproduction system. The convolved signal for each acoustic reflection in the real room is put into the nearest available loudspeaker (NLP from SDM data, see [23] and [25]). The reproduction system offers increased spatial accuracy, especially in the frontal hemisphere, to accurately play back early reflections while covering the whole sphere around the listener by providing speakers in every dimension.

The intention here was to study the direction of late sound, therefore the sound field must be controllable in time windows and distributed to different speaker groups. After de-composing the sound field from the concert halls to the laboratory setup, the 24 impulse responses, one for each speaker, were separated in early sound (0...80 ms) and (late) reverberation (80 ms...infinity). In the next step, only the late part was muted for certain channels. Direct sound and early reflections were played back *as is* in the original IRs whereas the late part would come only from chosen loudspeakers. The late energy was divided up so that the *front late reverb* was played back by 4 speakers at 0°, ± 22.5° and a frontal elevated speaker. The *side reverb* was coming from a total of 8 loudspeakers at ± 45°, ± 67.5° and ± 90°, as well as lower speakers at ± 45°. *Ceiling reverb* speakers were all above ear level at ± 45°, ± 90°, ± 135°, as well as immediately above the listener (7 speakers). Lastly, *rear reverb* speakers were 5 speakers at ear level or below (± 135° on ear level/below and at -180° on ear level), see Figure 4.18.

It can be seen that the number of speakers is not equal between conditions. A lot of consideration went into how to distribute the speakers especially after identifying this as a possible shortcoming in previous experiments. However, the chosen distribution was felt to be the closest to what a listener would consider to be a certain direction when sitting in that concert hall seat. A speaker at 45° azimuth and 45° elevation could be interpreted as both ceiling and side (ceiling chosen) or a speaker at 0° azimuth and 45° elevation as both frontal and ceiling (frontal chosen). It was also decided, not to normalize the energy between the four different direction groups as the real sound field conditions were to be investigated, in particular two real concert halls *as they are* opposed to an entirely equalized synthesized sound field.

An excerpt of Beethoven's Symphony No. 7 was used as an audio example, a good representative of classical to romantic repertoire (duration 22 s, bars 23 to 30 of the 1st movement) with mainly woodwind and strings playing a rather calm, slow to mid-tempo mezzo-piano. The excerpt was played repeatedly. Mostly running reverberation could be heard and a decay from the final stop chord of the excerpt. Playback was done with a MacBook Pro from Apple connected to a MOTU 16A sound card converting to analog outputs for 24 loudspeakers. The participants were controlling the experiment via an iPad that mirrored the screen of a MacBook and added a touch pad functionality. Two different experiments were conducted.

Both experiments 1 and 2 were interchanged and randomized in the order of stimuli as well. Each had a training session of roughly 5 minutes with a break for discussing open questions and the methodology. Informed consent was collected from each participant. Only members of the Aalto Virtual Acoustics Group and the Department of Signal Processing and Acoustics took part in the experiment, mostly informally trained skilled listeners (10 participants, average age 31 years, all male). Average duration for Experiment 1 was 16 min and 13 min for Exp. 2 summing up to 29 minutes for both tests. Listening Levels L_{Aeq} ranged between 59 and 61 dB (see below). In Exp. 2, when turning up the fader fully, a maximum L_{Aeq} of 68-70.5 dB or L_{AFmax} of 78 dB was possible. The listening room follows requirements of ITU-R BS.1116-1 for reverberation time (0.1 s at mid-frequencies) and a level difference greater than 10 dB between direct sound and room reflections. Loudspeakers were calibrated to 0.1 dB, different loudspeaker delays accounted for and the background noise was measured at $L_{\mathrm{Aeq}} = 31$ dB.

4.5.2 Experiment 1: paired comparison

Four altered sound fields and the original situation were presented for both seats. This results in a number of 10 stimuli overall which were all compared against each other, a total of 45 comparisons. The participants were asked if "stimulus A or B is more enveloping". The change between stimuli was possible instantly (files playing in parallel on endless loop). The choice was not forced, namely stimuli could be rated to be identical regarding envelopment (encoded as equal weight to both stimuli). After initial testing it became clear that eliminating the late sound of all room direction groups but one sounded too artificial as there would be an unrealistically strong contrast between different cases. Therefore, the opposite was chosen and the reverberation was removed from one certain direction group, specifically, one stimulus had no frontal late reverberation, the other had no side late reverberation etc. A somewhat similar effect occurs in a real hall with an absorptive ceiling or back wall.

4.5.2.1 Results

Figure 4.19 shows the probability for a certain stimulus to provide the most envelopment among the group of stimuli. Firstly, it can be seen that all Berlin Konzerthaus stimuli (BK) are judged to be more enveloping than Berlin Philharmonie examples (BP). Secondly, the order of envelopment can be read as follows: the most envelopment is offered as expected by the original sound field of Berlin Konzerthaus with late reverberation from all directions. Then follows the sound field without late energy from the front of the hall (BK-no front reverberation). Yet if either the rear, side, or the ceiling reverberation group is excluded it has a significant effect on the feeling of being enveloped ($p<0.05$). The five stimuli of Berlin Philharmonie being so close to zero is a matter of a low probability compared to BK. Scaling up these five stimuli (Fig. 4.20) it can be seen that there are few differences: again, the original reverberation situation appears to have the highest envelopment followed by excluded front and rear reverberation groups. However, the ceiling reverberation group seems to have a more important role as the envelopment goes down when it is excluded (borderline significant, $p = 0.08$).

The results can be interpreted as follows: Berlin Konzerthaus has more envelopment overall. Removing the front reverberation group has no important perceptual effect on the envelopment in BK. All other three directions of reverberation are noticed. Interestingly the side and ceiling reverberation are of equal importance in this seat. In Berlin Philharmonie, the only likely noticeable difference appears if the late ceiling energy is excluded, in other words: front, rear, and side groups have no significant influence on overall LEV and most audible late reverberation originates from above.

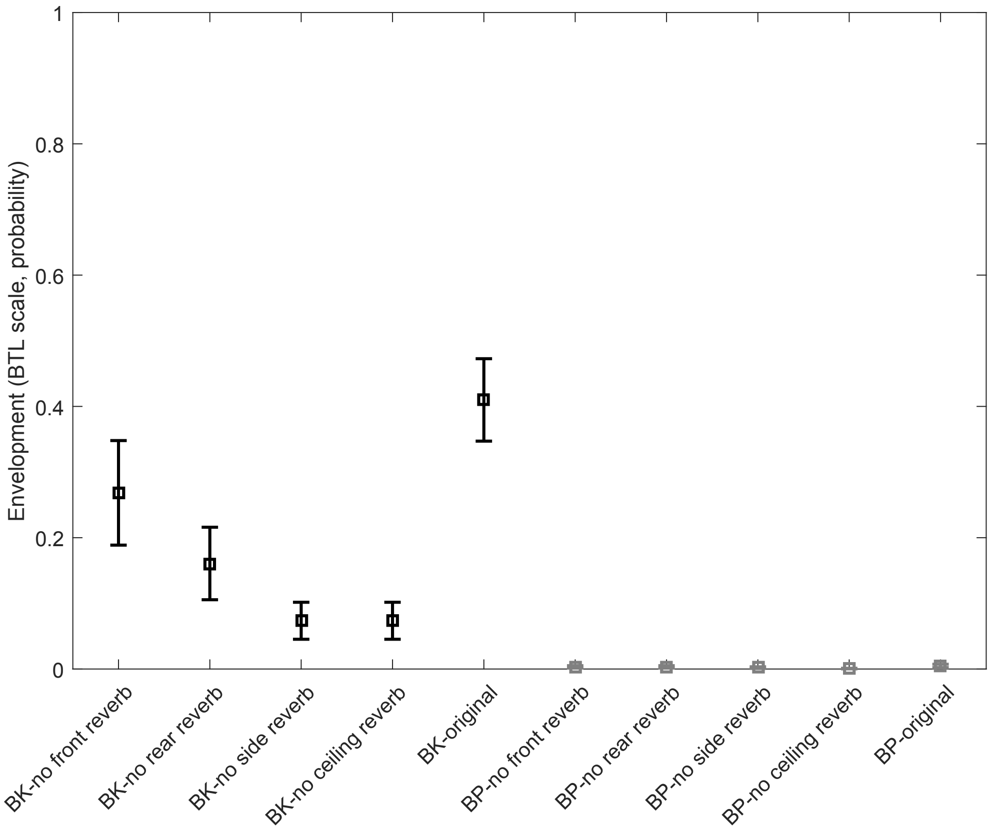


Figure 4.19. Envelopment when excluding late reverberation from different directions for Konzerthaus (black) and Philharmonie (grey). Error bars show ±1SE.

Table 4.3. Measured in situ stimulus levels (monaural, class 1 SPL-meter, averaged 3x per stimulus.)

Berlin Konzerthaus	Level L_{Aeq} [dB]	Berlin Philharmonie	Level L_{Aeq} [dB]
no front late reverb	60.6	no front late reverb	59.4
no rear late reverb	60.3	no rear late reverb	59.4
no side late reverb	60.1	no side late reverb	59.0
no ceiling late reverb	60.0	no ceiling late reverb	58.2
original	61.0	original	59.6

Sound levels measured in the listening room are shown in Table 4.3. It can be seen that the overall difference ranges from 1 - 1.5 dB between the two halls. As expected, the original stimuli are the loudest in each hall. Levels between BK original and rear, side and ceiling reverberation groups are within the just noticeable difference of 1 dB for G, yet significant perceptual differences were found. It can therefore be concluded

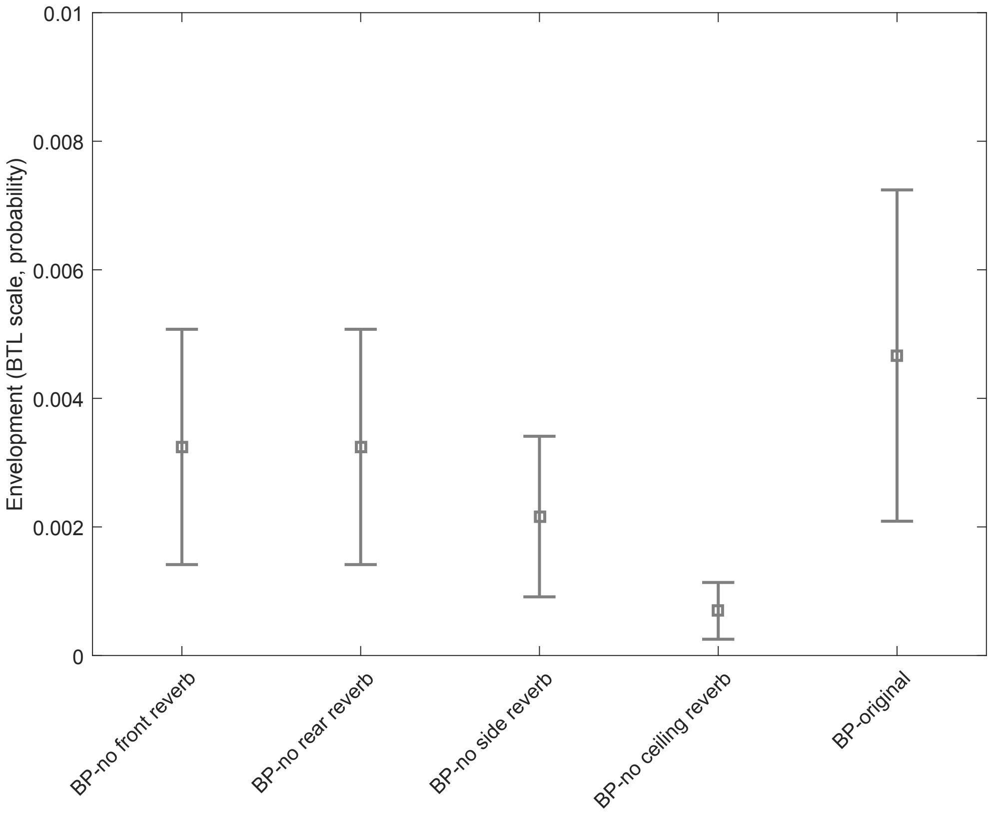


Figure 4.20. Envelopment when excluding late reverberation from different directions only for Philharmonie (grey). Error bars show ±1SE.

that the differences are due to spatial effects and not due to changes in loudness. In BP it can be seen that when excluding front, rear or side reverberation groups there is little change in the monaural level compared to 1.4 dB less for "no ceiling", suggesting that the most reverberation originates from above.

Level based acoustic parameters strength G and late lateral level L_J were calculated according to the ISO 3382-1 standard [6], averaging G over 500 Hz and 1000 Hz octave bands, and L_J over 125 Hz to 1000 Hz octave bands (energy average). Late strength G_{late} (80 ms - inf) was averaged similarly to L_J to facilitate comparison between the two parameters. Strength G, leftmost in Fig. 4.21, shows that the energy was reduced most for the both rooms without ceiling reverberation. This factor helps explain its importance in the test. Comparing these findings with the results from the listening test it becomes clear that both L_J and G_{late} do follow the trend in the results. However, an average between the two parameters would be closer to the test results in the two cases BK side/ceiling (black) and BP side/ceiling (grey). In other words L_J is underestimating and G_{late} is overestimating the influence of the ceiling reverberation group. For BK (black) it can also be seen from G_{late} that excluding the ceiling reverb group removes more late energy than excluding the side reverb group. Yet both were judged to be similar, hence late lateral/side reverberation is indeed more effective for LEV. It is interesting to note that excluding the BP-ceiling reverberation group also decreases the lateral energy, meaning there is a noticeable amount of late lateral energy in the ceiling group (see Discussion). Rank correlation values between listening test results and parameters strength, late lateral level and late strength of 0.78, 0.90 and 0.98 respectively are found, i.e. analyzing only the late reflections is more suitable than late lateral level with figure of eight directivity here. Correlation would increase further if the energy from the back were weighted more and the side energy weighted less.

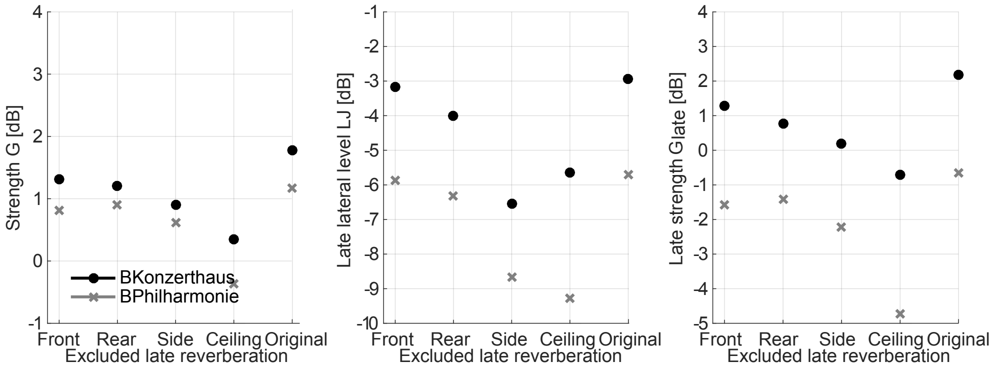


Figure 4.21. Parameters for the different conditions, Avg. 500/1000 Hz for G and 125-1000 Hz for L_J, G_{late} (Exp. 1).

4.5.3 Experiment 2: magnitude adjustment

In the second experiment listeners would actively change the sound field to their preference by controlling the late energy from particular direction groups. Participants could set a digital dial that changed the gain of one particular late reverberation direction group. For instance, the late reverberation from the sides changed while the rest of the sound field was original and constant. The user interface for this task is shown in Figure 4.22. The participant could increase or decrease the gain by moving their finger up and down on the touch pad and was only told that a component of the sound field was changing. The dial was endless to suppress visual bias (see Haapaniemi [66]) and initial gains of the variable controllable sound component were randomized between -7 to -13 dB below the original value 0 dB. The constant sound components stayed at the fixed original gain of 0 dB. Also, the maximum possible gain was limited both for safety as well as to keep the participant from getting out of a reasonable listening range. A mute button silencing the variable late component was introduced for reference and greater accuracy.

The main interest was whether certain directions of late energy would be dialed in louder, lateral or maybe directions of missing reverberation, e.g. in order to get a spatially equal reverberation. Another question was how close to the original situation (0 dB) participants would set their preferred gains. Each judgment was repeated three times resulting in 24 trials (3 * 8 stimuli).

Figure 4.22. User interface for adjustment test (Exp. 2). The blue endless dial changes the level of the variable late reverb direction, 'mute' silences it.

4.5.3.1 Results

In the magnitude adjustment test, participants adjusted one late component of the sound field blindly to preference. A Friedman Test (non-parametric ANOVA) was done in SPSS 22 showing statistically significant differences in the stimulus set ($\chi^2(7) = 55.9$, $p < 0.000$) and further post-hoc analysis with an included Dunn-Bonferroni multi-comparison. Analyzing the individually controlled reverberation direction groups it can be seen that participants set the gain higher than the original value in all cases. This result could be either due to a general taste for more reverberation or the stimulus listening level being too quiet (though the listening level is close to the real on-site level). None of the Konzerthaus stimuli are significantly different from each other, see Fig. 4.23. There is a trend that the BK front stimulus was raised more than the rear, side and ceiling reverberation groups. This finding is mostly in agreement with Exp. 1 where side and ceiling reverberation had the most prominent effect (and energy) and were here therefore not raised more. BP rear stands out and is significantly different from all stimuli ($p<0.05$) except BP side: Participants raised reverberation from the back of the hall a lot. This occurence might be due to the fact that there was little rear reverberation to start with, offering greater headroom (see discussion spatial plots and Fig. 4.26) and suggesting the need for reverberation also from behind the listener. It remains unclear if this adjustment is due to a desire for LEV, reverberance or level.

When grouping the four directions' levels, Berlin Philharmonie is made significantly more reverberant than Berlin Konzerthaus (Fig. 4.23 right, Wilcoxon Rank-Test: $Z = -4.994$, $p < 0.000$, effect size $r = -0.46$), i.e. Philharmonie is lacking late energy.

Level based acoustic parameters for both the original situation and the median values

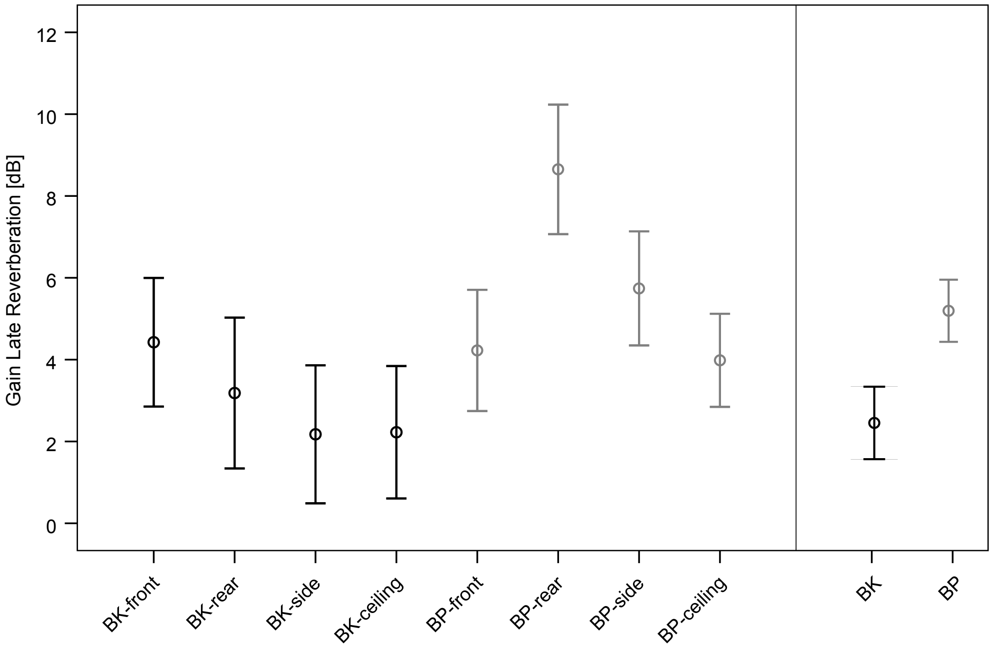


Figure 4.23. Mean and ±1 standard error of the late reverberation level in experiment 2 (magnitude adjustment test), asking for preference. Individual directions and grouped, n=10.

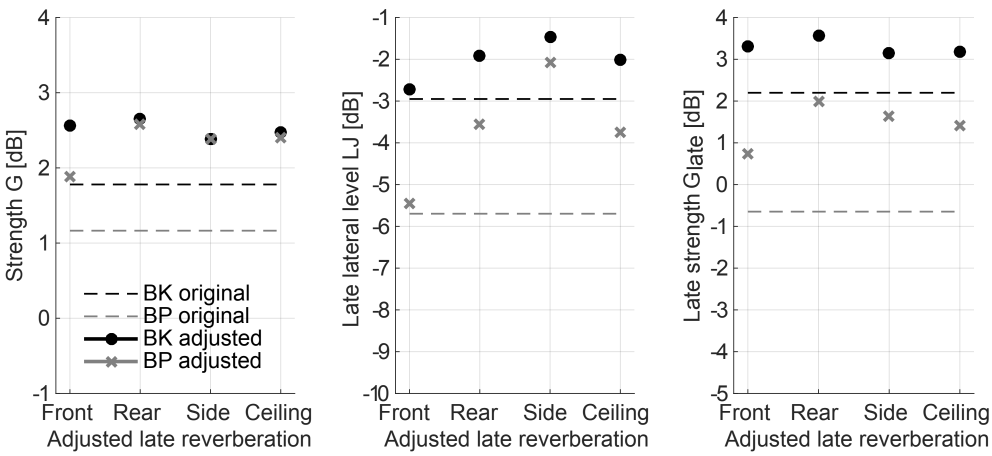


Figure 4.24. Parameters for the different conditions, Avg. 500/1000 Hz for G and 125-1000 Hz for L_J, G_{late} (Exp. 2).

set by the participants are compared to visualize how the adjustments reflect in the measures (Fig. 4.24). G and G_{late} show that the energy was raised similarly among direction groups in both halls (less for BP front): For G on average by +1.2 dB (BP) and +0.8 dB (BK), to roughly 2.5 dB (a difference of 1 dB is generally considered noticeable). Increase in late lateral strength L_J is not surprisingly the strongest for lateral/side late reverberation due to the figure of eight directivity with zero amplitude to front and back. It seems that by adjusting the late reverberation listeners increased LEV, the reverberance and/or strength on average, excluding BP front, possibly due to coloration or a loss of clarity.

Spatial plots show the directional distribution of sound energy in the lateral and median planes for the different adjustment cases (Figs. 4.25 and 4.26). Included are the

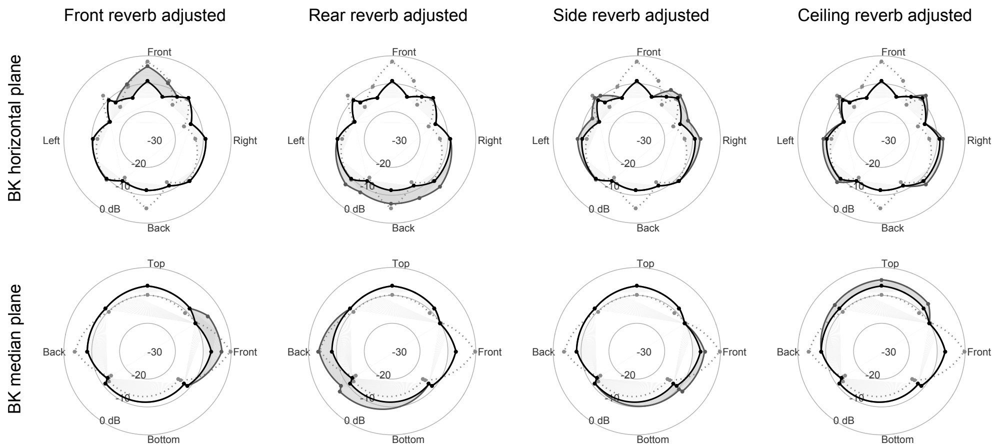


Figure 4.25. Spatial plots for Berlin Konzerthaus. Late energy (80 ms - inf) for the original situation in solid black, participants' median adjustments are shaded grey. Dotted grey lines show early energy (0-80 ms).

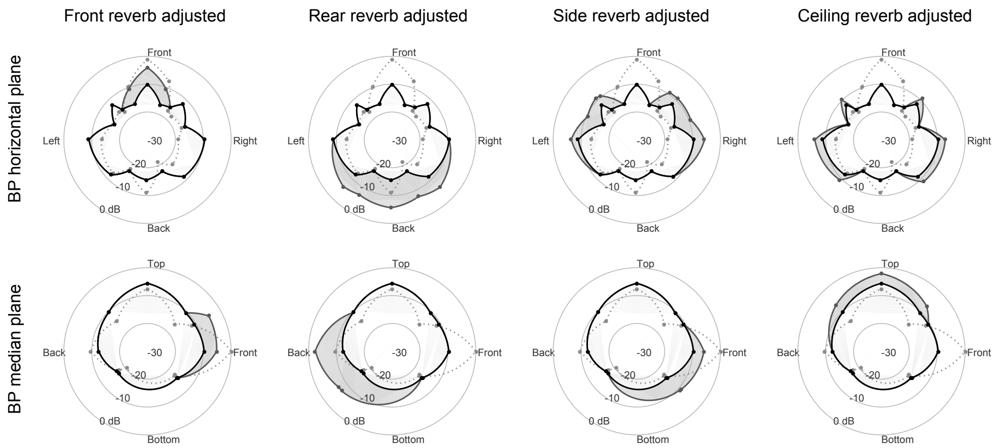


Figure 4.26. Spatial plots for Berlin Philharmonie. Late energy (80 ms - inf) for the original situation in solid black, participants' median adjustments are shaded grey. Dotted grey lines show early energy (0-80 ms).

original late reverberation (black dots/solid line), the adjustment based on the median result (gray dots/solid line), and the distribution of direct sound and early reflections (gray dots/dotted line). The dots show the actual sound energy coming from a direction corresponding to the reproduction system, and the lines interpolate between the dots for visual assistance. A toroidal weighting function was used for calculating the direction distributions, particularly a figure of eight rotated around its axis, so that for instance, for the lateral plane the top direction has zero weight. It can be seen that the original late reverb distribution both in BK and even more so in BP is somewhat ellipsoid shaped. In BP, there is noticeably less late energy from the rear group than in BK. Also, within BP there is less rear than side reverberation. In both halls the ellipsoid is adjusted to be somewhat more round.

4.5.4 Discussion

In this study the direction of late reverb of real concert hall material is manipulated within a more advanced sound field. The widely accepted notion of late side/lateral reverberation being more important is put into perspective. It depends on the sound field presented: late side energy *is* important in this test for Berlin Konzerthaus, a typical shoebox hall, but the reverberation from above seems equally important (Furuya [105] reported 35 % importance compared to side reverb). In Berlin Philharmonie, however, there is overall less energy in the side and rear reverberation group so the ceiling group becomes much more important. This emphasis is probably due to the fact that surround halls are prone to having less late rear and side energy because of the raked seating. Likely, the distribution of speakers influenced the results, but there are several ambiguities on how to assign a speaker (or reflective surface). Is sound from elevation 35° but 90° azimuth to be considered side or ceiling? Distinguishing clearly is not trivial and was done based on common sense and sound-mixing practice. Also, there was a good amount of lateral energy in the ceiling reverb group of both BK and BP (Fig. 4.21). The question arises how elevated lateral reflections are interpreted in general.

The remaining level differences between the stimuli might be influencing the results and therefore bias the listener. Yet equalizing loudness would eliminate the real balance in these two halls. Similar problems existed in several other laboratory envelopment tests where the amount of speakers was not the same per directional group. For a purely synthetic diffuse sound field approach the loudness matching can be done and is kept for further study.

Only a few participants noticed what was being manipulated and this mainly in Exp. 2 where it was possible to increase the level of the speakers. The changes in direction themselves are not overly obvious when listening. Also, audience members more often face forward towards the stage. In that listening situation sounds from frontal, above and behind are not easy to distinguish. It becomes much more apparent when the head is turned.

The first experiment was overall considered manageable, however Exp. 2 received some criticism due to difficulties both in the procedure (adjusting the level on the touch pad) and the fact that multiple things or cues changed, making this task harder. The bigger variation in Exp. 2 is due to this, particularly the participants not being used to the dial as a controller (mixing engineers might have performed differently) and asking for *preference*. Comments pointed towards the difference and multidimensionality of reverberance and envelopment. Some stimuli could be adjusted to a good reverberance that fits the musical flow but no envelopment was obtained and vice versa.

4.5.5 Conclusion

The spatial distribution of late reverberation plays a vital role in the concert hall: for shoebox halls reverberation from the side is important, confirming findings from previous literature. However, both in this shoebox as well as the vineyard hall late energy from the ceiling and to some extent the rear had a noticeable contribution to listener envelopment. This is not predicted fully by the figure-of-eight weighting in late lateral level L_J.

4.6 Influence of different seat backrest height on directional late energy

All previous experiments studying envelopment assumed an unobstructed sound field around the listener's head. However, in actual venues, inclining seating often requires higher backrests for a noticeable number of seats which might alter the sound field close to the listener's head. A shielding of reflections from behind the listeners could appear and at the same time an amplification due to additional reflections off the higher backrest bringing more energy to the listener. The present study quantifies some possible effects using virtual acoustics and directional measurement techniques.

4.6.1 Setup

The chair under investigation is intended for a medium-sized concert hall in Germany with a number of 1000 seats in total. One fifth of these, 200 seats, are installed with raised backrests for safety reasons. Normally, the acoustic consultant would prefer a construction that is partly sound transparent (e.g. perforated) but this was not possible here. Therefore, a raised backrest with the fabric glued onto the wood directly was chosen.

Figure 4.27. Chair with increased backrest (left) and standard height (right).

To quantify the level decrease due to the backrest as well as the effect in a diffuse sound field the three versions have been measured in an almost anechoic testing room equipped with enhancement system (inner dimensions of 4.8 x 3.4 x 2.2 m, room volume 36 m^3). The present reverberation time T_{30} is below 0.05 s at 250-4000 Hz reaching 0.15 s below 125 Hz. 16 Genelec 8130a loudspeakers (frequency response +/- 2 dB between 58 Hz and 2 kHz) are arranged on a semi-sphere at two heights around the listener and are used for creating the synthesized sound field with the room enhancement system Vivace, see Section 2.1. Loudspeaker signals are distributed and processed to maximize de-correlation. The preset “Concert hall 2-4” was used with an additional delay of 80 ms between direct sound and reverberation onset to clearly separate late energy (resulting T_{30} of 2.0 s at 500/1000 Hz). The speakers received a digital AES signal from a RME HDSP Madi soundcard through Nexus D/D-converters after processing in the Vivace system.

The chair was placed in the center of the room at a distance of 2.2 m from the front speaker. All speakers were calibrated to have equal SPL at this position and were delayed digitally to the same distance. For the attenuation measurement we chose the front speaker at 0 ° and a rear speaker at 180 ° in height of the head. For the diffuse sound field all 16 speakers were used. Monaural and binaural (Cortex dummy head) measurements were conducted for each of the three chair versions as well as directional array measurements, introduced in Section 2.3.3.

4.6.2 Results

Fig. 4.28 compares the sound from the frontal loudspeaker against the rear loudspeaker for the two versions. The differences between sound from the front and back is on average 0 dB for the normal backrest (blue). For the high backrest there is an increase of frontal sound energy visible but, as expected, also a noticeable decrease for sound from behind resulting in a front/back difference of 5-13 dB.

The front/back test is likely oversimplified in comparison to typical sound fields. Yet, similar effects can be observed when analyzing the situation in the synthesized diffuse sound field with direct sound from the front (Fig. 4.29): the introduction of the high backrest as a surface close to the head increases the overall energy by roughly 2 dB for the partly reflective (an ideally reflective surface would yield 3 dB). At the same time, late lateral level L_J hints that the spatial distribution is changed (Fig. 4.30). A decrease for the high backrest (red) can be seen compared to the normal condition (blue). This influence will likely result in decreased listener envelopment (LEV) as lateral energy was shown to be more effective than frontal or rear reverberation.

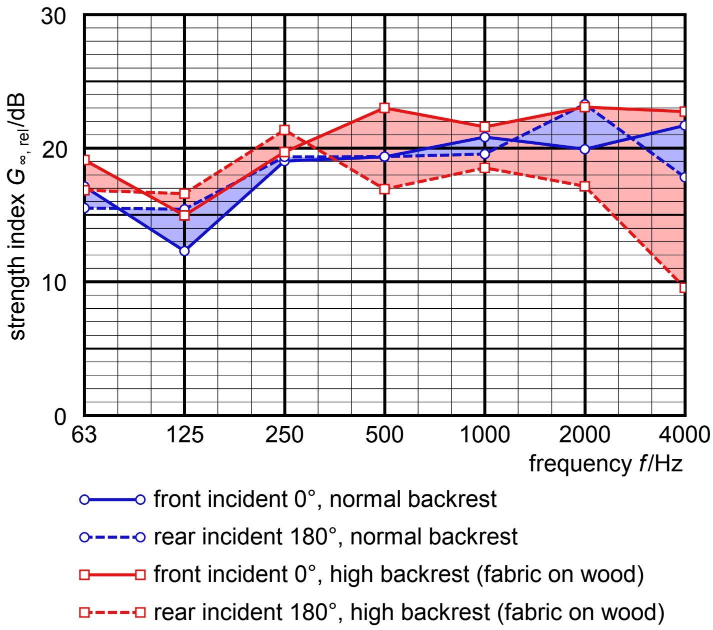


Figure 4.28. Strength G measured with the two chair versions and frontal and rear sound (binaural measurement, results from left ear only). Virtual room, thus arbitrary gain/strength values.

4.6.3 Discussion

The results are mostly as expected. The front/back comparison and the diffuse sound field show the shielding of the high backrest nicely. The effect for the free standing chair might be somewhat overestimated as in reality a seat is always within a block of other seats or audience members also influencing the sound field. The author is not aware of a comparable study for reference where the influence of the backrest height has been investigated.

Some periodicity hinting at comb filters was seen in the FFT spectrum between 2500 and 5000 Hz for the frontal pure tone/ high backrest. In the diffuse sound field this was not noticeable in measurements or during listening. More diffraction than in reality might have appeared because of the measurement position being further from the backrest than for a human receiver, see Fig. 4.31. No in situ listening test was conducted for LEV comparison, also due to practical constraints such as switching the seat rest. During informal music listening the higher backrest appeared to be noticeably less enveloping.

4.6.4 Conclusion

Higher backrests, often required in inclining seating areas, alter the sound field close to the listener considerably. This finding shows a level increase from frontal directions but noticeable shadowing effects to the rear and side, likely leading to reduced listener envelopment.

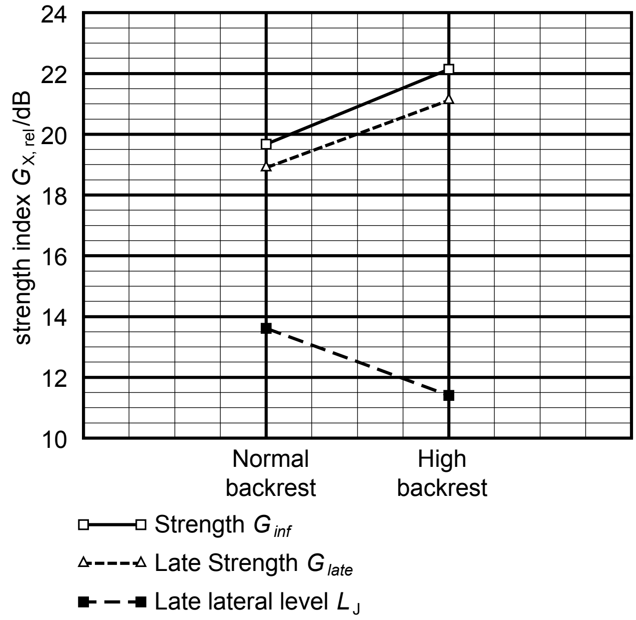


Figure 4.29. Strength indexes for the two chair versions in the synthesized diffuse sound field (Averaged 500/1000 Hz for G values and as well as for L_J for comparability). Virtual room, thus arbitrary gain/strength values.

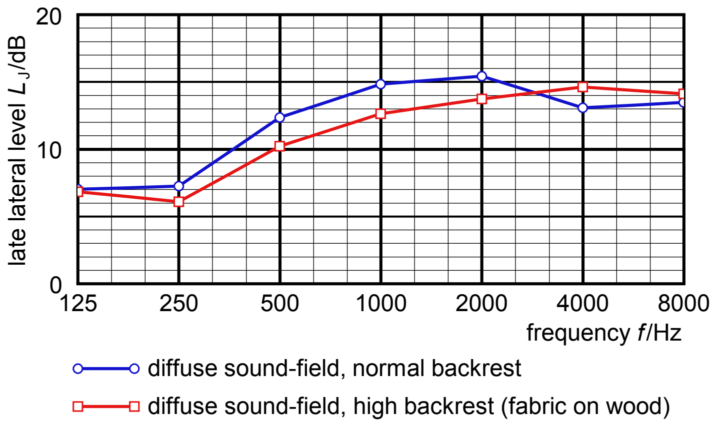


Figure 4.30. Late lateral Level L_J for the two chair versions in the synthesized diffuse sound field. Virtual room, thus arbitrary gain/strength values.

Figure 4.31. Measurement setup with increased backrest for array-(left) and dummy-head measurement (right).

4.7 Discussion and conclusion

Typically, previous experiments regarding listener envelopment were conducted in somewhat simplified sound fields. In this work, the influences of directional late energy were investigated in more realistic settings. The direction of reverberation was varied in two enhanced rooms of different sizes, a lecture hall and a medium-sized concert hall, as well as a comparison of two large venues (reproduced Berlin halls).

The results are in line among experiments and partly with literature: overall, the reverberation level was found to be dominant. If the reverberation level was constant, then side reverberation was the most important among the directional groups. However, the reverberation from the rear could also contribute to envelopment (LEV), and also but less so ceiling and frontal reverberation. These results together with previous evidence from literature should outweigh the currently standardized notion that only late lateral energy is decisive for envelopment, as the parameter in ISO 3382-1 [6] connected to LEV only measures lateral late energy.

Interaural cross-correlation (IACC) was not sufficient for predicting LEV, as it does not react to level differences. This lack is detrimental especially because the reverberation level was shown to be dominant for LEV. Also, as a result of this it does not seem to be a quality criterion for suitable reverberation to have a low (late) IACC. The outcome is in line with the current notion that IACC is not conclusive as a predictor for envelopment.

The two large halls were famous and well received representatives of the concepts rectangular vs. vineyard hall design. The results suggest that halls with greater volume and sloping audience are worse for envelopment because of the overall lower reverberation level and lower late side and rear energy. The implications for concert hall design are, therefore, to increase or at least maintain reverberation level to ensure good envelopment. If the location can be influenced, reverberation volumes should thus be at the side or behind from the listener as the other locations are less effective. Also for room enhancement systems some consequences can be derived: audience coverage with late energy has to be maximized. Side reverberation is more effective and could thus, followed by rear loudspeaker groups, receive more loudspeakers.

Lastly, some evidence was presented in one of the three experimental environments that the direction of late energy did not affect the clarity of sound. However, some more testing would be required.

5. Levels and dynamics of music in reverberation

5.1 Introduction

In chapter 3, the influence of different levels of reverberation on reverberance were studied. However, it was not discussed in that chapter how this alters the music signal at the receiver's ear. Lately, there has been some evidence that *dynamics* is also relevant in concert hall acoustics. For instance, that early reflections provided by the hall alter the dynamics of the music signal [114]. In the following chapter this is discussed further and extended to reverberation.

5.1.1 Definitions of level and dynamics

Level describes the absolute magnitude whereas dynamics refers to the range of values. Furthermore, one could differentiate between the physical dynamics (e.g. the range of SPL values) and perceived dynamics. Also, the meanings vary depending on the context.

In **music**, dynamics describes the concept of different steps of volume [115, p. 137] and is thus one of the main musical features next to melody, harmony and tempo. Changes between quiet and loud parts (crescendos) or sudden stops are used widely in most types of classical music to transport meaning and emphasize emotion. The dynamics notations such as *piano* or *forte* symbolize the steps of sound volume. The dynamics notations are per se relative descriptors, thus ultimately their translation to levels depends on the instruments playing, the piece and also the venue (see below). Notation of dynamics symbols found its way into classic music score notation around 1600 and was well established and used more regularly with the typical range of *pianissimo (pp)* to *fortissimo (ff)* only in Beethoven's time [116, p. 281].

In **music recording**, dynamics would be defined as the range between maximum and minimum levels (sometimes mean instead of minimum). The level itself is variable as every customer listens at another volume. Dynamics is equally important and discussed controversially in modern music recording. With increasing signal compression and peak limiting music production/broadcasting is supposedly suffering from a "loudness war" or over-compression. Even though there are no standardized definitions, some additional differentiation exists: the term macro-dynamics is used to describe longer parts carrying musical meaning such as crescendos or a verse/chorus-progression whereas micro-dynamics considers the dynamic range of short events or transients (e.g. single drum hits), apparent with short music excerpts of only some seconds. When the macro-dynamics, the dynamic progress over the course of time is altered, for example with compression, a previously quiet verse can become louder than the previously dominant chorus, thus changing the whole dramaturgy of songs [117].

In **room acoustics**, the frame of this thesis, dynamics as the variation in signal amplitude over time has not been explored much until recently. Thus the findings up to now are fundamental and must be developed and discussed further.

Some observations and experiences of the respective authors suggest that there existed some kind of dynamic difference between venues [17, p. 157]) but it seemed neither explainable ([19, p. 509], "immeasurably") nor overly relevant since the room itself is considered to be a linear system. There had been evidence that perceptual attributes such as *Apparent Source Width* depend on the source playback/listening level [118] and thus would vary with a different dynamic range. However, the dynamic range of music as such was not found to be an important attribute in concert acoustics until recently.

There are several auditory phenomena that are not revealed when analyzing impulse responses: this lack includes non-linear excitation by the <u>sources</u> (instruments producing more overtones when playing louder), thus changing the spectrum depending on the dynamics [17, p. 36], [13, p. 371]. Combined with the non-linear perception and binaural directivity of a human <u>receiver</u> [119] this connection led to the observation that early lateral reflections provided by the <u>room</u> enhance musical dynamics. In other words, the acoustics were said to enhance the dynamics and thus the music. Two more studies backed up this finding [114] [120]. A first conclusive model schematic has recently been given by Lokki (Fig. 5.1)[121]) where it is also separated between level and frequency dependent properties per stage. The effects found in above mentioned studies were attributed mainly to the level dependent source spectrum, early reflections in the room and non-linearity or directivity of hearing for the receiver.

5.1.2 Motivation

Several further attributes that were not covered yet in the above mentioned model do seem relevant for the discussion of dynamics in concert halls. The following overlap between reverberation and level or dynamics are discussed in the upcoming sections.

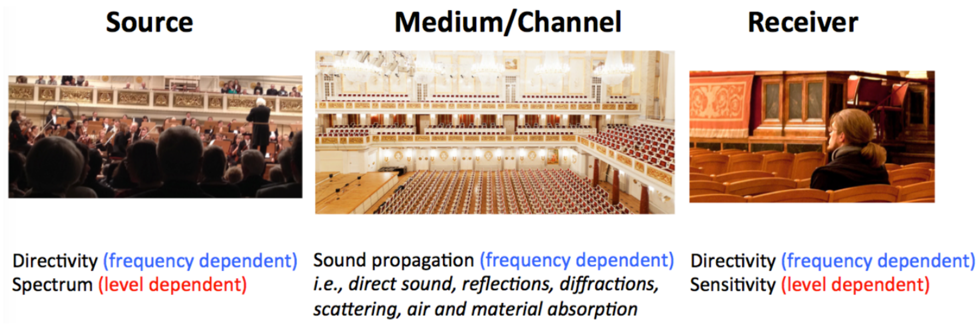


Figure 5.1. Source-Medium-Receiver communication process for a concert hall, from Lokki 2016 [121].

Less accounted for in the above model is the inherent, normal amplitude modulation of the signal due to reverberation. That is, a source with its amplitude varying over time alters the dynamic range of the signal depending on the reverberation in the room. Thus, the dynamic range (or modulation) shall be investigated by comparing the music signal of a constant piece and/or orchestra source in different venues.

As described, typical evaluation and analysis is done with impulse responses describing the system properties of the room as linear and independent of the source or receiver. Thus, concert rooms are amplifying the music linearly depending on the cubic room volume and absorption area, influencing the level. It has been found repeatedly that this room gain, measurable as strength G, is an important quality indicator [9]. Typical preference values were established, e.g. a criterion for strength G of -2 to 10 dB [6], more preferably in the range of 3-7 dB for large concert halls when analyzing popular venues [19], [122]). Likewise in the mentioned recent dynamics experiments, strength explained a majority of the perceptual results and yielded high correlations. But to what sound pressure levels do these strength values translate? There is only little research regarding actual listening levels. Therefore, for information on typically encountered levels and dynamic ranges a large corpus of in situ recordings from different venues and pieces is analyzed.

A last experiment repeats and extends the study presented by Pätynen et al. [120] comparing crescendos from different halls, in order to verify findings utilizing a different technique and set of stimuli as well as to also tackle the effect of overall level on dynamics in concert halls.

5.1.3 Visualizing dynamics

Figure 5.2 shows the sound pressure level over time for a 60 sec Lohengrin excerpt. Besides L_{eq} and L_{max} percentile levels are given. The histogram of the level parameters belonging to this example is given in Fig. 5.3 and offers a different view on the signal: the distribution of sound pressure levels is indicated in this representation instead of SPL over time. The quiet part in the beginning and the longer louder part become visible again. The shape of the histogram is slightly different for different percentiles, e.g. L_{max} and L_{eq}. The A-weighting filter curve is applicable for low volumes (30 phon) which is seldom the case for music, thus mostly unfiltered ("Z-weighted") SPL values are shown in accordance with Meyer [116, p. 279 ff.]. In this thesis, energy-equivalent level L_{eq} and sometimes the percentile level L_5 are mostly used with maximum-minimum ranges. L_{eq} and L_5 are calculated for windows of 1 s duration, the values are pooled and maximum-minimum ranges or 90-10% percentile ranges computed (no overlap, "fast" integration time of 0.125 s).

Meyer argued that the difference in sound power level between a quiet solo instrument and full loud orchestra tutti amounts to approximately 60 dB [17, p. 279]. Thus, one step in dynamic notation would be ≥ 10 dB. Subsequently, dynamic notations are introduced on the right side of graphs for better understanding and discussion (*p* for *piano*, *f* for *forte* etc.). The values are considered to be appropriately scaled for a full orchestra after subjective evaluation of all subsequently analyzed audio examples. However, it shall be noted there has been no formal investigation connecting sound pressure level and dynamic notatio., Thus, the notation serves for orientation only.

5.2 Initial considerations

5.2.1 Change of dynamic range due to reverberation is signal dependent

It is known and accounted for in hearing research that (late) reverberation reduces the modulation depth, namely the dynamic range of the signal over time. In the case of speech perception this is mostly undesirable as the signal-to-noise ratio (SNR) decreases. Subsequently, some parameters were introduced for assessing speech intelligibility that measure the degree of modulation. In concert hall research this effect has not been studied for the music signal itself but indirectly via energy considerations of the impulse response (e.g. ratio of early to late energy C_{80}).

In the following, the set of stimuli from the previous experiment of Section 3.2 is analyzed regarding dynamic range. Here, artificial reverberation was introduced in steps of 0.5-1 dB (energy after 80 ms only). Fig. 5.4 shows the dynamic range, calculated according to Section 5.4 for an 8 second saxophone stimulus convolved with seven increasingly reverberant impulse responses. As expected, the signal range is lowering. Fig. 5.5 shows the same acoustic situations, but convolved with a fairly steady violin phrase without breaks or a clear decay (15 s). Here, the same tendency can be seen, but the differences are much smaller. Thus the dynamic range, as the excited reverberation, depends on the input signal and its variation over time.

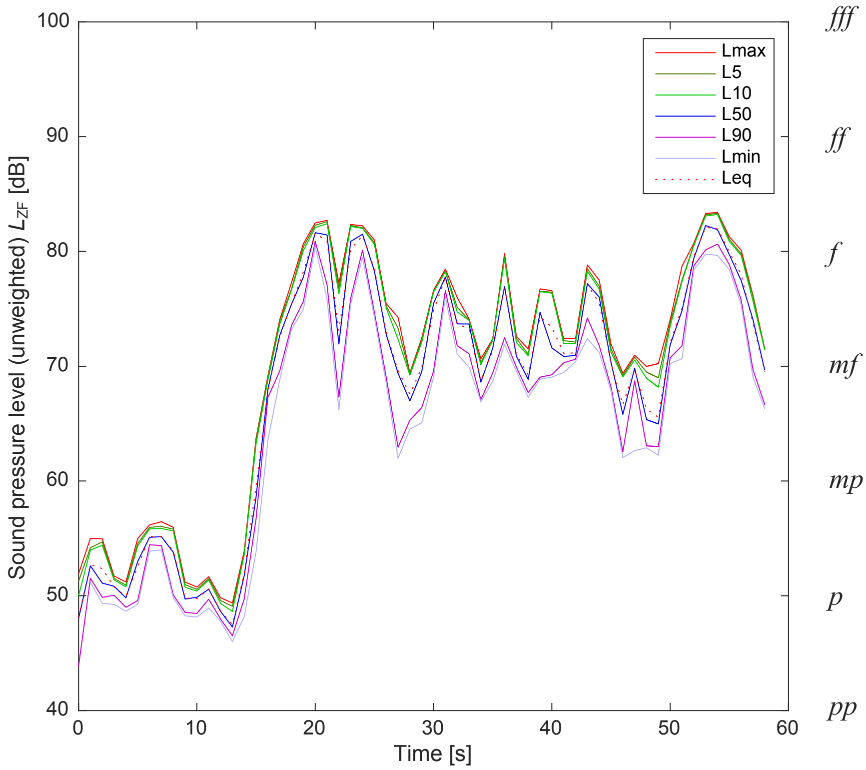


Figure 5.2. SPL over time, different percentile levels for a 59 sec excerpt of "Lohengrin" with a brief quiet section and a longer mezzoforte-forte part. Dynamic notations suggested to be appropriate for an orchestra as a whole.

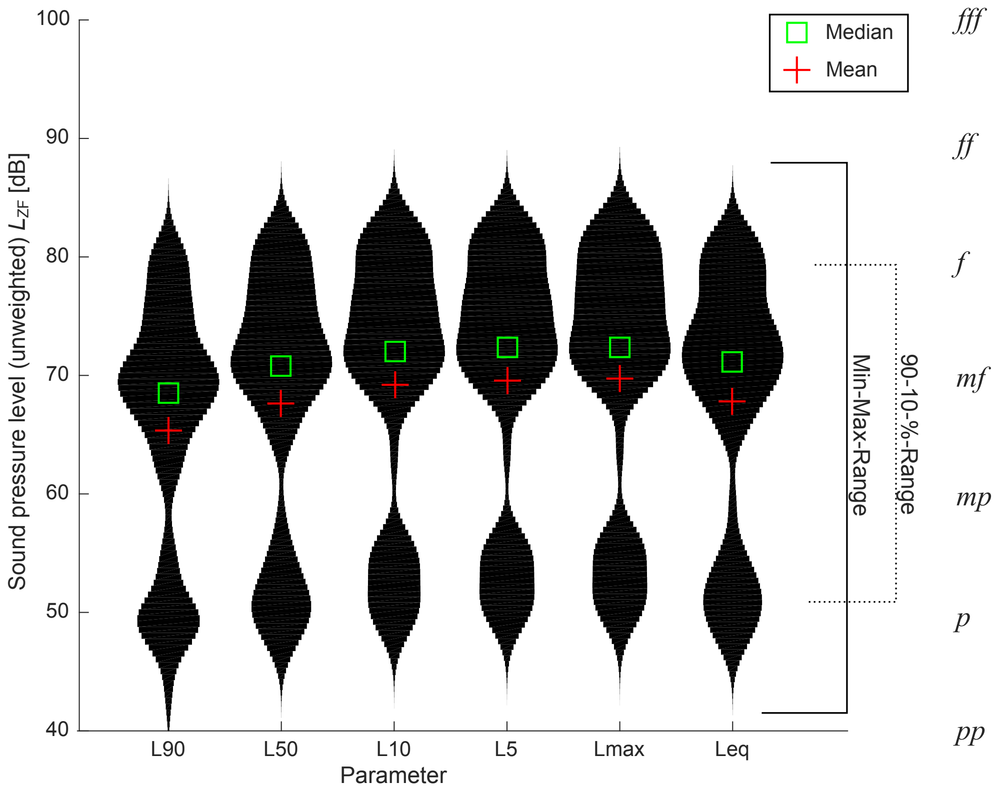


Figure 5.3. Histograms/distribution of SPLs for different percentiles for a 59 sec "Lohengrin" excerpt with a brief quiet section and a longer mezzoforte-forte part. Dynamic notations suggested to be appropriate for an orchestra as a whole.

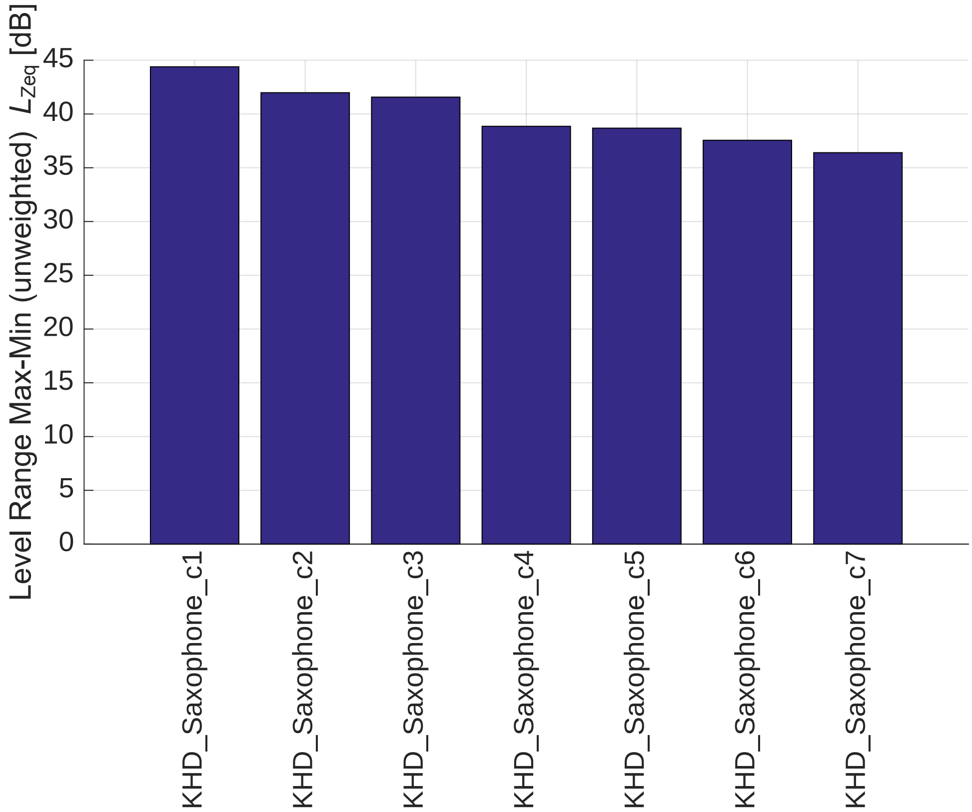


Figure 5.4. Increasing reverberation (c1 to c7) reduces the dynamic range of a lively saxophone signal.

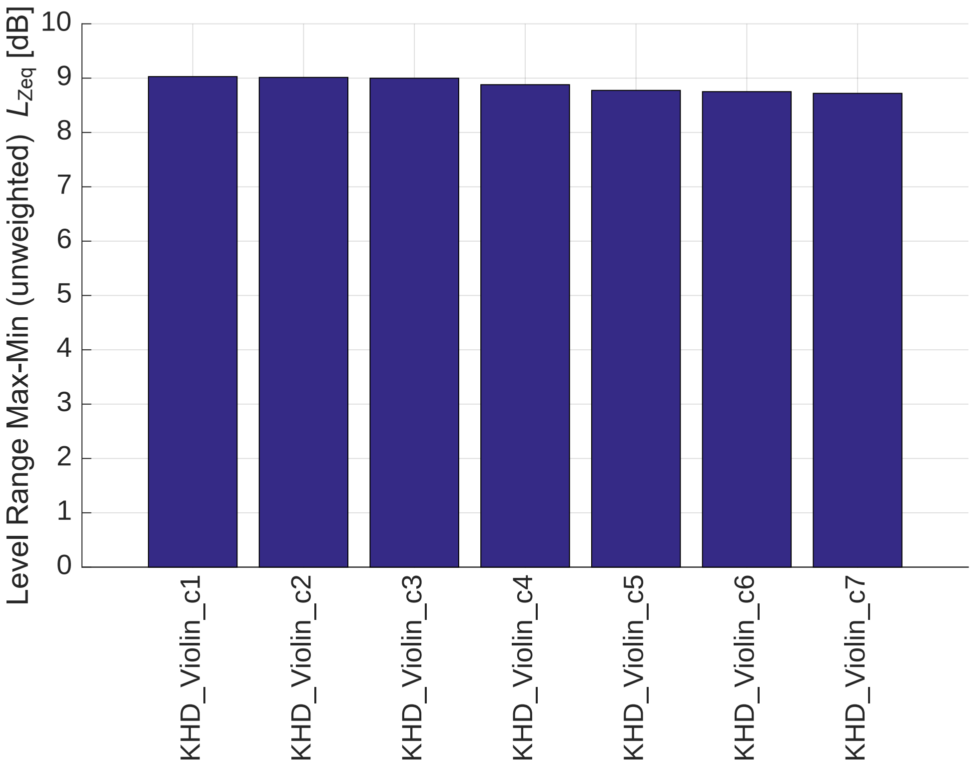


Figure 5.5. Increasing reverberation (c1 to c7) reduces the dynamic range of the Violin signal, but less so as the signal is not changing as much over time.

5.2.2 Different influence of early and late energy

As the energy in the sound field decreases over time, the contributions of early and late reflections to the signal are different. Early reflections together have more energy than the direct sound at most distances in a hall. The ratio between early and late energy, known as clarity C_{80}, can range typically from -5 to +5 dB for performance spaces [6, p. 18]). Figure 5.6 gives an example how the different components of the impulse response alter the signal over time: analyzed is a 45 second excerpt from a solo Oboe piece shown in Fig. 5.7, convolved in a simulated concert hall (15,000 cbm, receiver distance 7 m from the conductor). The setup is explained in more detail in Section 5.4.

It can be seen that the direct sound (dotted line) indeed does not contribute much to the overall energy (solid black). Furthermore, the early energy (red) seems to often contribute to the maxima or *peaks* in the signal, whereas the late energy (blue) more often changes the height of the minima or *dips*, especially in short breaks (i.e. around 6, 19, 24 and 26 seconds).

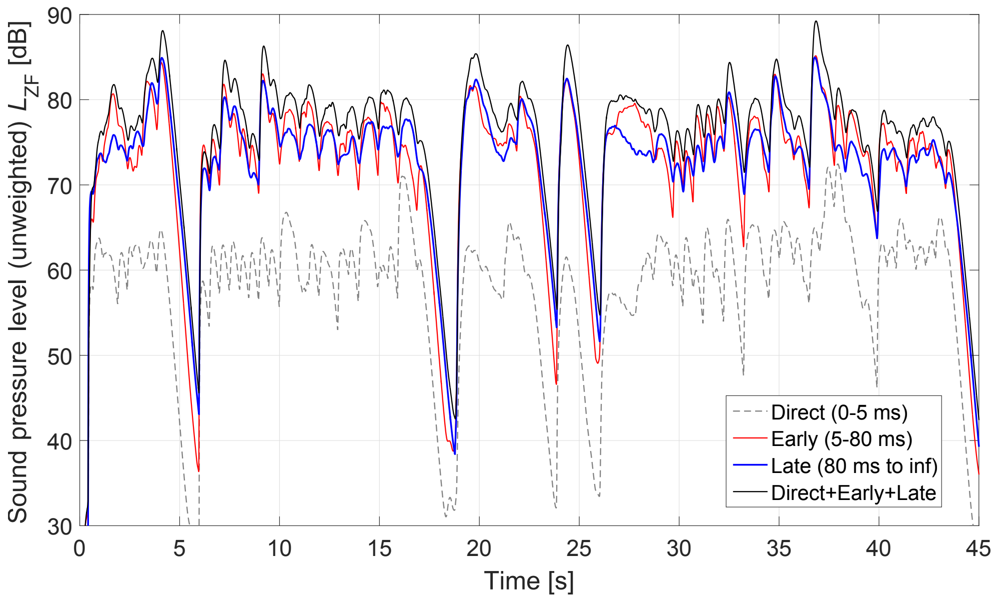


Figure 5.6. SPL over time for an Oboe playing in a simulated concert hall, convolved with the different time windows of the impulse response. "inf." stands for infinity, i.e. end of the impulse response.

5.2.3 Dynamic range changes over distance

Barron's revised theory [10] offers an approach to calculate separately the level of direct, early and reverberation energy at different distances assuming a diffuse sound field. It becomes clear that the modulation depth of the signal would lower with increasing distance.

Three receiver positions from four large unoccupied concert halls are compared (see Fig.5.8, real recordings, same orchestra). The calculated dynamic ranges shown in Fig.

Figure 5.7. Oboe score from Brahms' 4th symphony, 3rd movement for the audio excerpt, public domain from IMSLP.org.

5.9 are for a 67 seconds excerpt from Bruckner's 3rd Symphony with strong dynamical contrasts. It can be observed that the closest position to the stage (1) has a somewhat higher range of levels than a position further back in the stalls (2). However, then an elevated balcony position (3), that is indeed even further away, does not necessarily have a smaller range unless it is much more distant and reverberant such as position 3 in Musikverein (*).

The reason for the effect is that close to the orchestra, the range is the highest because direct sound and early reflections likely dominate. Moving away from the sound source these influences get smaller or are modulated by the reverberation. However, with the elevated positions the instruments are less obstructed which leads, as Meyer pointed out [116, p. 281], to stronger early sound and a larger dynamic range. Thus, realistic sources and sound field conditions must be accounted for.

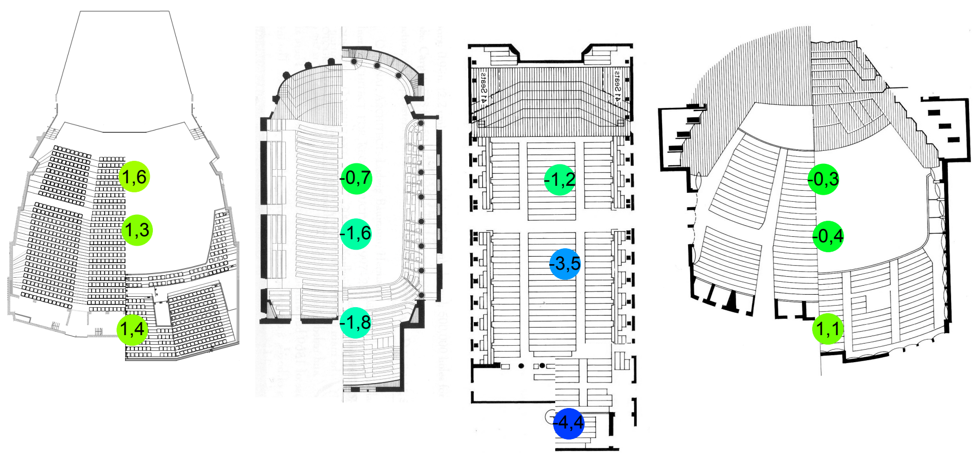


Figure 5.8. Groundplans and receiver positions of the four halls with each receivers 1, 2 and 3 from stage to the back of the hall. Data from the author [43], plans from Müller-BBM and literature [19]. For reference, clarity C_{80}, the energy ratio between late and early energy is shown.

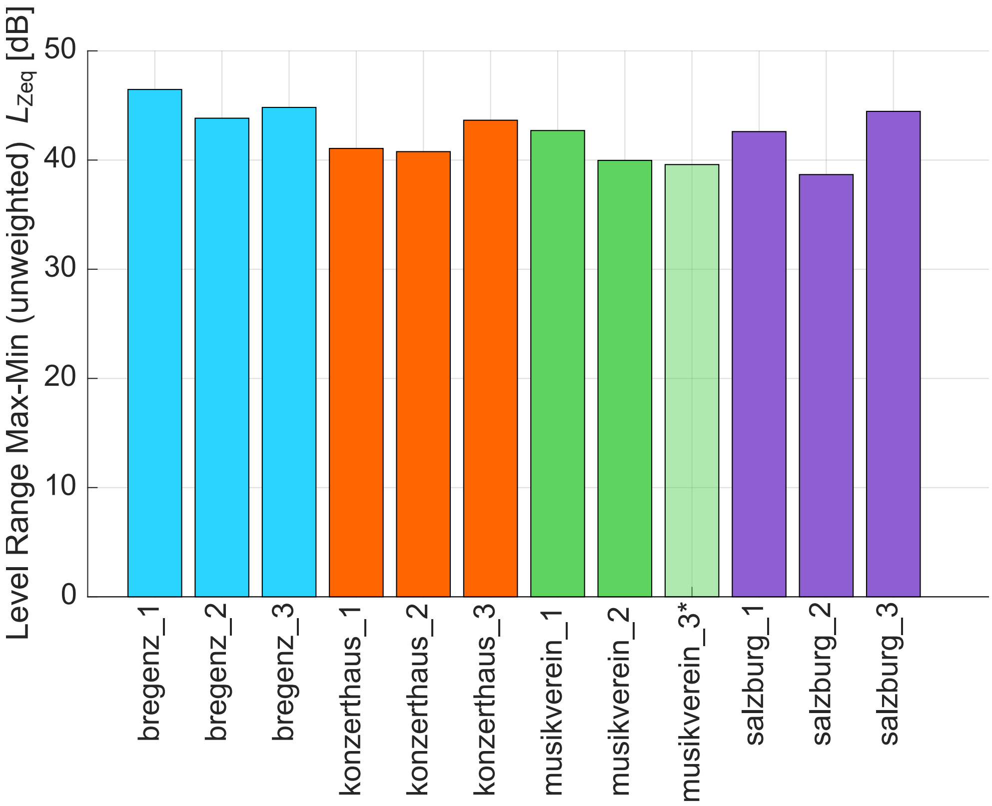


Figure 5.9. Level range for a 67 second Bruckner excerpt, recorded in four halls at each three comparable receivers positions.

5.3 Relevance for music performance demonstrated with Wagners' Lohengrin

Beethoven is known to have considered and changed the amount of orchestra players depending on the acoustics of a venue [2]. If Beethoven was concerned about one aspect, Wagner is the example of a composer caring for the entire production chain of his works: a specific venue, the Festspielhaus in Bayreuth, was dedicated and built for a handful of his own compositions. But through their popularity Wagner's operas are nowadays played in theaters and opera houses all over the world with widely different acoustic conditions. How does this affect the performance and musical perception and what can be observed in recordings or measurements? See also [123].

5.3.1 Introduction: considerations regarding performance practice

In can be claimed that the music on concert halls' program sheets stayed rather consistent over the last 70 years with classical and romantic music making up most of the repertoire. On the other hand, hall geometries and thus venue acoustics changed considerably during the last century [17, pp. 317-318]. This change must have influenced the overall loudness and dynamics. To the *Wiener Klassik* and beyond it was quite normal for composers to be dealing with different rooms and performance conditions for their works since there simply were no dedicated concert halls as spaces for music yet. Opera houses and theaters, on the contrary, had developed earlier around 1650, along with the increasing popularity of operas as a moneymaking spectacle and the demand for staging them. Also for these operatic pieces it was normal and welcomed to be played in different venues and configurations since public success of a piece in one venue would hopefully result in staging the production in other places.

Therefore, Richard Wagner's idea of creating Bayreuth as a single venue for his works stands out. This concept probably comes from a more fundamental change of the self-conception of the composers during this romantic period, evolving from music producer/entrepreneurs and virtuosos to *the artist and genius* showing his or her idea(l)s and work of art [1]. The idea and concept of the Festspielhaus must have developed gradually through various influences and thoughts such as the desire for more artistic control, a dedicated theater festival and perfectly suitable conditions for his works. As a part of this, the acoustics, quite intentionally or not, turned out to be unique. Nowadays, the popularity of Wagner's operas is unbroken. Thus, there is no doubt that the works are performed in a variety of acoustic environments. In the 2011/12 season (no special anniversary year) five out of the top 20 performed operas in Germany were Wagner pieces, totaling 63 productions or venues and 363 shows with 270,000 visitors [124].

Lohengrin was the second most performed among the Wagner operas in the above mentioned survey. Even though the piece had been completed more than 20 years before the Festspielhaus was opened, it is only a few years from the Rheingold, and Wagner's first documented but substantial ideas about the Festspielhaus. His ideas included preference for a democratic seating arrangement and a covered orchestra pit [125]). Lohengrin is therefore surely not composed specifically for Bayreuth, but it can be assumed that Wagner followed a sound ideal for his theater that fits most of his works.

5.3.2 Setup

An excerpt from Lohengrin is compared for Festspielhaus Bayreuth and a german mid-sized opera/theater, both occupied in situ-recordings without artificial reverberation or spot microphones. The comparable recording methods are described in section 2.3.1.2 (Bayreuth) and Section 2.3.1.1 (theater). This comparison allows for switching between two rather different rooms - the famous example on one hand and the more common "average" opera house on the other. In the second house, in the following referred to as "Theater 2", measurements according to ISO 3382 [6] have been performed using

[1] In Wagner's case this longing was most likely emphasized by a history of unpleasant conditions during his life such as financial problems and unsteady environment.

omni-directional sources and receivers. Room acoustic measurements for Bayreuth are taken from literature [19] and in-house-data [126]. Two musical pieces have been chosen for audio analysis, each covering different situations: an excerpt with quiet orchestra accompanying a solo tenor, piano increasing to forte (act 3, duration: 1 min) and a second excerpt with a quick crescendo to fortissimo with full orchestra and tenor, then choir and soprano (act 3, duration: 1 min). The two venues could be described as follows:

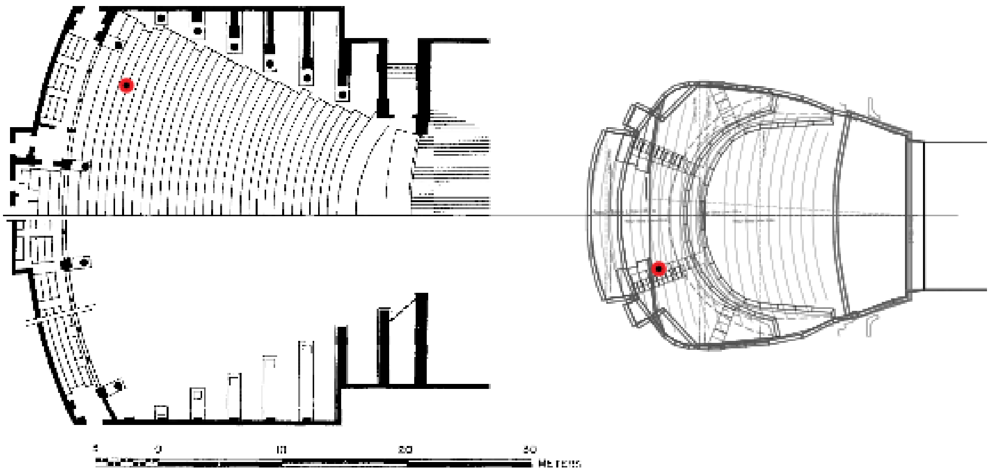


Figure 5.10. Ground plans of the Festspielhaus Bayreuth [(1)] (left) and the medium-sized "Theater 2" (right). The circle marks the receiver position in each venue.

Festspielhaus Bayreuth

The amphitheater style auditorium with inclined seating offers equal sight lines and supposedly uniform acoustics in general. The side-wall reflections are somewhat reduced by the entrance door niches, probably returning those as late reverberation. The orchestra pit is half-covered with a lid to prevent the visitors from observing the musicians. This situation is special for the musicians which therefore rely on the conductor more than in other venues, who himself is balancing stage and orchestra sound to be synchronized within the audience area. A performance becomes more challenging than in other venues. The proscenium surfaces as well as the ceiling are helping to project the sound from the stage. The ensemble size is bigger than the average opera orchestra with highly skilled players and singers with powerful voices.

Theater 2: A medium-sized opera house

Theater 2 is a typical two-balcony, multipurpose house in a medium-sized city used for opera, ballet and (amplified) plays. The orchestra pit has an uncovered size of about 15 m by 5.5 m and only a small covered part below the stage. The audience on the balconies can see most of the musicians in the orchestra pit. In the inclining stalls area the orchestra is shielded by the balustrade of the pit. The proscenium walls and ceiling elements deliver strong reflections which create some sort of spatial impression as well as some erroneous localization perceptions for the audience in the stalls area. This effect vanishes in the two balconies. The ceiling of the hall shows folded elements which project the sound from stage and orchestra pit directly to the audience with a relatively even distribution. Parts of the first balcony are overhung by the second balcony which causes the sound energy to decrease. A comparison of the two rooms is given below (Table 5.1).

Table 5.1. Room data

	Bayreuth	Theater 2
Volume audience [cbm]	10,300	6,000
No. of Seats	1800	1120
Rec. Distance to stage [m]	25.0	20.6

A small, informal listening session was conducted. Outtakes of the two recordings were presented over headphones (AKG K701 connected to a Fireface UC) to four colleagues from within the room acoustics department. They were asked to listen to the following music and comment on the sound, and give their impression at first in general and then more in detail regarding balance, reverberation, loudness and preference. It was not revealed ahead where the audio examples were from.

5.3.3 Results

5.3.3.1 Measurements

When analyzing the ISO-measures it can be seen that the two houses differ noticeably in their reverberation times (Table 5.2). Values for Bayreuth are roughly 0.8 seconds longer, probably mainly because of the bigger volume. The strength value is only slightly different between the receivers. Even though early energy is likely higher in Theater 2, there is less late energy which overall leads to similar strength values.

Fig. 5.11 shows the sound pressure level over time: First of all it can be seen that the level in Theater 2 is higher. Furthermore, it appears as if Theater 2 has a bigger dynamic range (height of the peaks and dips). Interestingly, the difference is even bigger in the first 15 seconds (pianissimo). Analyzing the statistical level distribution in Fig. 5.12a, the occurrences of levels, reveals that the overall range and the shape of the histogram are quite different. Bayreuth (left) has in fact a smaller dynamic range and 2-3 main areas of activity. Theater 2 (right) has a much wider distribution with 1-2 main areas and louder parts around 10 dB. This information cannot be read from level averages mean which are different by ca. 3 dB. In fact, when computing the maximum-minimum level range between L_{eq} values for successive windows (Fig.

5.12b) the audio file from Bayreuth returns an average range of ca. 35 dB and Theater 2 a range of 45 dB – a big difference. Note that the range is computed in windows of one second and thus does not simply equal the difference between overall maximum and minimum. This observation stays qualitatively the same also when excluding potential outliers by applying a 5%-percentile. Similar effects are observed for the second sound sample: Bayreuth has more running reverberation, filling up the gaps with sound, whereas this is more separated in Theater 2. The difference can also be observed in the second excerpt e.g. around 55 s (Appendix, Fig 6.8 on page 151). It is interesting to see for both examples that the brief quiet parts (*dips*) are roughly at the same height, whereas the louder parts (*peaks*) stand out more in Theater 2.

The analysis was also done for the whole duration of the complete act 1, which yielded range values of around 50.5 dB (Bayreuth) and 53.7 dB (Theater 2). Thus, over the course of the whole piece the difference is smaller but still present (applause was removed, silences etc. included).

Table 5.2. Acoustic Measures

	Bayreuth$^{(1),(2)}$	Theater 2
T_{30} occ. 0.5/1kHz [s]	1.7	0.85
T_{30} occ. 125/250Hz [s]	1.8	1.05
T_{30} occ. 2/4kHz [s]	1.45	0.75
Strength G 0.5/1kHz [dB]	4.0	3.2

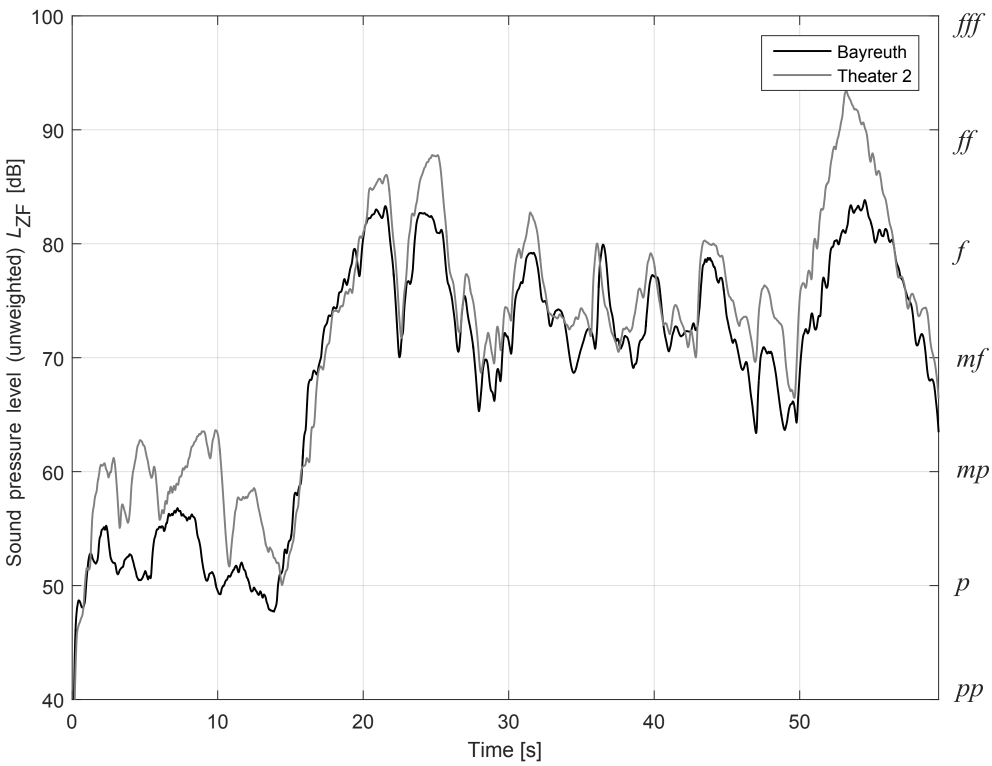
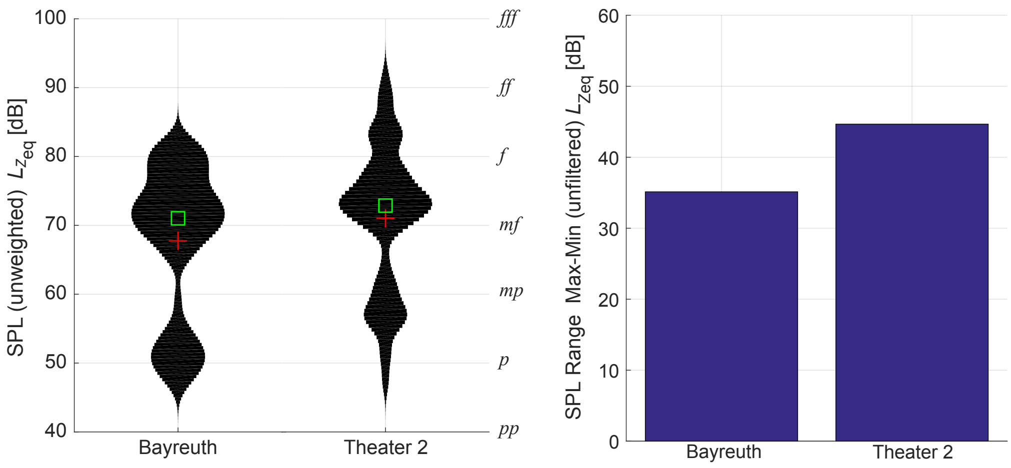


Figure 5.11. Sound pressure level over time for Bayreuth (black) and Theater 2 (grey).



(a) Level distributions and means (red cross). **(b)** Level ranges for the two samples.

Figure 5.12. Analysis of sound pressure levels for the “Lohengrin” audio excerpt 1.

5.3.3.2 Listening Impression

Listening impressions were collected from the informal, blinded listening sessions with four experts (experienced acousticians). It is given as supporting information as it does not represent a formal listening test:

The difference in the listening impression is large and was mentioned mostly with vocabulary related to attributes such as reverberance/decay, loudness, orchestra sound and dynamics as well as the balance between orchestra and singers.

In Bayreuth the sound is quite massive and immediately stands out compared to the other venue. It could be described as diffuse and less defined with more ease for the singers to compete with the orchestra. This follows literature [19, p. 287] describing the overall better balance between singers and the orchestra, which itself is desirably limited in its upper dynamics due to the covered orchestra pit. There is a sense of the big volume. Voices and orchestra have a “halo” of reverberation, carrying every phrase, connecting musical lines and mixing together nicely. The reverberation seems almost too much at times. During the listening session this might be emphasized by the lack of visual and room information. Bayreuth’s seating capacity and small row distances would not be possible today; this keeps the overall absorption area in the favorable ratio.

The musical performance in Theater 2 is overall more dynamic, forte passages are louder, the sound decays more quickly. The smaller room volume is certainly audible which makes it more intimate. Orchestra sound in theater 2 is more clear and direct, both from orchestra and singers. Yet the voices are not carried by the dry room well which lets musical connections fall apart at times.

Preference was different between listeners with a tendency to the overall more “organic” and well balanced sound in the Festspielhaus, supporting the drama.

Figure 5.13. Score of Lohengrin, Excerpt 1, public domain from IMSLP.org

5.3.4 Discussion

The extent to which the signals differ is somewhat surprising. It is of course known that the acoustics alter the signal but the translation as observed as here was not reported before. It could be interpreted that the additional reverberation in Bayreuth is reducing the dynamic range. Secondly, Bayreuth is generally said to be superior or at least more authentic, a fact not objectively investigated here. Assuming that this is correct, not a large dynamic range but an appropriate dynamic range is desirable; this is somewhat in line with state of the art in room acoustics, where e.g. clarity C_{80} (early/late energy ratio) is required to be in a certain range and not as high as possible.

The sound pressure level distributions are as observed because Bayreuth is larger, which lowers the overall level, especially apparent in soft passages. At the same time, the late reverberation, which is likely responsible for the decreased dynamic range, is stronger because of the long reverberation time excited fully in loud parts and smaller ratio of absorptive/reflective surfaces.

A number of limitations are to be mentioned: the recording system was similar and not identical, yet sound pressure levels are correctly calibrated and thus comparable providing the basis for the results. Also, neither the same orchestra, conductor and singers nor production are given. These variables are thus not constant and also not negligible. Still, the overall differences between the acoustic situations are considered to be dominating which is emphasized in the next section where the source will be kept absolutely constant. The study also compares two fairly extreme cases. However, the conditions are common practice and represent a realistic case as encountered daily. It opens a very relevant discussion from the stand point of musical performance practice.

5.3.5 Conclusion

Audio excerpts of Richard Wagner's Lohengrin were analyzed in two opera venues: Festspielhaus Bayreuth and an average, medium-sized theater. Large, measurable differences in the dynamics of the signals can be observed, attributed to the different acoustics of the two venues. Assuming that Bayreuth, considered the ideal Wagner venue, is more suitable and authentic, dynamics that are too large in the other theater do not seem desirable. The smaller venue is louder, less reverberant but clearer. The lack of reverberation is not always noticeable during performance, but for every pause gaps are apparent. In comparison to Bayreuth, these gaps impede the harmonic progress and dramatic character.

The case-study proves to be a prominent example for the link between acoustics and musical performance practice.

5.4 Constant orchestra source in simulated rooms

In the previous examples the sound sources were representative for each venue but not entirely constant in their sound power or exact receiver position. In this data set the source, an anechoic orchestra, is kept constant and the audio files are created by means convolution of computer simulated IRs. The distance to the receiver is fixed. The study investigates how the dynamics of a full orchestra compares between concert halls and a smaller rehearsal space, with and without acoustic treatment. The situation under consideration is a very commonly encountered scenario: even though larger rehearsal spaces would be desirable, financial and logistical considerations limit the choice. When subsequently designing the rehearsal space, there is a conflict between achieving a decay or reverberation time similar to a concert hall and keeping the overall sound level in the room low by absorbing sound energy – two contradictory requirements.

5.4.1 Setup

Three rooms have been simulated in the software Odeon Auditorium v13: a large concert hall, a medium-sized concert hall [127] and an orchestra rehearsal space in three conditions: No absorption, 190 sqm and 250 sqm of absorption (see Table 5.3 and Fig. 5.14). The sound sources were from a full medium-sized orchestra consisting of 56 instruments with given directivity arranged similarly on the available stage area in each venue (orchestra instruments: 2,2,2,3-3,2,0,0,timp,perc,12/10/8/6/4). The large and the medium-sized hall were simulated as occupied (α_w=0.75(L) and α_w=0.9), the rehearsal room without any additional audience. One binaural receiver was placed in each of the three venues at a distance of 7 m. It was not feasible to increase the distance more due to the limited space in the relatively compact rehearsal room (total length ca. 25 m).

Table 5.3. Room data for the five hall models/versions.

Room	Volume [cbm]	Seats	T_{30} 0.5/1 kHz [s]
Large concert hall	23,000	2650	2.0
Medium-sized concert hall	15,000	1300	2.2
Rehearsal hall, no absorption	2,500	/	2.45
Rehearsal, max. absorption	2,500	/	1.15
Rehearsal, medium absorption	2,500	/	1.3

A 90 second long excerpt of the 4th Brahms symphony, 3rd movement was chosen for convolution, see Fig. 5.15. The last two seconds of the excerpt with a decay to silence were excluded for the analysis. Both the orchestra and the convolution data comes with Odeon. However, the size of the orchestra and arrangement on stage were adjusted. The simulations were done in two stages of accuracy (Odeon presets

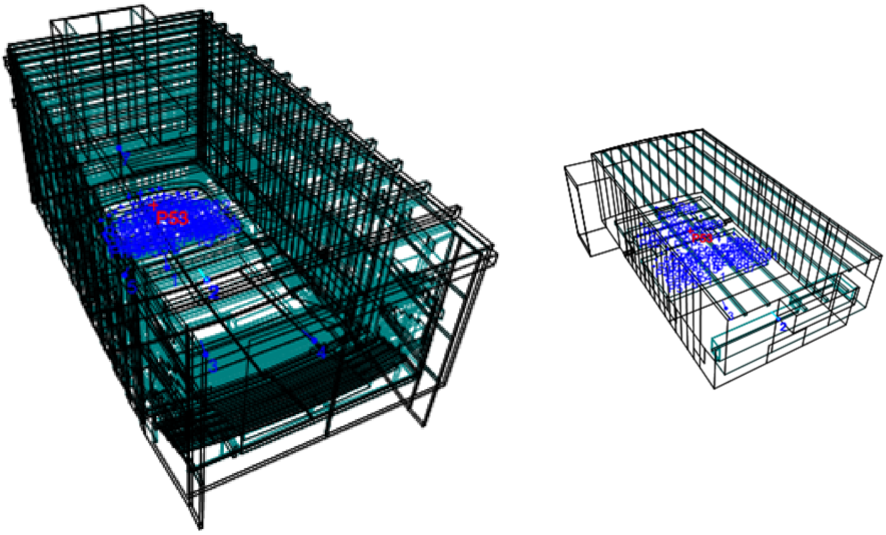


Figure 5.14. 3D-view of the medium-sized concert hall (left) and the smaller rehearsal space (right), both with orchestra on stage.

engineering and *precision*), however no audible or measurable changes were noticed, so the lower resolution was subsequently kept. The level calibration was done close to an Odeon release note [128].

5.4.2 Results

The following single values and level ranges are calculated from the audio files of the convolution mixes for the whole orchestra (Table 5.4). As expected, it can be seen that the mean sound pressure level is smallest in the large hall and biggest in the untreated rehearsal space (80 compared to 85 dB). The minimum level is even more different (61 vs. 68 dB), namely “rehearsal, no absorption” does not get nearly as quiet. This results in different level ranges: the concert halls vary in level ranges between 27-30 dB. The rehearsal space without absorption yields a range of 23 dB, i.e. it has 4-7 dB less of a dynamic range due to the reverberation, while at the same time generating overall higher levels due to the smaller volume. This range can be influenced by adding absorption in the rehearsal room (“Rehearsal max. absorption”). Mean levels drop by 2 dB, minimum levels by 4 dB, leading overall to a bigger dynamic range of 29 dB. With more pronounced breaks as apparent during a whole piece or rehearsal the difference would be even greater. When the last 2 seconds with the final decay to silence were included, the minimum levels were 10 dB lower with added absorption. That means, quiet parts became half as loud.

In other words, by adding absorption to a smaller room, the dynamic range of the concert hall performance situation can be offered in a smaller space. By adjusting the absorption, reverberation time is balanced against level range and overall levels. The

Figure 5.15. First page of the Brahms excerpt analyzed (4th symphony, 3rd movement), public domain from IMSLP.org.

version with medium absorption is still fairly similar in terms of mean and range values but offers a 0.15 s longer reverberation time. Interestingly, the maximum values do not change too much, probably because the early energy dominates which is influenced only a little by the absorption.

The level distributions (Fig. 5.16) confirm this finding and reveal more interesting facts: the rehearsal room without absorption in fact has the smallest range. The rehearsal conditions seem to have two fairly distinct main level peaks or areas whereas in the concert halls these are broadened more. In particular, there is a bigger variety in dynamics in the concert halls. This fact can be observed in table 5.3 from the parameter "90%-10% range" (the difference between percentiles 90 and 10): These *medium* levels, making up 80% between either extremes, have a larger range in the concert halls (17 dB) but are similar among all conditions of the rehearsal room (14 dB).

Comparing the level range to measures from ISO-3382 (not all shown) it becomes clear that there is a strong correlation with the reverberation time T_{30} (r=-0.92) and less so with EDT (r=-0.8) or clarity C_{80} (r=0.6 for simulated C_{80}, theoretical value following Barron's revised theory: r=0.88). This finding is a clear indicator that here the reverberation decay mainly influences the level range.

Table 5.4. Levels and dynamic ranges in the five room models and the anechoic mix equivalent to a free-field recording.

	SPL L_{Zeq} [dB]			Level Range L_{Zeq} [dB]	
Room	Mean	Min	Max	Max-Min	90%-10%
Large concert hall	80	61	88	27	17
Medium-sized concert hall	82	61	91	30	17
Rehearsal hall, no absorption	85	68	92	23	14
Rehearsal, max. absorption	83	64	93	29	14
Rehearsal, medium absorption	83	66	92	25	14
Anechoic mix	91	65	98	33	20

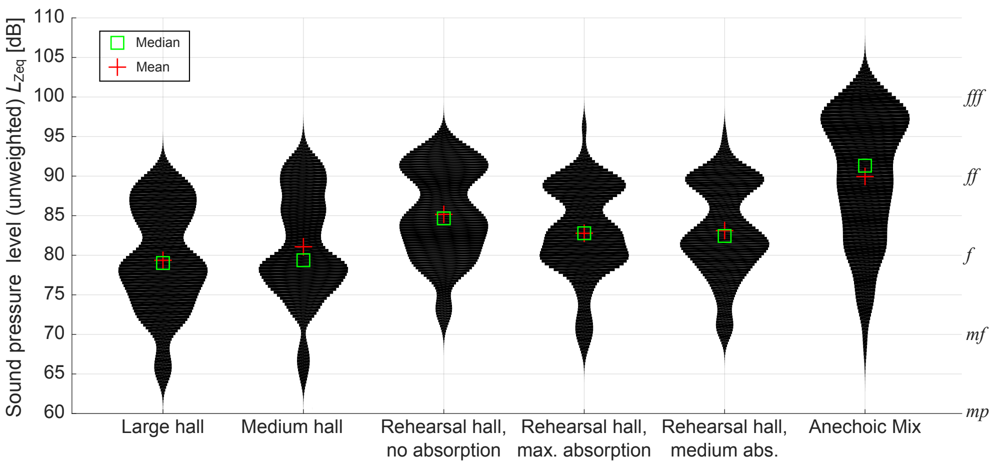


Figure 5.16. Level histograms for the five convolved signals and the anechoic source material.

5.4.3 Discussion

The results are conclusive. However, the connection between acoustics and signal is somewhat outside the normal way of thinking in the field of room acoustics. As an example, the level distribution for a mix of the anechoic tracks is shown in Fig. 5.16 on the right. The mix is convolved following the correct level balance between instruments and has a range of 33 dB (90-10% range: 20 dB). This dry, closely recorded audio material, would decrease linearly in level under free field conditions. In other words the range or shape of the histogram would not change. However, in the room with its modulation due to reflections the dynamic range is decreased, limiting the range of dynamics.

A thorough understanding and assessment is difficult as the purely level-based analysis seems to face limits: perceptually relevant differences between early and late reflections cannot be properly identified and separated. Auditory inspired approaches such as the method from Schuitman et al. [79] can possibly overcome this gap. However, for the time being, these approaches require extensive testing and evaluation of the method itself as it became clear in an article co-written by this author [92].

5.4.4 Conclusions

A full orchestra was simulated in rooms of varying cubic volume and absorption, the convolved audio signals were analyzed. It is observed that reverberation reduces the range of levels, additional absorption in the room re-establishes the dynamic range. Larger concert halls maintain a wider distribution of levels than the smaller rehearsal space is able to offer, the resulting musical dynamics would be larger while being less loud overall (e.g. *pp* to *f* instead of *mp* to *ff*).

The study shows that reverberation as a property of the linear system *room* between source and receiver does alter the signal dynamics.

5.5 Typical levels and dynamic ranges in venues for classical music: analysis of complete concert and opera performances

Most of the previous sample sets were relatively short excerpts compared to a typical concert experience. Thus complete performances from various classical music venues are analyzed with respect to commonly encountered dynamic ranges and levels. To facilitate comparison, levels are given here both A-weighted and unweighted.

5.5.1 Setup

The following data stems from recordings made with microphones attached to a human head (next to the ears but not in the ear channel and thus not strictly binaural, see Section 2.3.1.2). The data has been collected over the course of several years and represents a good average of what a concert-goer would experience in larger venues for classical music. All recordings were done with the same equipment and live, in situ, by the author. A variety of positions in the venues were covered at different positions, therefore an immediate comparison between venues as done before is not reasonable. Re-calibration was done after periods of one year, where no major deviations were found over time, the measurement accuracy is thus estimated to ±1.5 dB. Relative values such as the dynamic range are not affected by this change. Pre-amplifier clipping occurred with continuous pink noise at a level of L_{Zeq} = 100 dB and L_{Aeq} = 96 dB. The noise floor of the recording system including microphone, pre-amplifier and converters was measured at L_{Zeq} = 31 dB and L_{Aeq} = 24 dB.

A number of internationally renowned venues are included in the study. Abbreviated information about venues and the musical program is given in Figs. 5.17 and 5.18 in the *Results* section. The musical program consists mainly of typical pieces of the classical and romantic era including full orchestra. For several venues, the same hall was measured at different performances in varying seats. The audio material was divided between concert and opera performances:

For “concert”, there were 15 different venues included. Concert halls ranged in size from 3,000 to 30,000 cbm (mean and median around 15,000 cbm) and reverberation times T_{30} from 1.5 s to 3 s (mean and median approximately 1.9 s). 31 audio files were recorded, namely separate acts or pieces with a duration of 18 h 29 min (average file duration 36 min).

For “opera”, 12 venues were analyzed, resulting in 42 files or 40 h 12 min of material (average file duration 57 min). Room volumes ranged around 11,000 cbm with reverberation times T_{30} of ca. 1.4 seconds on average.

Lastly, start and end applause were separated, combined and analyzed for 25 venues (37 audio files, duration 3 h 23 min, 5-6 min clapping per file). Interim applause was removed from the music analysis, even though it was seen that for the whole duration of a 45 min ballet piece this changed the L_{eq} of the piece by only 0.1 dB and the level range by 0.8 dB (percentile level L_5 did not change at all).

5.5.2 Results

Results from the measurements are given as A-weighted sound pressure level in Table 5.5 and unweighted/linear SPL in Table 5.6. The values are power averages from the levels of the individual audio-files. All values are L_{eq}-values, specifically "Min" stands for the average minimum L_{eq}-value of all one second-windows of all recordings [2].

It can be seen that the mean level is around L_{Aeq} = 80 dB for both concert and opera, the opera being a few dB quieter. Minimum values are measured around L_{Aeq} = 34-40 dB. Maximum values go up to ca. L_{Aeq} = 95 dB with the applause reaching 100 dB, where pre-amplifier clipping could have started to occur (*). Lastly, the level range lies in the area of L_{Aeq} = 60 dB. 80% of the levels are found within a range of 24-30 dB (90-10% range). When no weighting is applied (linear SPL, "Z"), values are a little different, most noticeably the range gets smaller as the minimum values are higher.

Table 5.5. Average L_{Aeq}-Level values for a corpus of in situ audio recordings from classical music venues.

	SPL L_{Aeq} [dB]			Level Range L_{Aeq} [dB]	
Performance	Mean	Min.	Max.	Max-Min	90%-10%
Concert	81	34	95	63	30
Opera	77	40	94	56	24
Applause	92	81	100*	46	7

Table 5.6. Average L_{Zeq}-Level values for a corpus of in situ audio recordings from classical music venues.

	SPL L_{Zeq} [dB]			Level Range L_{Zeq} [dB]	
Performance	Mean	Min.	Max.	Max-Min	90% -10%
Concert	84	49	98	50	25
Opera	79	54	95	45	21
Applause	92	81	100*	41	7

Figures 5.17 and 5.18 show sound pressure levels for the individual audio files attributed to "concert" and "opera" performance respectively. Analyzing the data points, it appears that the differences between individual files are also to be attributed to the different music source and program. For example, in Concertgebouw, the *Wiener Klassik* piano concerto (data point 3) is approximately 5 dB quieter than the *romantic period*, full orchestra piece (data point 4, Fig. 5.17). As there are several possible influences for the differences (musical program, venue, receiver position), it does not seem reasonable to draw conclusions regarding the acoustic conditions. Level ranges are given in the appendix (Fig. 6.10 on page 153 and Fig. 6.11 on page 154).

5.5.3 Discussion

The results above are in good agreement with literature and support earlier findings.

Preferred listening levels of around 79 dB in peak ranges, A-weighted, were reported for two fairly different musical motives [109, p. 356] in a laboratory test environment.

[2] The Max-Min-Range is similarly computed per time window and averaged and is thus not equal to the difference between minimum and maximum values from the table.

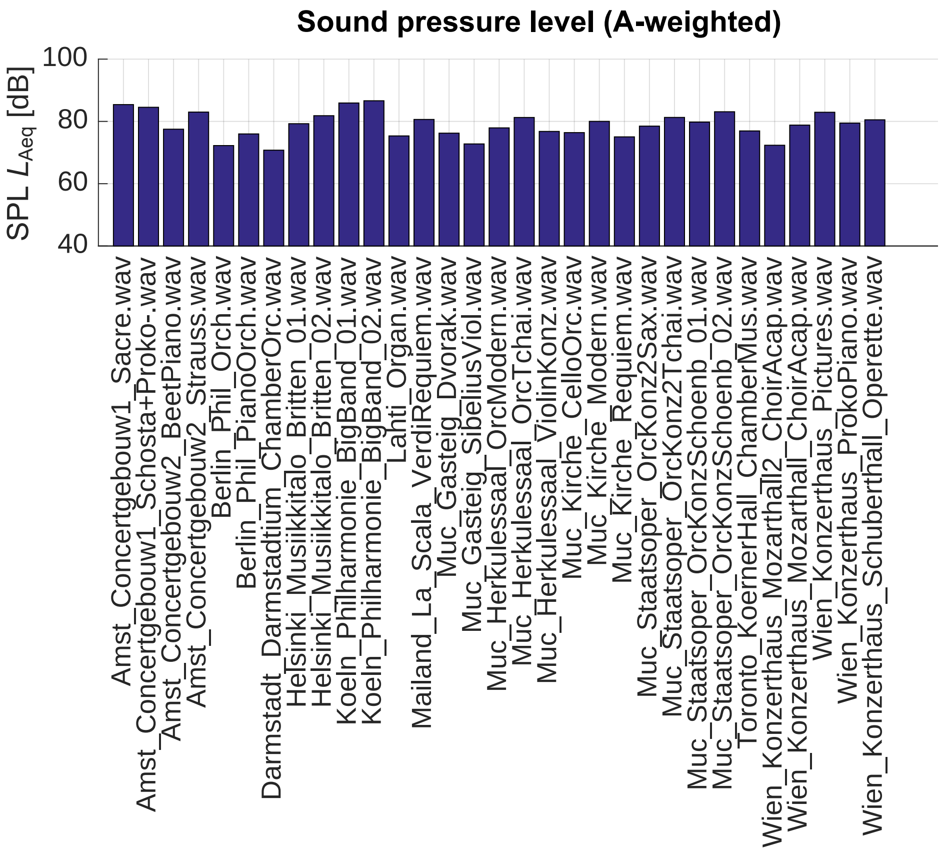


Figure 5.17. A-weighted sound pressure level and level ranges, Concert performances.

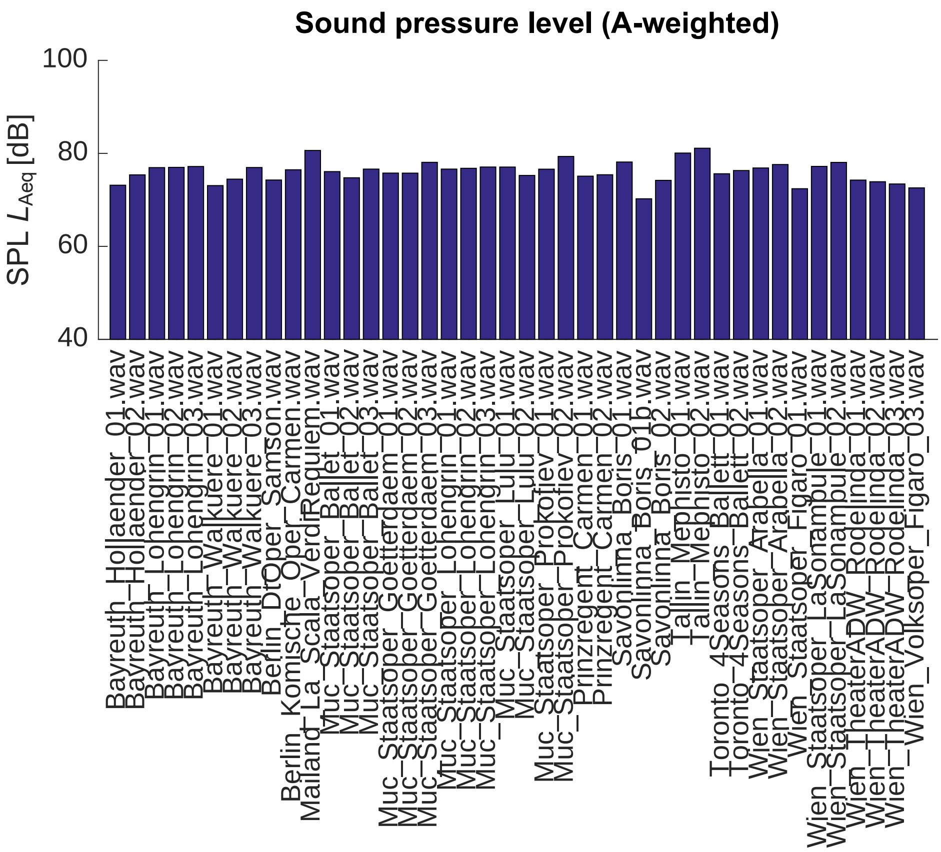


Figure 5.18. A-weighted sound pressure level and level ranges, Opera performances.

The exact duration and type of level/averaging is unknown. Further tests with different stimuli by the same authors showed preferred levels of around 83 dB (A-weighted, in a simulated sound field, p. 371) and in situ testing with music from a loudspeaker on the stage (83-86 dB, A-weighted, p.383). Recently, preferred levels of L_{Aeq} = 72-75 dB for chamber quartet and piano duo were reported in an in situ concert hall group experiment where live musicians played 1-min-excerpts which were artificially amplified to different levels [41]. Likely, a full orchestra would be playing louder and preferred at different values. The mentioned levels all come from experimental setups. Meyer conducted [116, p. 256, 281] parallel SPL measurements at 5 positions in one auditorium and found differences in the level ranges between those receiver positions, which he attributed to the different directivities of the orchestra but not investigated further or compared to other venues. Also, he discussed typical dynamic ranges for instruments: a typical range between single instrument *pianissimo* and an orchestra tutti *fortissimo* would be around 60 dB, derived from sound power differences (no weighting, RMS-Level with custom/manual smoothing). It is argued that a range of 60 dB can be said to be an average but the ideal case is not always met. A dynamic range of 50 dB can be assumed (no explanation given) [17, pp. 278-281]. Here, ranges of ca. 60 dB (A-weighted) or 50 dB (linear) were found. Regarding minimum levels, Winkel argued that when the noise level was at 40 dB, the orchestra *pianissimo* was at 42 dB [17, p. 278]. For the same piece with a noise level of 50 dB in a different room the *pianissimo* could only be played at an SPL of 55 dB. It is not clear if these values are A-weighted or linear and only have observational quality. Lastly, silence in the signal does affect values such as L_{eq}. However, only natural and relevant pieces of silence where kept in the analysis, such as breaks between movements, just as it would be experienced on-site.

The data might also serve as input to increase performance of auditory models for room acoustic evaluation or when applied in a music context.

5.5.4 Conclusion

A pool of binaural in situ audio recordings from several concert and opera venues was analyzed regarding levels and level ranges. For the 60 hours of music material, a value of approximately L_{Aeq} = 80 dB was found as an average music level. Maximum music values reached 95 dB, applause reached 100 dB. A dynamic range or range of levels of ca. 60 dB (A-weighted) or 50 dB (linear) was found. For future research, the corpus should offer valuable input data for room acoustic analysis utilizing auditory modeling.

5.6 Level and dynamics when judging "impact" of crescendos in different acoustics

In [120] it was reported that rectangular concert halls have a greater emotional *impact* on listeners. This finding was attributed to differences in dynamics, by comparing crescendos from different halls. However, the experiments were done with a loudspeaker orchestra as a source. In order to prove that the findings are also valid for real musicians, a real orchestra is used here instead of the synthesized version. Additionally, two different listening distances are compared as well as a loudness equalization attempted to minimize overall level differences. Even though perceived dynamics is not directly measured, it shall be observed if the role/importance of dynamics changes when altering the overall level. See also [129].

5.6.1 Setup

5.6.1.1 Stimuli

A professional orchestra, the Staatskapelle Berlin, was accompanied during a pair of concerts in the Konzerthaus am Gendarmenmarkt (BK, Berlin Konzerthaus) and the Philharmonie Berlin (BP). The orchestra performed Egmont Ouverture by L. v. Beethoven, a typical Wiener Klassik piece, during the final rehearsal before the first concert as well as the short dress rehearsal of the second concert. The same excerpt, containing a 15 s crescendo, was recorded simultaneously in two receiver positions in both rehearsals with the halls unoccupied. The first position in BK was at 6 m distance from the conductor and 2 m off-center to the right (row 5, seat 15, BK1). The second position was on the first balcony at 22 m distance (seat 9, BK6). In BP, the first position (row 6, seat 19) was approximately at the same physical distance as in BK. As BP does not have a balcony, the corresponding position was at 27 m distance in the rear stalls. The receiver positions are shown in Fig. 5.19. These are the same halls and fairly similar receiver positions as were used in the study investigating musical dynamics [114], but not quite as in the thematically more related study investigating emotional impact [120]. The influence of the different receiver distances are discussed later. The orchestra was seated along the German arrangement with 1st and 2nd violins on opposite sides, double basses to the far left, and celli and violas to the center and far right, respectively.

5.6.1.2 Recording and Reproduction

The setup for the orchestra recordings was a head-like recording system as introduced in Section 2.3.1.1. Two calibrated recording systems were used simultaneously in both concert halls, thus allowing a matched A/B-comparison afterwards. The reproduction

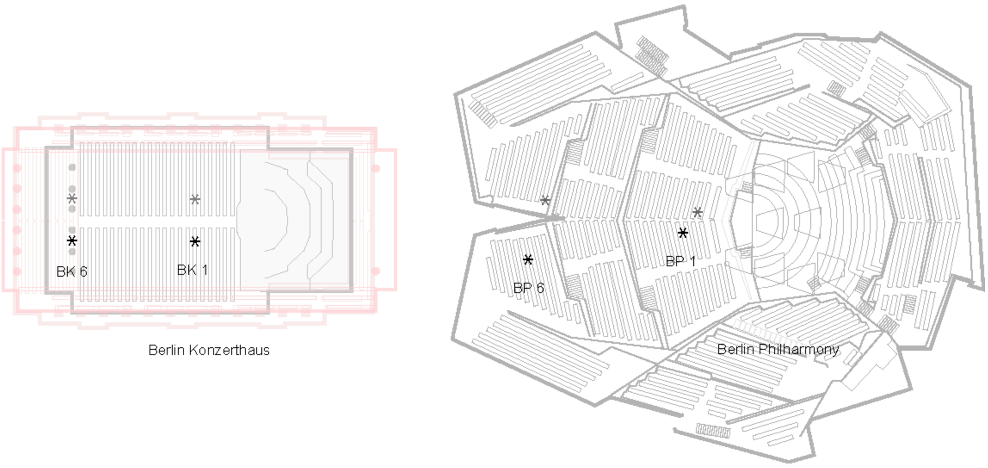


Figure 5.19. Groundplan of the Berlin Konzerthaus (left) and Philharmonie (right) with each two receivers. Closeby receiver position marked in gray were used in a related study [114].

was accomplished with four loudspeakers (Genelec model 8020B) positioned at ±45° angles for the front and ±135° for the rear speakers. The radius of the loudspeakers from the listening position was 1.4 m, and the sound level of each loudspeaker was calibrated. To block sound coming from the left loudspeaker to reach the right ear (cross-talk cancellation) a 4 cm thick absorptive and isolating panel was placed directly in front of the listener's head. The level for the listening test was adjusted to L_{Aeq} of 78 dB or L_{AFmax} of 86 dB for the loudest stimulus. The presented sound levels were approx. 6 dB higher than in situ due to calibration issues. However, the difference was constant for all stimuli. The laboratory calibration was done with a class 1 SPL-meter and a human head with two DPA 4060 microphones attached next to the ear for further processing of loudness and levels (calculations in PsySound 3 [89] using the DLM loudness model). A post-hoc analysis of the sound pressure levels showed that BK1 is around 2 dB stronger than BP1 and BK6, which again are another 2 dB stronger than BP6. The loudness normalization for all stimuli led to an average loudness of 26 sone or an L_{Aeq} of 76 dB.

5.6.1.3 Experiment

The listening task was designed for quick comparison between the four stimuli by paired comparison (see Section 2.4), following the approach in the above mentioned comparative studies. Of the pair of stimuli, the listener was instructed to choose the one which felt to have more impact. The subjects could also indicate a tied response for "no difference". The user interface was provided on a touch screen (Apple iPad, connected to a MacBook Pro). The audio was played back from the computer via Motu 16A audio interface, which feeds the signals to the loudspeakers. The stimuli could be switched seamlessly as the compared recordings were played synchronously. The paired comparison was a full-rank design with two repetitions, each pair was presented

Figure 5.20. Reduced score of the music excerpt. Full orchestral score in Appendix, Fig. 6.6, p. 150, public domain from IMSLP.org.

twice in random order.

A brief training session was conducted before the experiment with four stimuli. Two of the stimuli were used in the actual test and the other two were from the same halls but different seats. A maximum of six training comparisons were offered, though most participants felt comfortable with the task after four pairs. The verbal test instructions informed the participant about the proper seating position and explained the objective of the study on the experienced impact. The brief description defined the term *impact* as having more influence, being more interesting, or more effective on oneself. The subjects were recommended to initially listen to both completely instead of switching quickly back and forth. After the training, the correct understanding of the task was confirmed in a brief discussion. During the test, the subjects were also asked to write on paper the principal differences driving the decision for certain stimuli for each pair. The test session was concluded with a discussion on the collected criteria in order to resolve ambiguous answers and to narrow down the vocabulary. For instance, a mention of spatial impression was defined further to an indication of envelopment, perceptual distance, or source width. Other general remarks on the test were collected as well. Consent was given by the participant to ensure accordance to ethical regulations.

The experiment was conducted in two test periods. In the first test, the stimuli were presented at the original relative sound levels. 18 subjects (2 female, 16 male) participated in the first test. All were Aalto University staff, and half of them could be considered skilled listeners due to their work in room acoustic or signal processing research groups. The average age of the subjects was 31 years. In addition to the paired comparison between four stimuli with two repetitions, a control pair with two identical stimuli was hidden in the test sequence. Hence, there were 13 comparisons in the test. All subjects reported the pair without difference as a tied comparison, and the control pairs were omitted from the subsequent analysis. The average duration for the paired comparison task was 13.6 minutes.

In the second test period, conducted on a different day, the presented stimuli were matched with regard to overall loudness. As this process rendered the task more difficult, relatively more experienced listeners were invited to contribute to the experiment. A total of 10 subjects (1 female, 9 male) participated in this listening test, 7 subjects from the acoustics or signal processing groups and three untrained listeners. The hidden identical pair was removed but three other control pairs were included. Namely, for comparison between the original and the loudness matched version of each BK1 and BP6 as well as a re-test of the two fairly similar stimuli BK6 vs. BP6. The 50% increase in comparisons extended the test duration to 18.5 minutes. Most of the subjects were already familiar with the task from the first experiment. However, they were not told how the stimuli differed from the first experiment.

5.6.2 Results

5.6.2.1 Listening Test

The choice probabilities for offering the most impact according to the BTL model are shown in Fig. 5.21. First, we consider the results of the experiment with the original relative sound pressure levels. Position BK1 clearly has a higher perceived impact than the other positions. The error bars indicate ±1 standard error around the mean BTL value. Statistically significant differences ($p<0.05$) appear between all stimuli except for BK6 and BP6 ($p=0.11$). With Bonferroni-correction applied ($p<0.0083$) only the first stimulus is significantly different from all others. It can be concluded that first, front positions produced higher impact than farther positions in both halls, and second, positions in Konzerthaus yielded in general a higher impact with regard to the respective positions in Philharmonie. This result regarding the sound pressure levels observed in the compared positions was expected.

The results from the loudness matched comparison are shown in lighter shade in Fig. 5.21. Despite the equalized overall loudness, the overall rank order as well as the significant differences is equal to the first test. In general, the magnitude differences between BTL probabilities are smaller. The difference between BK6 and BP6 is not significant ($p=0.34$). Thus, matched loudness equalized the impact in the farther positions but not in the front positions.

For the additional control pairs comparing the original with the loudness matched version of each BK1 and BP6 and original versions of BK6/BP6 the probabilities were in the order of 0.9 for the louder stimulus to 0.1 (not shown here).

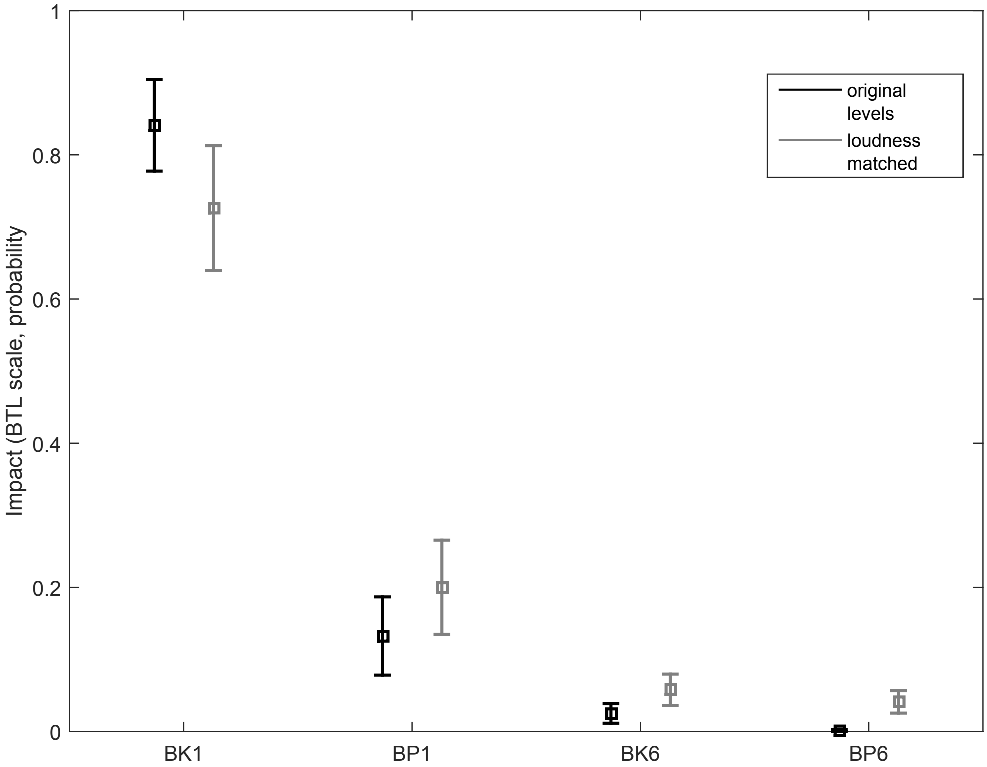


Figure 5.21. Probability of impact rating for stimuli in original condition (black) and loudness matched (grey). Error bars represent ±1 SE.

5.6.2.2 Vocabulary Profiling

For each pair, one or more descriptive adjectives suggest the reason for the particular choice. This data also reveals the perceptual dimensions which the subjects used in evaluating the stimuli. The refined adjectives collected after the discussion with each participant were manually categorized into groups of similar attributes. For the first test, almost 90% of the adjectives can be grouped into seven attribute groups. These groups were further combined into two general categories related to strength or dynamics, and spatial properties. This procedure is shown with the respective results in Table 5.7. For the first test with original relative levels, 25% of the decisions were based on strength attributes. The second most frequent group is proximity, followed by spatial attributes, bass, and clarity. In essence, approximately one third of the decisions were based on dynamics or sound level differences. For the second test with matched loudness, the prominent attribute groups are altered. The strength/dynamics/crescendo group combined accounts for only 11% of the choices. In contrast, spatial attributes are more frequent, and proximity appeared as the most frequent single attribute.

It can be concluded that when the loudness is normalized, differences in strength (and closely related attributes such as dynamics) are less decisive. Other cues then outweigh the decision, but not sufficiently enough to alter the order of stimuli regarding impact.

Table 5.7. Vocabulary Profiling: Adjectives, collected from the participants and grouped to attributes, driving the impact ratings. Percentages greater than 15 are highlighted in bold.

Attribute Groups	Adjective Count			
	Percentage		Number	
	Test1 (original)	Test2 (loudness matched)	Test1 (original)	Test2 (loudness matched)
dynamics+crescendo	9%	3%	32	4
envelopment+spaciousness	10%	9%	33	13
clarity	8%	13%	28	18
bass	12%	**15%**	42	21
width	5%	13%	17	18
proximity	**18%**	**22%**	60	30
strength/level	**25%**	8%	86	11
timbre		4%		6
SUM	88%	88%	298	121
other	12%	12%	40	17
env+spaciousness+width	**15%**	**22%**	50	31
strength+dyn.+cresc.	**35%**	11%	118	15

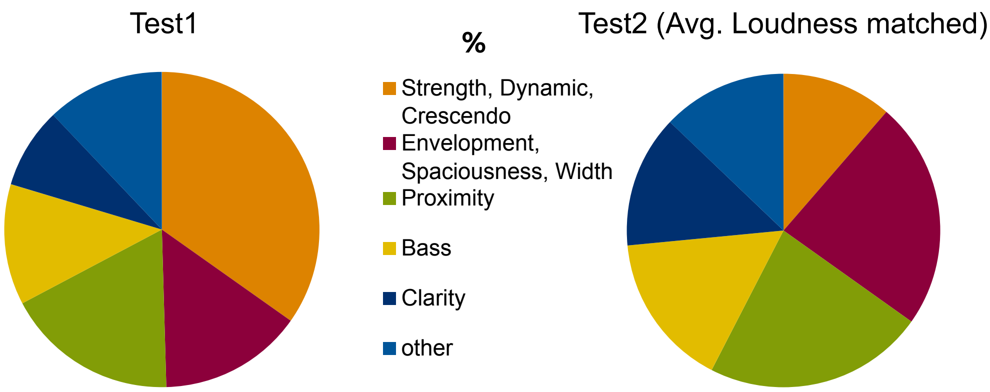


Figure 5.22. Vocabulary Profiling: Adjectives, collected from the participants and grouped to attributes, driving the impact ratings.

5.6.2.3 *Level analysis*

When analyzing the sound pressure level over time measured in the lab at ear position, it becomes obvious that BK1 is around 2 dB louder than BP1 and BK6 while BP6 is 2 dB quieter than the latter. It can also be noted that BK1 overall has a greater variation in dynamics over time and that the difference in dynamics between two seats of the same hall, e.g. BK1 and BK6, is bigger than in Berlin Philharmonie. This difference has been plotted separately in Fig. 5.25. Whereas, the modulation or dynamic difference over time is bigger in BK1, the overall dynamics, namely from second 0 to second 14, is in fact larger in the distant seat of Berlin Philharmonie. This result indicates that it is the micro dynamic range and not the overall dynamic range that should be analyzed separately. Since these level differences could have a big effect, the loudness between stimuli was adjusted to the same mean loudness. The signal envelope is not completely the same as only the overall loudness has been matched. Note that the dynamic range is not changed by the normalization in level offset.

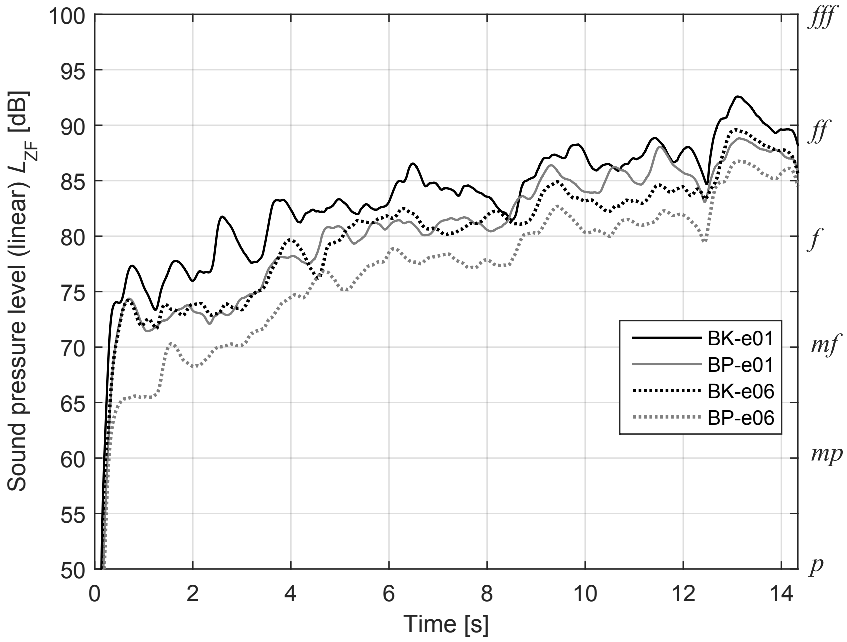


Figure 5.23. Linear sound pressure level over time for the four stimuli in original condition

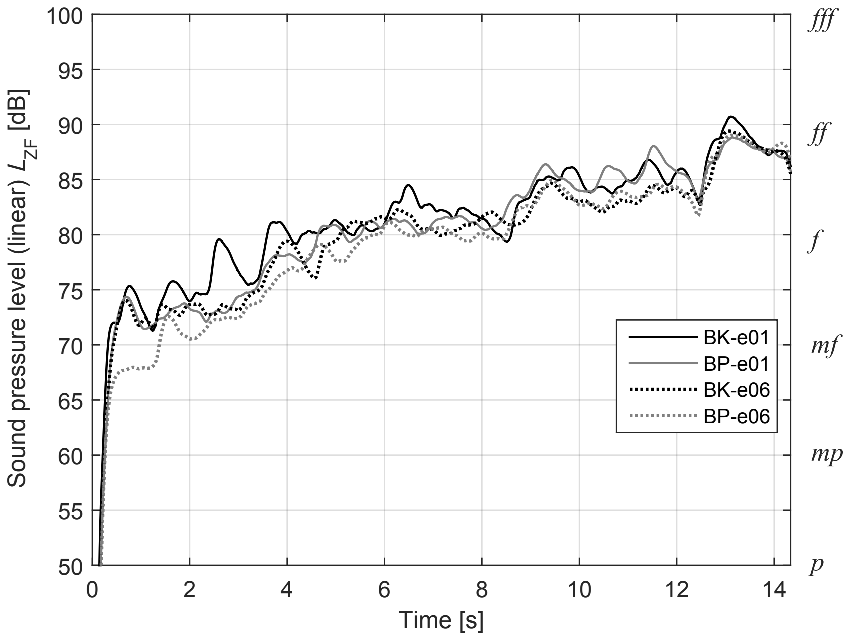


Figure 5.24. Linear SPL over time for the four stimuli with their overall loudness matched

5.6.3 Discussion

The physical distances to the receiver positions were not equal between the two halls, and the absolute judgments for these halls are not entirely justified. However, similar

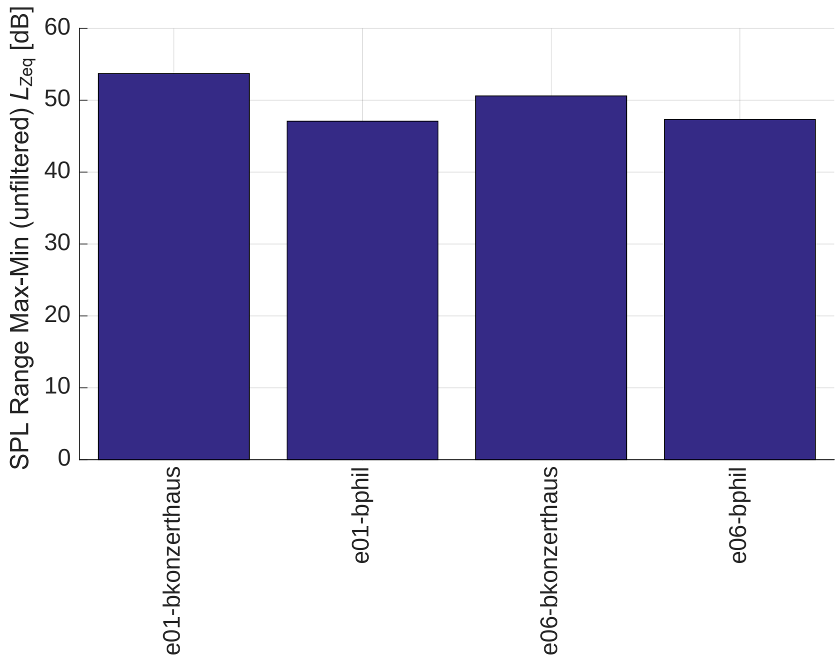


Figure 5.25. Difference between maximum and minimum L_{eq} values (max-min level range) for the four stimuli, no weighting.

hall areas have been used (e.g. R1 is situated in Row 6 for both halls) and the findings are in line with the study using the artificial orchestra [120]. In that study, distances between pairs of receivers were exactly the same, and a significant difference was found between the halls. In other words, the 3 and 5 meter offset in distance between Konzerthaus and Philharmonie receivers is probably not to change the overall outcome.

Previous research utilized convolutions of anechoic orchestra recordings and spatial room impulse responses measured from the halls [120],[114]. Although the present study employs recordings from a live orchestra, the results are in agreement with the previous findings. This finding underlines that both approaches are likely valid for studying subjective impact by concert hall acoustics. The earlier study ([120]) proposed early lateral energy as the main cause for the enhanced impact.

For the more remote seats, the level is more important than for close seats. This assertion is emphasized in the loudness matching of the second test and also observed in the stimulus-dependent vocabulary profiling of the comparative study [120]. The comparison between differently distanced receivers was included and, as expected, the closer seats have more impact. Also, the loudness matching provides, to some extent, a hint that overall trends might not change even when the important level/loudness cue is missing.

Regarding the connection between dynamics and level, it seems in this instance that level is the more decisive choice criterion for impact. When normalizing the loudness, the vocabulary responses for dynamics went down even though the physical dynamic range did not change. This result could also be a hint that the overall SPL influences the perception of dynamics. Furthermore, the physical dynamic is greater in BK1/6 than in BP1/6 (Fig. 5.25) which does not fit the impact ratings. In other words physical dynamic range might not be that decisive for impact. Untouched by this issue, the spatial attributes still also change dynamically (with varying sound pressure level) as shown in [118].

The influence of level is not surprising. Overall SPL has been found to be dominant in a lot of studies, for instance above cited experiments by Pätynen et al., and has led to level normalization before (e.g. concert hall study from the Göttingen-group [130]). These studies strongly encourage analyzing levels and strength in performance venues.

5.6.4 Conclusion

This study with two high-profile concert halls indicated that positions near to the orchestra, and particularly those in Berlin Konzerthaus, produce prominent subjective impact. When eliminating the level differences between stimuli, the order is not affected since other perceptual cues take over. Impact as a measure of the emotional effect on the listener is affected greatly but not solely by level. Differences in dynamic range were found but might not be the sole decisive criterion. Using a real orchestra, similar results are found as in a study conducted with an artificial orchestra, suggesting the validity of both approaches.

5.7 Discussion and conclusion

In this chapter connections between acoustics and dynamics were discussed. Several studies were conducted to investigate this relationship, an area only recently found to be of interest. To visualize the changes sound pressure level histograms were employed.

An introduction to the topic was made and the different definitions of dynamics explained for the domains music, audio recording and acoustics. Several comparisons were conducted, firstly showing how the dynamic range (or range of levels) is affected by late reverberation. Secondly, it was observed that the decrease of dynamic range due to reverberation depends strongly on the signal. Thirdly, increasing distance does not simply decrease the dynamic range as might be expected but depends on the directivity of the instruments. Elevated seats with less obstruction for direct sound and early reflections had a larger dynamic range. The findings are in line with current knowledge from psychacoustics and signal theory. As similar studies have not been discussed in room acoustic research, with the exception for basic observations by Meyer, there are few experiments for comparison. The relevance of late energy for the perception of dynamics might have been overlooked in the two recent studies related to dynamics by

Pätynen: In both studies [114, Tab.2+3], [120, Tab.3] there were noticeably more and higher correlation values for the parameter that includes late lateral energy than the parameter for early energy, which was hypothesized to be more important. The higher correlations especially occurred at increased receiver distances.

Subsequently, a case study was conducted comparing Bayreuth as the original venue for Wagner's operas and an "average" medium-sized theater. By analyzing audio signals for the same piece, it was observed that the level range differs noticeably between the two venues. It was argued that this change presents a major difference regarding sound aesthetics and thus performance practice. Despite only small differences in average level, the theater appeared louder in *forte* parts, resulting in an extended dynamic range. It is suggested that this is not desirable in this case. Studies connecting or discussing performance practice and acoustics of historical venues have been conducted for Beethoven (Weinzierl, [2]) and Haydn (Meyer, [17, pp. 166, 267-269]), but without analyzing music signals or signal dynamics. Overall, the venues where classical and romantic pieces premiered were smaller and had somewhat shorter reverberation times. Thus, a larger dynamic range will have appeared than in today's venues. Both authors also pointed out that composers might not have preferred the acoustics of the given venues and therefore, as documented through letters, adapted a composition. In the case of Bayreuth it can be assumed that the venue is very close to the ideal of the composer Richard Wagner.

It might be argued that the observed changes can be primarily attributed to the slightly different sound sources in each venue. However, by employing a simulated orchestra in several virtual venues it was shown that the *acoustics* change the signal envelope considerably. The smaller the room, the less the absorption, the smaller the dynamic range. To get a better understanding about typical sound levels and dynamic ranges in concert venues a large corpus of 60 hours of binaural in situ audio recordings was analyzed. Finally, we confirmed previous findings by Pätynen et al. that rectangular halls create a stronger emotional impact. Loudness normalization did not change the outcomes considerably. Then however, it was not possible to attribute this effect to differences in dynamic range. The feeling of proximity and perception of bass were more important.

The idea that the room influences the signal dynamics might be unfamiliar as the room is considered to be a linear system. However, with the combination of the source signal characteristics and the perceptually different influences of early and late energy this effect is present and explainable. It remains to be seen if it is one of the missing links in explaining concert hall preference.

6. Summary and general discussion

Acoustic research has established a well-founded theory for reverberation over the last decades. By comparing perceptual data to physical measurements, we have seen progress in prediction of acoustical quality of concert venues. However, until now the standardized physical measurement parameters only explain portions of the perceptually relevant effects. This fact becomes obvious, for example, when new concert halls are successes or failures even though the *objective*, measurable parameters suggest otherwise.

Partly, the deficiency is due to the fact that the established experimental methods have properties that influence the listener in one way or another. In this thesis, a combined approach of in situ and laboratory testing was applied to circumvent this issue. A lecture hall and a concert hall equipped with an electronic acoustics system were used as experimental environments, effectively combining a real and an artificial sound field. This semi virtual condition offers experimental freedom while retaining a realistic scenario. A substantial part of this thesis was therefore conducted utilizing electronic acoustics systems or *room enhancement*. Room enhancement systems are normally utilized to improve the acoustical quality of a performance venue by providing additional sound reflections played back from loudspeakers. Hereby, the reverberation can be lengthened, making a venue suitable for more than one type of music. Even though a number of installations and different systems existed, there was little documentation or scientific experiments utilizing room enhancement. Our related experiments, not shown in this thesis, also encouraged the usage of enhancement for investigating specific questions where the realistic context seems to play a very important role. First, the role of early ceiling reflections on the perception of proximity could be studied more thoroughly [131], [132]. In an experiment that depends so much on a proper room response (feedback between acoustics and the organist [133]) hardly any other environment than the semi-virtual one could provide a reasonable setting. At the same time, it became clear during experiments for this thesis, that there are some inherent limitations when combining existing and virtual sound fields that can hinder the investigations. These limitations include especially the changes in the reverberation decay depending on the system gain. Together with the fact that the enhanced environment is somewhat more costly and time consuming than a pure laboratory situation, it can serve perfectly for well defined, high-quality experiments in a semi-virtual condition, but not as a complete replacement.

Firstly, we investigated the influence of reverberation level on the related perception *reverberance*. A listening experiment was conducted in the enhanced medium-sized concert hall. Preferred reverberation was asked for while varying the contribution of the enhancement. It was observed that a small influence was preferred over the real situation. The results suggest that only a well set system provides the desired positive effects. When the enhancement is exaggerated, negative effects occur, namely coloration or artificiality. Also, it did not seem easily possible to keep the reverberation time constant while changing the reverberation gain. This claim was made by some system engineers. We believe this is not possible in an enhanced sound field due to the overlap of real and enhanced reverberation.

Subsequently, we performed two experiments in the enhanced lecture hall to investigate the relationship between reverberation level and decay. In the first experiment, stimuli of varying reverberation level and time were offered and the reverberance of each was quantified. In the second experiment, equal reverberance had to be set for stimuli of different reverberation levels and times. It was observed that reverberance was more consistently predicted with the energy parameters than the decay parameters. Thus both level and decay are needed to properly describe the perception *reverberance*. Two solutions are suggested: a combination or decay time such as EDT and energy parameter strength G. The author assumes that this connection would have also provided good correlation in the study by Lee [75] or the follow up study by Jian [77]. Secondly, a loudness-based reverberation analysis was investigated, motivated by said studies. While the approach works in principal, reverberance prediction was not noticeably improved here. The results were not worse than conventional decay times but neither were they noticeably better. Also, estimates from the three loudness models DLM, TVL and ISO532-1:2016 differed among each other, which makes standardized evaluation difficult. It is arguable if the additional practical and computational effort of incorporating a loudness or auditory model in general room acoustics would be worth the effort based on these results as the model or procedure would also have to be standardized.

The loudness-based impulse response analysis approach seemed promising and attractive, because as an intermediate, it incorporates both well established techniques from room acoustics as well as some auditory inspired features by applying an auditory model for instationary loudness. In another study loudness analysis of impulse responses did also serve well for predicting disturbing echos [72]. Further evaluation would thus certainly be worth doing, for real-life measurements that are backed with listening experiment data, such as the concert hall measurements corpus by Lokki et al.

The second aspect, direction of reverberation, was investigated mainly because of conflicting evidence regarding the related percept *listener envelopment* ("LEV"). The standardized measure only accounts for late energy arriving from lateral directions whereas other researchers suggested rear reverberation or other room directions to be important as well. In particular for this aspect, most previous experiments had been done in simplified laboratory envelopment with a limited number of loudspeakers. The approach here was to spatially distribute late energy in the enhanced lecture and concert hall as well as two re-synthesized large venues, thus being more realistic. The investigations were fairly straightforward. Results are conclusive and well arguable. It was found that envelopment can be influenced by late energy directions other than lateral. While lateral reverberation is dominant, late energy from behind and also the ceiling were observed to contribute to the feeling of envelopment as well. The overall late energy was the most important factor for listener envelopment, a fact not predicted by correlation-based parameters such as IACC. These results confirm previous findings and provide some new aspects.

The previous studies motivating the investigations presented here were well founded. However, the remaining discrepancy is somewhat exemplary for the research field with a notion gathered from a limited number of simplified experiments being generalized and standardized. That is to say that previous research biases hinder the state of mind. However, the time limit used to separate early from late energy was fixed in order not to introduce another variable. It should be emphasized that it is a strong possibility and has been shown, that other frequency-dependent time limits might be better suited.

Thirdly, the aspect of levels and dynamics in concert halls has been studied for the influence of reverberation. In room acoustics, analyzing the signal at the listeners ear is not very common. Accordingly, the two previous chapters of the thesis analyzed impulse responses measures, not the signals. Likewise, the discussion of dynamics in concert halls is at its beginning. Only recently it was found that this might be a major differences between concert halls. Thus, the relevance of the effects observed in this work is not entirely clear. Here, an approach was taken to analyze the measurable level range of the signal as an evident parameter, in order to collect experiences that might be relevant for discussion of dynamics. This signal-based approach showed the physical influences of reverberation on the dynamic range in various scenarios. Audio recordings from real, enhanced and simulated venues were gathered for analysis of signal envelopes and further comparison to acoustical properties. The observations include that the dynamic range is decreased by reverberation and that this affects the overall musical dynamic. A case study was shown for Richard Wagner's Lohengrin from two acoustically distinct venues resulting in a change of dynamic and expressivity. It could be argued that some of the investigated effects are self-evident, but it is already a more sophisticated approach to evaluate the signal that is actually reaching the listener than remaining in the system domain.

State of the art for understanding reverberation has been studied and developed further. It appears that the standardized room acoustic parameters should be optimized to explain perceptually more complex situations. With repeated assessment, selection and refinement of the conventional parameters, and increasing consideration of source and auditory parameters, a growing understanding and explanation of room acoustic perception is achieved. This time-consuming procedure is in progress and advanced in the present work. Overall, methods related to perception do appear as more promising and have potential to reveal new insight. This outlook is especially the case where psychoacoustic techniques are used, that mean to represent the perceptual events very closely. For instance, concepts such as discrimination of direct/foreground and reverberant/background streams are utilized successfully.

Appendix

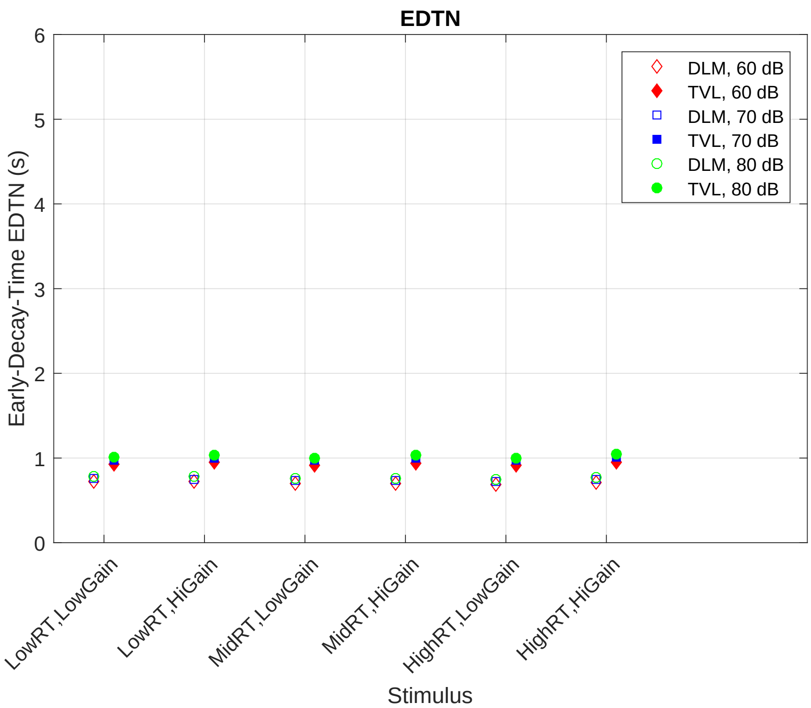


Figure 6.1. Loudness-based early decay time (EDT_N) for Loudness models DLM (left) and TVL (right, filled) for different input levels **Experiment 2**

Table 6.1. Conventional and extended ISO-Impulse Response Parameters (Averaged 125 - 4000 Hz).

	T_{30} [s]	EDT [s]	G_{inf} [dB]	G_{80} [dB]	G_{late} [dB]	C_{80} [dB]	C_{50} [dB]	C_5 [dB]
Real room	0.70	0.65	4.0	3.5	-6.2	9.7	5.5	-1.5
Reference	1.30	1.50	10.6	6.9	7.6	-0.6	-3.2	-10.5
Reverb 1	1.55	1.70	9.6	6.1	6.4	-0.4	-3.2	-9.1
Reverb 3	1.75	1.80	9.2	5.8	5.9	-0.1	-2.8	-8.6
Reverb 4	1.30	1.45	11.1	7.4	8.3	-0.9	-3.9	-10.7
Std. Dev.	0.18	0.14	0.75	0.64	0.94	0.30	0.41	0.90

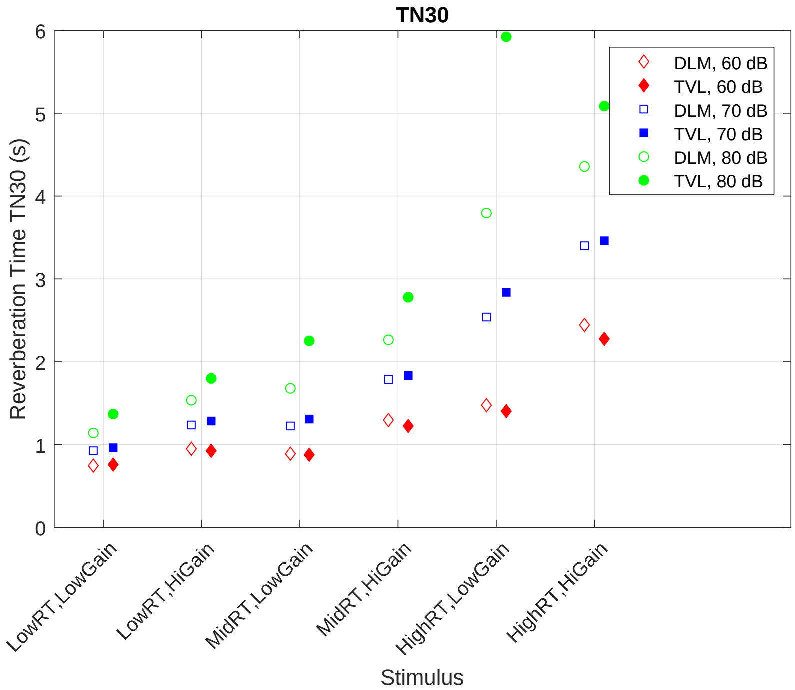


Figure 6.2. Loudness-based Reverberation Time (T_{N30}) for Loudness models DLM (left) and TVL (right, filled) for different input levels **Experiment 2**

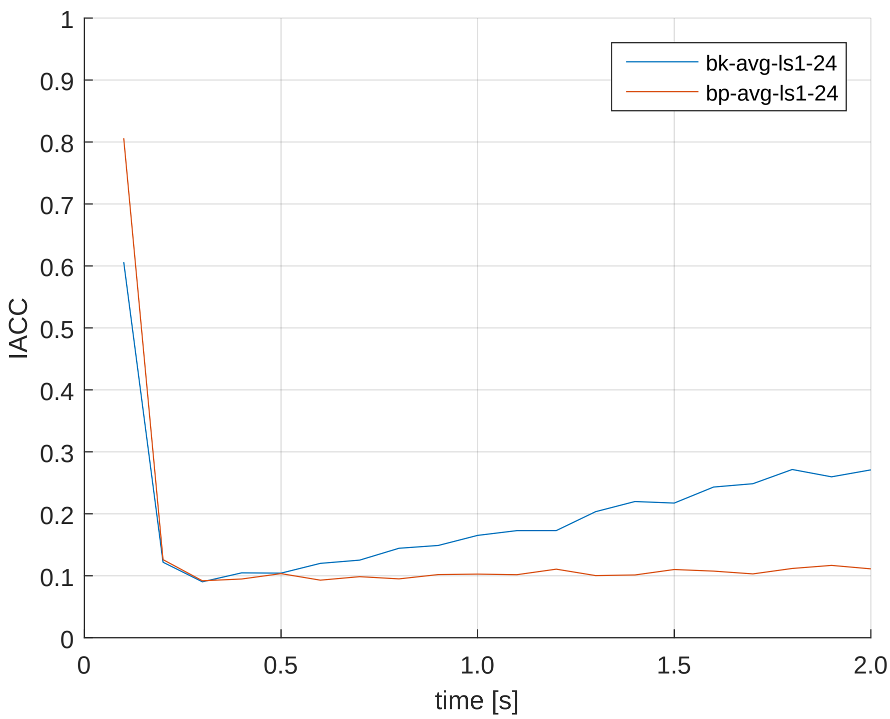


Figure 6.3. IACC over time in Berlin Konzerthaus (bk) and Berlin Philharmonie (bp), unfiltered, for one receiver and 24 loudspeakers (averaged).

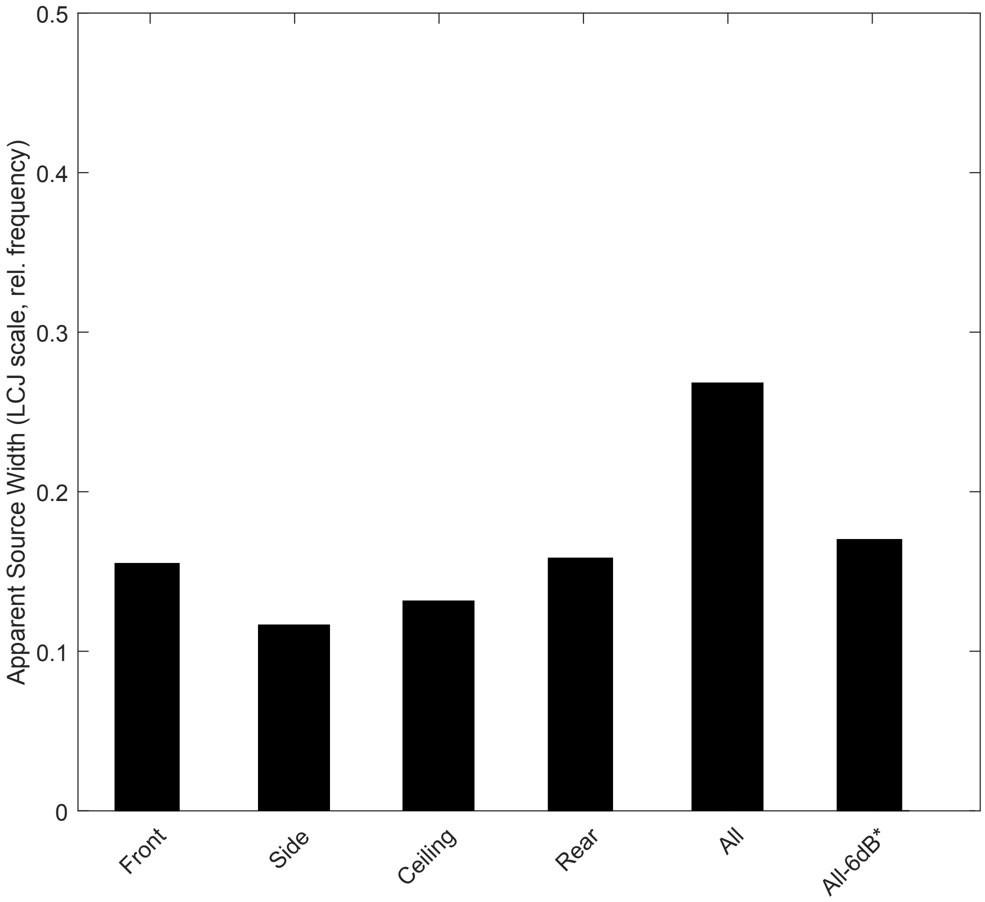


Figure 6.4. Apparent Source Width (ASW) for different directions of late energy.

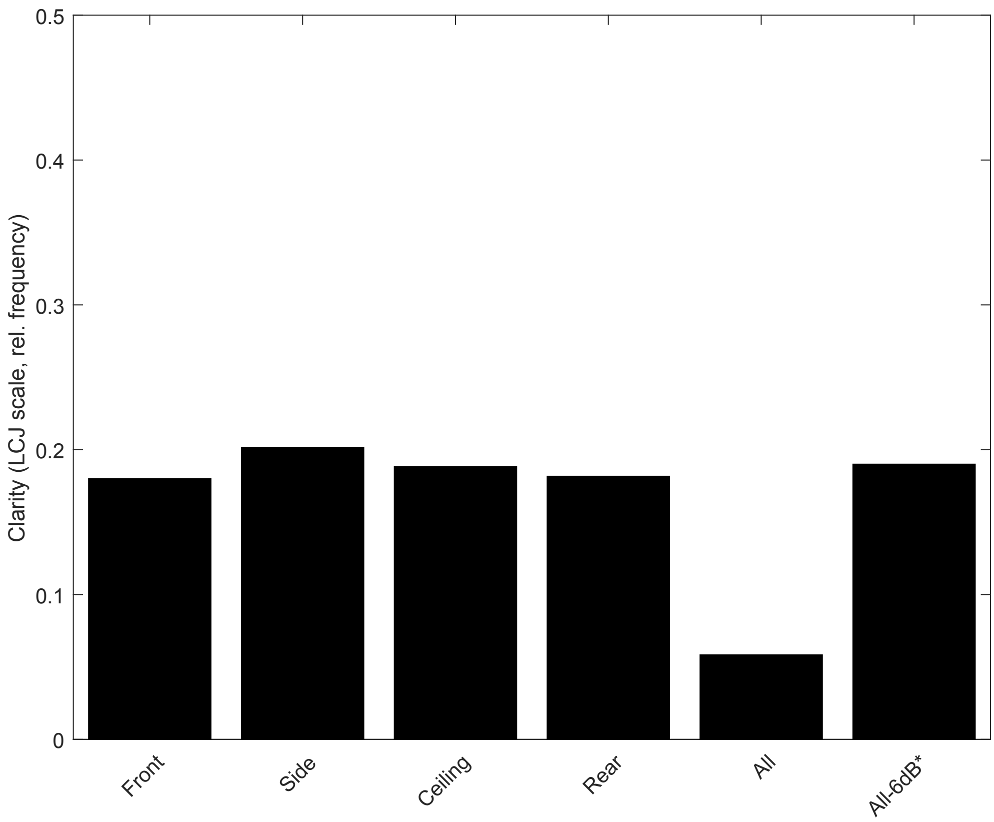


Figure 6.5. Test 2: Perceived Clarity of Sound for different directions of late energy.

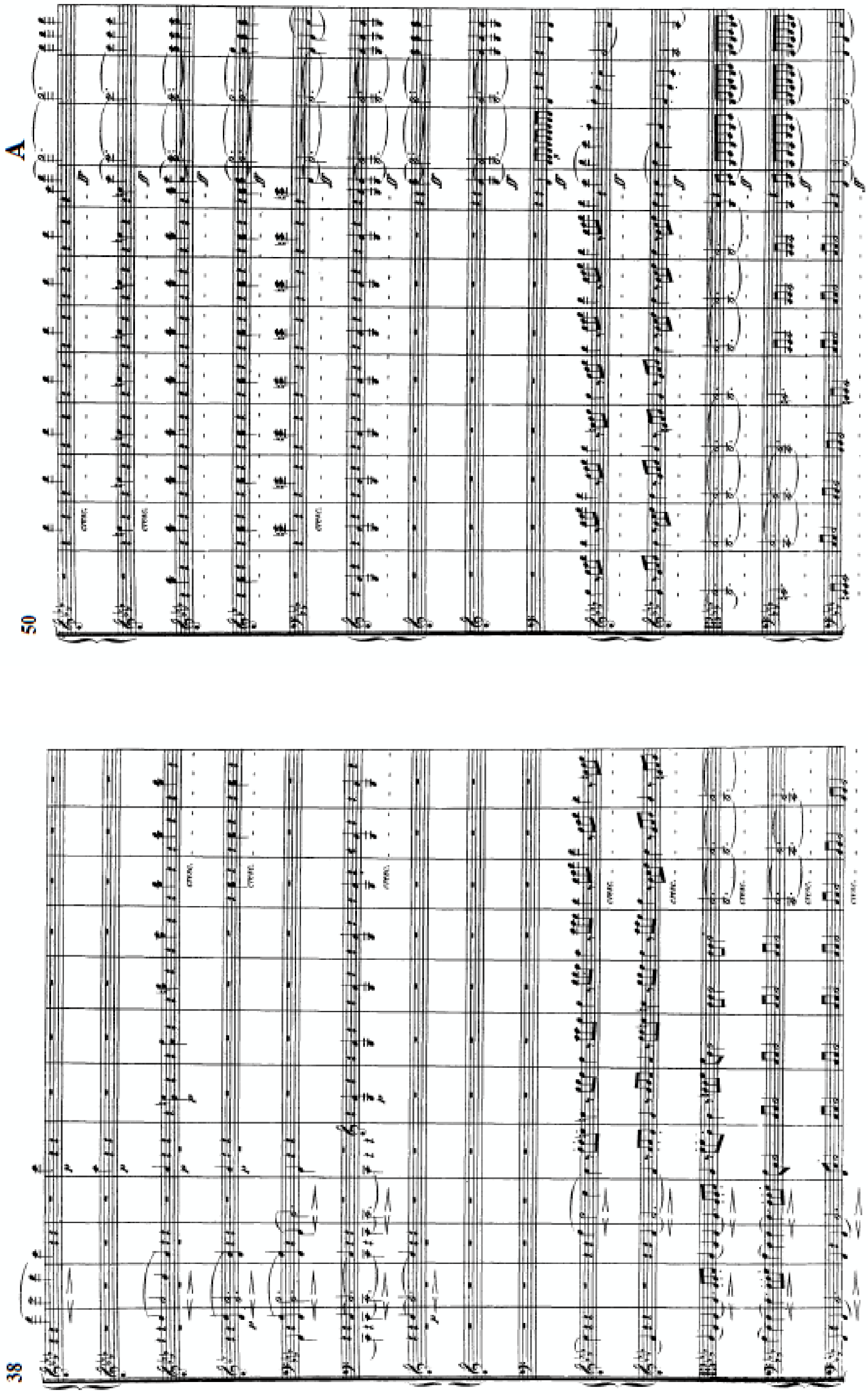
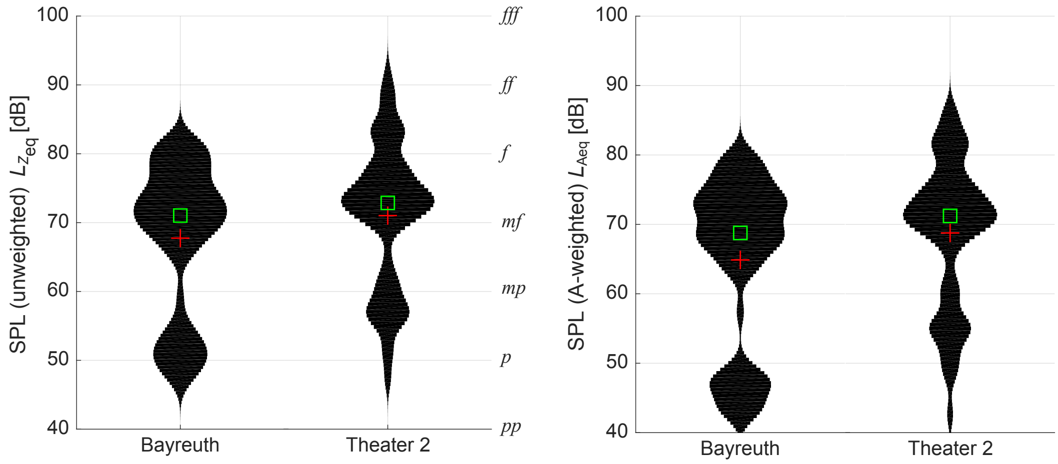


Figure 6.6. Score Excerpt Beethovens Egmont Overture, as tested in Section 5.6.



(a) no weighting (linear)

(b) A-weighted

Figure 6.7. Comparison of level histograms without and with frequency weighting for a 59 seconds excerpt of Lohengrin.

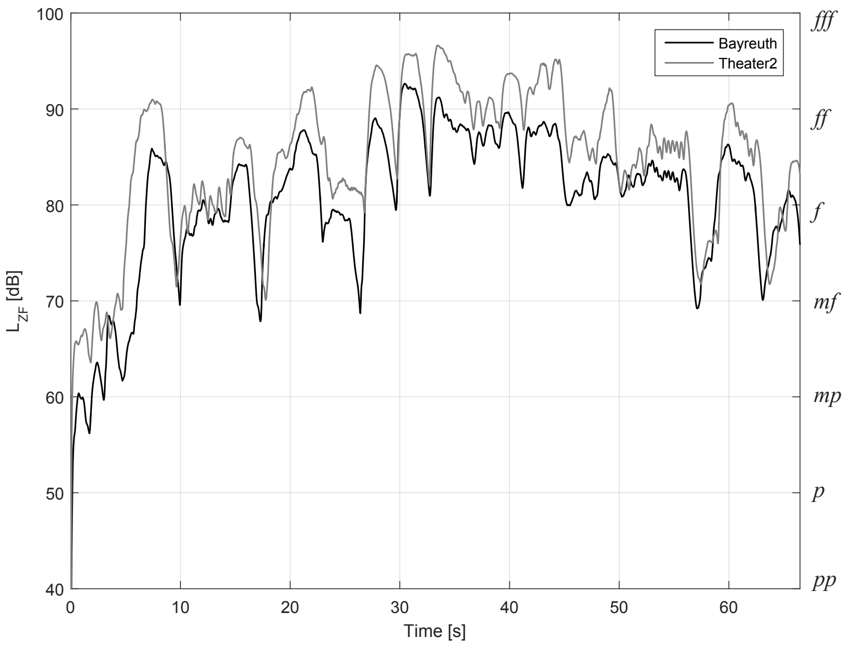
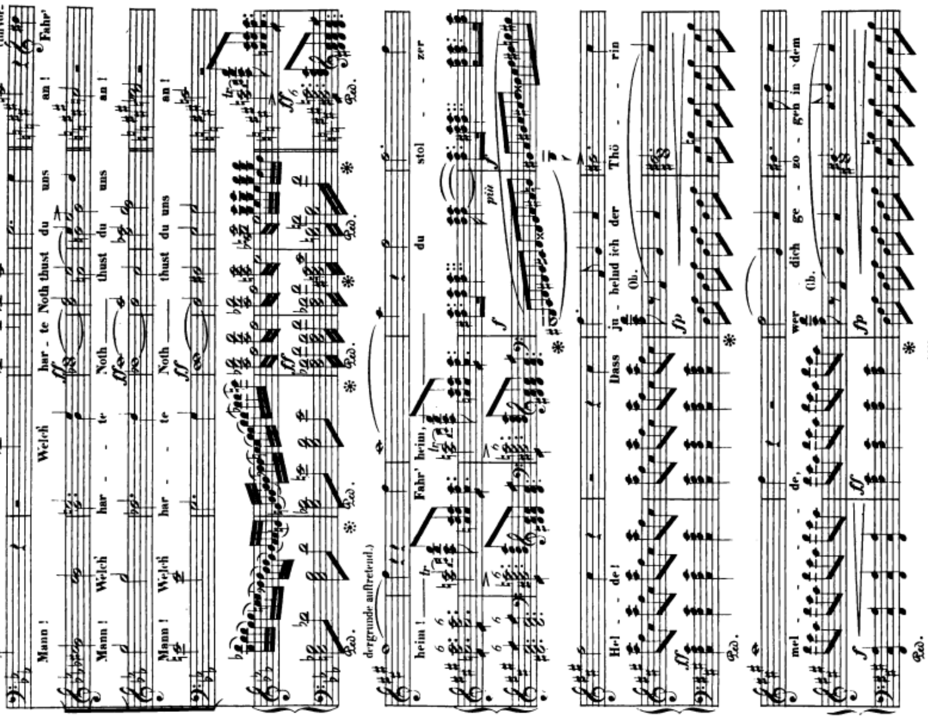


Figure 6.8. Comparison of SPL over time for Bayreuth and Theater 2, Lohengrin excerpt 2 (unweighted).



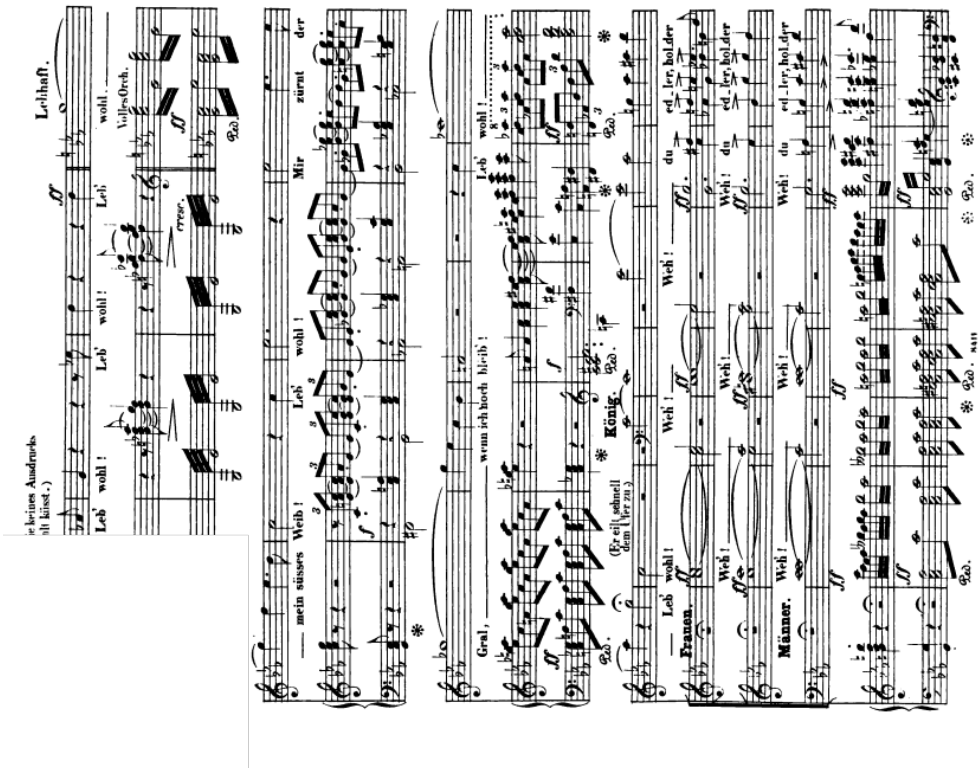


Figure 6.9. Score Lohengrin Excerpt 2.

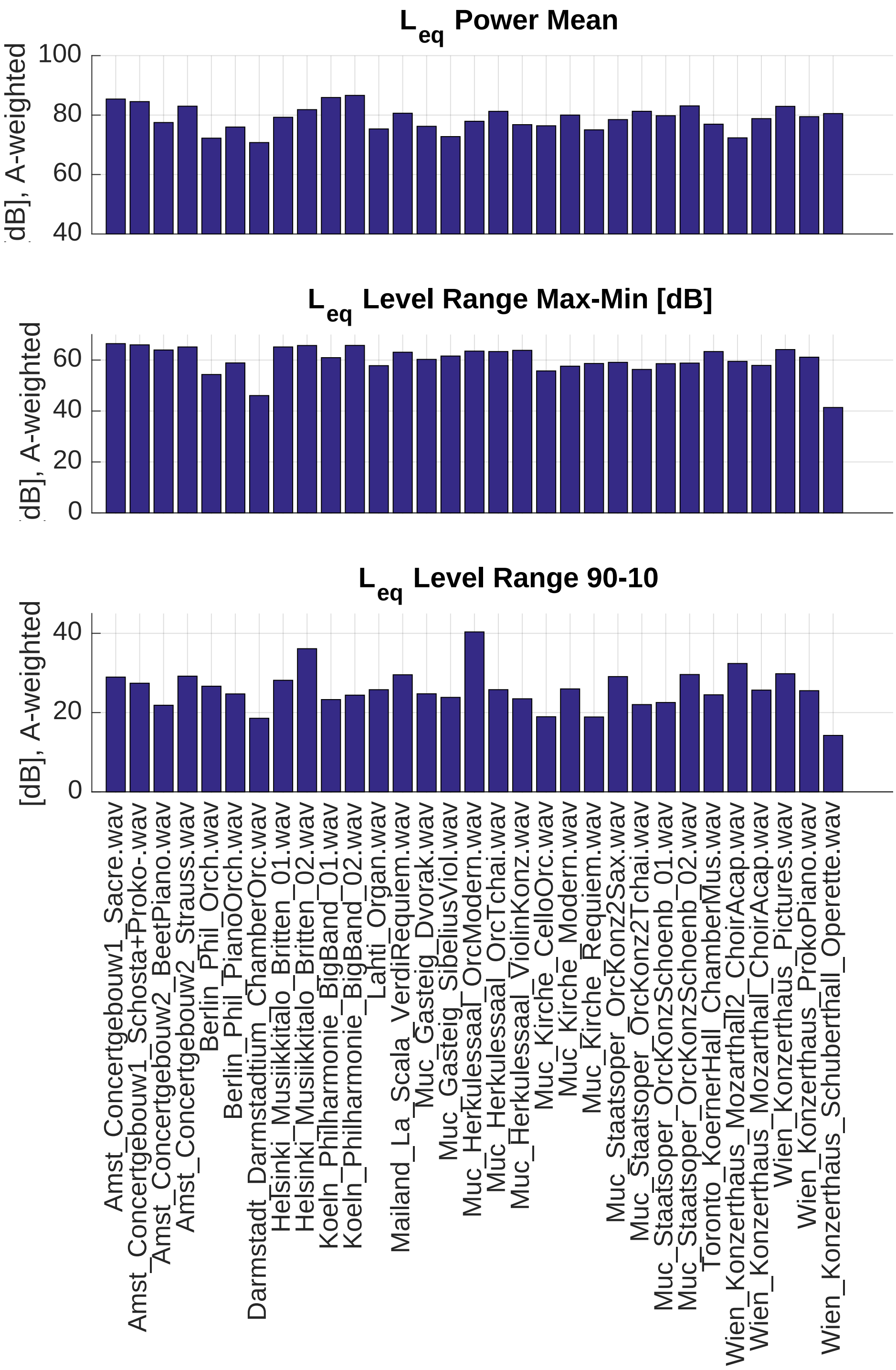


Figure 6.10. A-weighted sound pressure level and level ranges, Concert performances. The last stimulus was excluded from further analysis.

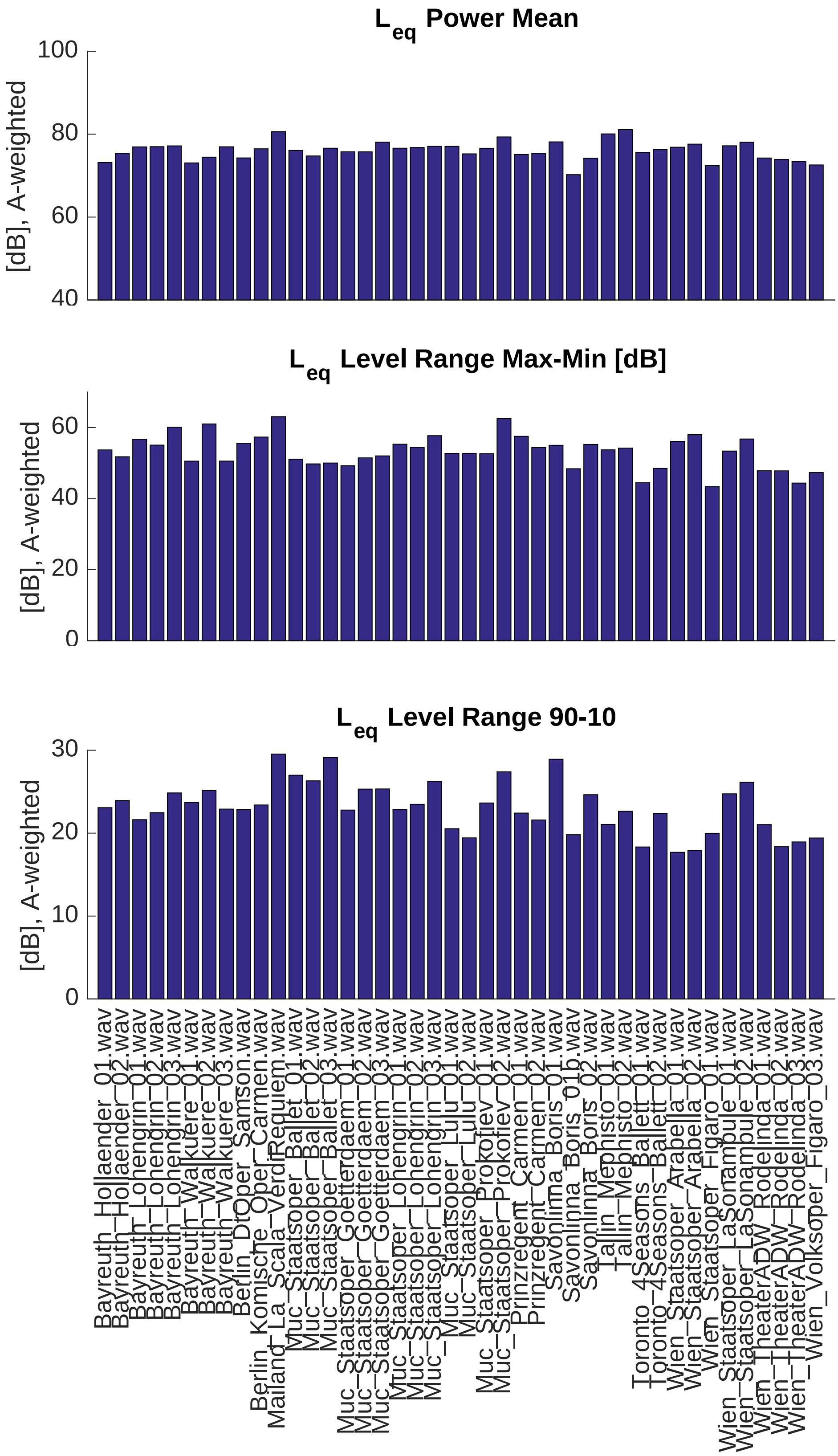


Figure 6.11. A-weighted sound pressure level and level ranges, Opera performances.

Bibliography

[1] A. Kircher, Neue Hall- und Thonkunst oder Mechanische Gehaim-Verbindung der Kunst und Natur durch Stimme und Hall-Wissenschafft gestifftet. Nordlingen: Heylen, Reprint 1983 ed., 1685. Th. Schäfer GmbH, Hannover.

[2] S. Weinzierl, Beethovens Konzerträume: Raumakustik und symphonische Aufführungspraxis an der Schwelle zum modernen Konzertwesen. Frankfurt am Main: E. Bochinsky, 2002.

[3] H. Möller, Europäische Musikgeschichte, Bd. 1, ch. 5: Die modernen Musiker des 14. Jahrhunderts, pp. 146–148. Kassel: Bärenreiter, 2002.

[4] J. L. Clarke, "Catacoustic enchantment: The romantic conception of reverberation," in Grey Room, Inc. and Massachusetts Institute of Technology, vol. 60, p. 36–65, 2015.

[5] H. Kuttruff, Room Acoustics. London: Spon Press, 4th ed., 2000.

[6] DIN EN ISO 3382-1:2009-10, "Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces," October 2009.

[7] E. Mommertz, Akustik und Schallschutz. Detail Praxis, Edition Detail, München: Institut für internationale Architektur-Dokumentation, 2008.

[8] J. Blauert, Spatial Hearing: The Psychophysics of Human Sound Localization (Revised Edition). Cambridge: The MIT Press, 2nd ed., 1999.

[9] P. Lehman and H. Wilkens, "Zusammenhang subjektiver Beurteilung von Konzertsälen mit raumakustischen Kriterien," Acustica, vol. 45, pp. 256–268, 1980.

[10] M. Barron and L.-J. Lee, "Energy relations in concert auditoriums," Journal of the Acoustical Society of America, vol. 84, no. 2, pp. 618–628, 1988.

[11] T. Lokki, J. Patynen, A. Kuusinen, and S. Tervo, "Disentangling preference ratings of concert hall acoustics using subjective sensory profiles," Journal of the Acoustical Society of America, vol. 132, no. 5, pp. 3148–3161, 2012.

[12] T. Lokki, J. Pätynen, A. Kuusinen, and S. Tervo, "Concert hall acoustics: Repertoire, listening position, and individual taste of the listeners influence the qualitative attributes and preferences," Journal of the Acoustical Society of America, vol. 140, no. 1, pp. 551–562, 2016.

[13] L. Cremer and H. A. Müller, Die wissenschaftlichen Grundlagen der Raumakustik. Stuttgart: S. Hirzel Verlag, 2nd ed., 1978.

[14] N. Kaplanis, S. Bech, S. H. Jensen, and T. van Waterschoot, "Perception of reverberation in small rooms: A literature study," in 55th International Conference of the Audio Engineering Society, 2014. paper no. P-3, Helsinki, Finland, 2014 August 27-29.

[15] A. Lindau, V. Erbes, S. Lepa, and S. Weinzierl, "A spatial audio quality inventory (SAQI)," Acta Acustica united with Acustica, vol. 100, no. 5, pp. 984–994, 2014.

[16] S. Weinzierl, A. Lindau, and S. Lepa, "Ein psychologisches Messinstrument für die akustische Wahrnehmung von Räumen für Musik und Sprache I: Fokusgruppe und Konsensvokabular," in 42nd Convention of the German Acoustical Association (DAGA), 2016. Aachen, Germany, 2016 March 14-17.

[17] J. Meyer, Akustik und musikalische Aufführungspraxis: Leitfaden für Akustiker, Tonmeister, Musiker, Instrumentenbauer und Architekten, vol. 6. Bergkirchen: PPV Medien GmbH, edition bochinsky ed., 2015.

[18] T. Lokki, "Sensory evaluation of concert hall acoustics," in Proceedings of Meetings on Acoustics (ICA), vol. 19, 2013. paper no. 4aID1, Montreal, Canada, 2013 June 2-7.

[19] L. L. Beranek, Concert Halls and Opera Houses: Music, Acoustics and Architecture. New York: Springer, 2004.

[20] A. Lindau, Binaural Resynthesis of Acoustical Environments.Technology and Perceptual Evaluation. PhD thesis, Technische Universität Berlin, Germany, 2014.

[21] M. Vorländer, D. Schröder, S. Pelzer, and F. Wefers, "Virtual reality for architectural acoustics," Journal of Building Performance Simulation, vol. 8, no. 1, pp. 15–25, 2015.

[22] M. Aretz, Combined Wave and Ray Based Room Acoustic Simulations of Small Rooms. PhD thesis, RWTH Aachen, Germany, 2012.

[23] J. Pätynen, S. Tervo, and T. Lokki, "Amplitude panning decreases spectral brightness with concert hall auralizations," in Proc. 55th Audio Eng. Soc. conference, 2014. paper no. p-13, Helsinki, Finland, 2014 August 27-29.

[24] J. Pätynen, A Virtual Symphony Orchestra for Studies on Concert Hall Acoustics. PhD thesis, Aalto University, Finland, 2011.

[25] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," Journal of the Audio Engineering Society, vol. 61, no. 1/2, pp. 17–28, 2013.

[26] A. Kuusinen, J. Pätynen, S. Tervo, and T. Lokki, "Relationships between preference ratings, sensory profiles, and acoustical measurements in concert halls," Journal of the Acoustical Society of America, vol. 135, pp. 239–250, 2014.

[27] S. Tervo, P. Laukkanen, J. Paetynen, and T. Lokki, "Preferences of critical listening environments among sound engineers," Journal of the Audio Engineering Society, vol. 62, no. 5, 2014.

[28] J. Blauert, "A perceptionist's view on psychoacoustics," in Archives of Acoustics, vol. 37, pp. 365–371, 2012.

[29] M. Kob, A. Meyer, C. Kuhn-Rahloff, and G. Kubanek, "Konzept, Umsetzung und musikalische Anwendung der Wellenfeldsynthese an der Hochschule für Musik Detmold," in Proceedings of the 25th VDT International Convention (TMT), pp. 779–783, 2008.

[30] M. Kob, M. Kieslich, S. Antonin, D. Grigoriev, and G. Weinberger, "Realisation of an antiphonal division of a concert organ using WFS," in the International Conference on Spatial Audio 2011 (ICSA), 2011. Detmold, Germany, 2011 November 10-11.

[31] R. Bakker and S. Gillian, "The history of active acoustic enhancement systems," in Proceedings of the Institute of Acoustics, vol. 36, pp. 56–65, 2014.

[32] G. Engel, "Acoustic enhancement systems for classical concerts - examples and experiences," in Proceedings of the Institute of Acoustics, vol. 33, pp. 185–188, 2011.

[33] T. Watanabe and M. Ikeda, "Various applications of active field control," in the 134th Convention of the Audio Engineering Society, 2013. paper no. 8859, Rome, Italy, 2013 May 4-7.

[34] D. Griesinger, "The importance of the direct to reverberant ratio in the perception of distance, localization, clarity, and envelopment," in Proceedings of the 126th Convention of the Audio Engineering Society, vol. 1, pp. 348–361, 2009.

[35] D. Griesinger, "Concert audience engagement - the perception of closeness," in Meeting of the Audio Engineering Society, Washington DC Section, 2010. 2010 April 21.

[36] D. Griesinger, "Improving room acoustics through time-variant synthetic reverberation," in the 90th Convention of the Audio Engineering Society, 1991. Paris, France, 1991 February 19-22.

[37] R. Bakker, "Design and tuning of active acoustic enhancement systems," in Proceedings of the Institute of Acoustics, vol. 36, pp. 76–82, 2014.

[38] M. Klatte, T. Lachmann, and M. Meis, "Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting," Noise & Health, vol. 12, pp. 270–282, 2010.

[39] T. Watanabe, M. Ikeda, and S. Kim, "Active field control using sound field generation techniques case study of a live concert at a virtual renaissance church," in Proceedings of Meetings on Acoustics (ICA), vol. 19, 2013. paper no. 3aAAa2, Montreal, Canada, 2013 2-7 June.

[40] K. Brawata, P. Malecki, A. Pilch, and T. Kamisinski, "Subjective assessment of commercial sound enhancement system," in the 138th Convention of the Audio Engineering Society, 2015. paper no. 9306, Warsaw, Poland, 2015 May 7-10.

[41] A.-R. Tereping, "Listener preference for concert sound levels: Do louder performances sound better?," Journal of the Audio Engineering Society, Engineering Reports, vol. 64, no. 3, pp. 138–146, 2016.

[42] DIN 18041:2004-05, "Hörsamkeit in kleinen bis mittelgroßen Räumen," May 2004.

[43] W. Lachenmayr, "Höreindruck und raumakustische Kriterien in Konzertsälen," Master's thesis, Institut für Komposition und Elektroakustik. Universität für Musik und darstellende Kunst Wien, 2012.

[44] P. Martignon, A. Azzali, D. Cabrera, A. Capra, and A. Farina, "Reproduction of auditorium spatial impression with binaural and stereophonic sound systems," in the 118th Convention of the Audio Engineering Society, 2005. paper no. 6485, Barcelona, Spain, 2005 May 28-31.

[45] D. Protheroe and C. Day, “Validation of lateral fraction results in room acoustic measurements,” in 43rd International Congress on Noise Control Engineering, 2014. paper no. 425, Melbourne, Australia, 2014 November 16-19.

[46] E. Vickers, “Metrics for quantifying loudness and dynamics.” online article, http://www.sfxmachine.com/docs/loudnesswar/metrics_for_quantifying_loudness_and_dynamics.pdf , accessed 2016-12-27, 2010.

[47] J. Rennies, J. L. Verhey, and H. Fastl, “Comparison of loudness models for time-varying sounds,” Acta Acustica united with Acustica, vol. 96, pp. 383–396, 2010.

[48] G. A. Soulodre, “Evaluation of objective measures of loudness,” in Canadian Acoustics, vol. 32, pp. 152–153, 2004.

[49] L. Thibault, “CRC loudness meter chosen as international standard,” Tech. Rep. 10, Communications Reserach Centre Canada, Spring 2009. ISSN 1717-7294.

[50] F. Schmidt, B. Kollmeier, and S. Uppenkamp, “Stärken und Schwächen von Pegelmaßen und Lautheitsmodellen bei der Bewertung von Musik,” in Proceedings of the German Acoustical Association (DAGA), pp. 285–286, 2016. Aachen, Germany, 2016 March 14-17.

[51] EBU, “Tech 3342-2016: Loudness range: A measure to supplement EBU R 128 loudness normalization,” 2016. Geneva, Switzerland, January 2016.

[52] D. Cabrera, D. Jimenez, and W. Martens, “Audio and acoustical response analysis environment (AARAE): a tool to support education and research in acoustics,” in 43rd International Congress on Noise Control Engineering, 2014. paper no. 235, Melbourne, Australia, 2014 November 16-19.

[53] J. Gerich, Handbuch der sozialwissenschaftlichen Datenanalyse, ch. Thurstone-und Likertskalierung, pp. 259–281. Wiesbaden: VS Verlag für Sozialwissenschaften, 2010.

[54] R. A. Bradley and M. E. Terry, “Rank analysis of incomplete block designs: I. the method of paired comparisons.,” Biometrika, no. 39, pp. 324–345, 1952.

[55] R. D. Luce, Individual choice behaviour: A theoretical analysis. New York: Dover Publications, 2005. (Reprint of the John Wiley Sons, Inc., New York, 1959).

[56] F. Wickelmaier and C. Schmidt, “A matlab function to estimate choice model and parameters from paired-comparison data,” Behavior Research Methods, Instruments, & Computers, vol. 36, no. 1, pp. 29–40, 2004.

[57] M. Semenou, “CompR-package for paired comparison data analysis in *R*.” available on CRAN: “Comprehensive R Archive Network”, https://cran.r-project.org/web/packages/CompR/, accessed 2016-12-27, July 2015. Version 1.0.

[58] S. Nakagawa, “A farewell to bonferroni: The problems of low statistical power and publication bias,” Behavioral Ecology, vol. 15, no. 6, pp. 1044–1045, 2004.

[59] R. J. Cabin and R. J. Mitchell, “To bonferroni or not to bonferroni: When and how are the questions,” Bulletin of the Ecological Society of America, vol. 81, no. 3, pp. 246–248, 2000.

[60] J. Bortz, G. A. Lienert, and K. Boehnke, Verteilungsfreie Methoden in der Biostatistik. Berlin: Springer, 2nd ed., 2000.

[61] M. Barron, "Subjective study of british symphony concert halls," Acustica, vol. 66, no. 1, pp. 1–14, 1988.

[62] E. Kahle and J.-P. Jullien, "Some new considerations on the subjective impression of reverberance and its correlation with objective criteria," in Sabine Centennial Meeting of the Acoustical Society of America, pp. 239–242, 1994.

[63] G. A. Soulodre and J. S. Bradley, "Subjective evaluation of new room acoustic measures," Journal of the Acoustical Society of America, vol. 98, no. 1, pp. 294–301, 1995.

[64] J. S. Bradley, "Using ISO3382 measures to evaluate acoustical conditions in concert halls," in the International Symposium on Room Acoustics: Design and Science (ISRA), pp. 1–10, 2004. Hyogo, Japan, 2004 April 11-13.

[65] J. J. Dammerud, "Details on G_{late} on stage versus in the stalls." online, 2011. `https://stageac.files.wordpress.com/2014/10/dammerud_glateindetail_2011.pdf`, accessed 2016-12-27.

[66] A. Haapaniemi and T. Lokki, "The preferred level balance between direct, early, and late sound in concert halls," Psychomusicology: Music, Mind & Brain, vol. 25, no. 3, pp. 306–316, 2015.

[67] M. Morimoto and A. Asaoka, "Multi-dimensional analysis of 'reverberance'," in the 18th International Congress on Acoustics (ICA), no. 4, pp. 2469–2472, 2004. paper no. Th2.B1.6, Kyoto, Japan, 2004 April 4-9.

[68] J. Pätynen, S. Tervo, and T. Lokki, "Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses," Journal of the Acoustical Society of America, vol. 133, no. 2, pp. 842–857, 2013.

[69] M. Barron, Auditorium Acoustics and Architectural Design. London: E & FN Spon, 1st ed., 1993.

[70] D. Lee and D. Cabrera, "Basic considerations for loudness-based analysis of room and impulse responses," Building Acoustics, vol. 16, no. 1, pp. 31–46, 2009.

[71] S. Frey, "Zur Prognose von Echowirkungen in Räumen," Master's thesis, TU Munich, Germany, 2013.

[72] D. Rauber, "Echovorhersage anhand von Lautheitsmodellen," 2015. Bachelor's thesis, TU Munich, Germany.

[73] D. Lee, D. Cabrera, and W. L. Martens, "Equal reverberance matching of music," in the Annual Conference of the Australian Acoustical Society (Acoustics), 2009. paper no. p17, Adelaide, Australia, 2009 November 23-25.

[74] D. Lee, D. Cabrera, and W. L. Martens, "Equal reverberance matching of running musical stimuli having various reverberation times and SPLs," in the 20th International Congress on Acoustics (ICA), 2010. Sydney, Australia, 2010 August 23-27.

[75] D. Lee, D. Cabrera, and W. L. Martens, "The effect of loudness on the reverberance of music: Reverberance prediction using loudness models," Journal of the Acoustical Society of America, vol. 131, no. 2, pp. 1194–1205, 2012.

[76] D. Lee, D. Cabrera, and W. L. Martens, "Accounting for listening level in the prediction of reverberance using early decay time," Acoustics Australia, vol. 40, no. 2, pp. 103–110, 2012.

[77] G. Jiang and L. Wen, “Predicting reverberance in flat rooms using Time Varying Loudness model,” in the 21st International Congress on Sound and Vibration (ICSV), 2014. paper no. 314, Beijing, China, 2014 July 13-17.

[78] J. v. D. Schuitman, Auditory modelling for assessing room acoustics. PhD thesis, TU Delft, Netherlands, 2011.

[79] J. v. D. Schuitman, D. de Vries, and A. Lindau, “Deriving content-specific measures of room acoustic perception using a binaural, nonlinear auditory model,” Journal of the Acoustical Society of America, vol. 133, no. 3, pp. 1572–1585, 2013.

[80] P. Luizard, Les espaces couplés: Comportement, conception et perception dans un contexte de salle de spectacle. PhD thesis, L’Université Pierre et Marie Curie, Paris, France, 2013.

[81] E. Kahle, T. Wulfrank, Y. Jurkiewicz, and N. Faillet, “Philharmonie de Paris – the acoustic brief,” in Proceedings of the Institute of Acoustics, vol. 37, pp. 105–110, 2015.

[82] B. Katz, Y. Jurkiewicz, T. Wulfrank, G. Parseihian, T. .Scélo, and H. Marshall, “La Philharmonie de Paris – acoustic scale model study,” in Proceedings of the Institute of Acoustics, vol. 37, pp. 431–438, 2015.

[83] M. Barron, “Interpretation of early decay times in concert auditoria,” Acustica, vol. 81, pp. 320–331, 1995.

[84] M. Barron, “Non-linear decays in simple spaces and their possible exploitation,” in Proceedings of the Institute of Acoustics, vol. 34, 2011.

[85] J. O’Keefe and D. Ruvalcaba, “Queen Elizabeth Theatre, Vancouver: acoustic design responding to financial realities,” the 8th International Conference on Auditorium Acoustics (IOA), vol. 33, no. 2, pp. 210–218, 2011. Dublin, Ireland, 2011 May 20-22.

[86] D. Griesinger, “Phase coherence as a measure of acoustic quality, part two: Perceiving engagement,” in Proceedings of the 20th International Congress on Acoustics (ICA), 2010. paper no. 603, Sydney, Australia, 2010 August 23-27.

[87] J. Paulus, C. Uhle, J. Herre, and M. Höpfel, “A study on the preferred level of late reverberation in speech and music,” in 60th International Conference of the Audio Engineering Society, 2016. paper no. 1-3, Leuven, Belgium, 2016 February 3-5.

[88] W. Lachenmayr, “Loudness-weighting of reverberation using electronic room enhancement,” in the 60th International Conference of the Audio Engineering Society, 2016. paper no. 3-3, Leuven, Belgium, 2016 February 3-5.

[89] D. Cabrera, S. Ferguson, and E. Schubert, “PsySound3: An integrated environment for the analysis of sound recordings,” in the Australian Acoustical Society Conference (Acoustics 2008), 2008. paper no. 42, Geelong, Australia, 2008 November 24-26.

[90] ISO/DIS 532-1: 2016, “Acoustics - methods for calculating loudness - part 1: Zwicker method,” April 2016.

[91] F. Englberger, “Nachhall in Musikräumen - Implementierung eines Lautheitsmodells nach der Methode von Zwicker mit anschließender Berechnung einer gehörsgerechten Nachhallzeit,” Master’s thesis, TU Munich, Germany, 2016.

[92] A. V. Osses, W. Lachenmayr, E. Mommertz, and A. Kohlrausch, “Predicting the perceived reverberation in different room acoustic environments using a binaural auditory model,” Journal of the Acoustical Society of America Express Letters (conditionally accepted), 2017.

[93] B. De Man and J. D. Reiss, "Perceptual evaluation of reverb in music production," in 60th International Conference of the Audio Engineering Society, 2016. demo 11, Leuven, Belgium, 2016 February 3-5.

[94] D. Lee, E. Gong, D. Cabrera, M. Yadav, and W. L. Martens, "Intelligibility of reverberant speech with amplification: Limitation of speech intelligibility metrics, and a preliminary examination of an alternative approach," Journal of Applied Mathematics and Physics, vol. 3, pp. 177–186, 2015.

[95] DIN 45631/A1:2010, Änderung von DIN45631:1991-03, "Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum – Verfahren nach E. Zwicker – Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche," March 2010.

[96] W. Lachenmayr, A. Haapaniemi, and T. Lokki, "Direction of late reverberation and envelopment in two reproduced Berlin concert halls," in the 140th Convention of the Audio Engineering Society, 2016. paper no. 9503, Paris, France, 2016 June 4-7.

[97] M. Morimoto and K. Iida, "A new physical measure for psychological evaluation of a sound field: Front/back energy ratio as a measure of envelopment," Journal of the Acoustical Society of America, vol. 93, no. 4, p. 2282(A), 1998.

[98] J. S. Bradley and G. A. Soulodre, "The influence of late arriving energy on spatial impression," Journal of the Acoustical Society of America, vol. 97, no. 4, pp. 2264–2271, 1995.

[99] J. S. Bradley and G. A. Soulodre, "Objective measures of listener envelopment," Journal of the Acoustical Society of America, vol. 98, pp. 2590–2597, 1995.

[100] P. Evjen, J. S. Bradley, and S. G. Norcross, "The effect of late reflections from above and behind on listener envelopment," Applied Acoustics, vol. 62, pp. 137–153, 2001.

[101] M. Morimoto, K. Iida, and K. Sakagami, "The role of reflections from behind the listener in spatial impression," Applied Acoustics, vol. 62, pp. 109–124, 2001.

[102] T. Hanyu and S. Kimura, "A new objective measure for evaluation of listener envelopment focusing on the spatial balance of reflections," Applied Acoustics, vol. 62, pp. 155–184, 2001.

[103] G. A. Soulodre, M. C. Lavoie, and S. G. Norcross, "Objective measures of listener envelopment in multichannel surround systems," Journal of the Audio Engineering Society, vol. 51, no. 9, pp. 826–840, 2003.

[104] G. A. Soulodre, "New objective measures of listener envelopment," in the 18th International Congress on Acoustics (ICA), pp. 2465–2468, 2004. paper no. Th2.B1.5, Kyoto, Japan, 2004 April 4-9.

[105] H. Furuya, K. Fujimoto, and A. Wakuda, "Psychological experiments on listener envelopment when both the early-to-late sound level and directional late energy ratios are varied, and consideration of calculated lev in actual halls," Applied Acoustics, vol. 69, pp. 1085–1095, 2008.

[106] L. L. Beranek, "Listener envelopment LEV, strength G and reverberation time RT in concert halls," in the 20th International Congress on Acoustics (ICA), 2010. paper no. 296, Sydney, Australia, 2010 August 23-27.

[107] E. Kahle, C. Mulder, and T. Wulfrank, "Perceptual relevance of location of reverberation in a concert hall," Psychomusicology: Music, Mind & Brain, vol. 25, no. 3, pp. 326–330, 2015.

[108] M. Morimoto, M. Jinya, K. Nakagawa, and K. Sakagami, "Effects of frequency characteristics of reverberation time on listener envelopment," in the 3rd European Congress on Acoustics (Forum Acusticum), 2002. paper no. RBA-02-001-IP, Sevilla, Spain, 2002 September 16-20.

[109] Y. Ando, "Concert hall acoustics based on subjective theory," in Handbook of Acoustics (T. D.Rossing, ed.), ch. 10, pp. 351–384, Springer, 2007.

[110] K.-H. Lorenz-Kierakiewitz, "Werte binauraler Parameter in Europäischen Konzertsälen," in Proceedings of the German Acoustical Association (DAGA), pp. 343–346, 2015. Nuremberg, Germany, 2015 March 16-19.

[111] J. Y. Jeon and J. You, "Effect of sound strength and IACC on perception of listener envelopment in concert halls," in Proceedings of the 20th International Congress on Acoustics (ICA), 2010. paper no. 905, Sydney, Australia, 2010 August 23-27.

[112] S. Klockgether and S. van de Par, "Just noticeable differences of spatial cues in echoic and anechoic acoustical environments," Journal of the Acoustical Society of America Express Letters, vol. 140, no. 4, 2016. EL352.

[113] T. Lokki, "Tasting music like wine: Sensory evaluation of concert halls," Physics Today, vol. 67, no. 1, pp. 27–32, 2014.

[114] J. Pätynen and T. Lokki, "Perception of music dynamics in concert halls," Journal of the Acoustical Society of America, vol. 140, no. 5, pp. 3787–3798, 2016.

[115] F. Hirsch, Das große Wörterbuch der Musik. Weyarn: Seehamer Verlag, 1996.

[116] J. Meyer, Kirchenakustik. Frankfurt: Edition Bochinsky, 2003.

[117] B. Katz, Mastering Audio - the Art and the Science, ch. How to Manipulate Dynamic Range for Fun and Profit, Part 1: Macrodynamics, pp. 113–118. London: Focal Press, 2nd ed., 2007.

[118] R. Wettschurek, Über die Abhängigkeit raumakustischer Wahrnehmungen von der Lautstärke. PhD thesis, TU Berlin, Germany, 1976.

[119] J. Pätynen, S. Tervo, P. W. Robinson, and T. Lokki, "Concert halls with strong lateral reflections enhance musical dynamics," Proceedings of the National Academy of Sciences, vol. 111, no. 12, pp. 4409–4414, 2014.

[120] J. Pätynen and T. Lokki, "Concert halls with strong and lateral sound increase the emotional impact of orchestra music," Journal of the Acoustical Society of America, vol. 139, no. 3, pp. 1214–1224, 2016.

[121] T. Lokki, "Why do some concert halls render music and more expressive and impressive than others?," in the International Symposium on Musical and Room Acoustics (ISMRA), 2016. paper no. 72, Buenos Aires, Argentina, 2016 September 11-13.

[122] L. L. Beranek, "Concert hall acoustics - 2008," Journal of the Audio Engineering Society, vol. 56, no. 7/8, pp. 532–544, 2008.

[123] W. Lachenmayr and G. Engel, "Comparing room acoustics for the performance of Wagners Lohengrin," in the 138th Convention of the Audio Engineering Society, 2015. e-Brief 191, Warsaw, Poland, 2015 May 7-10.

[124] V. Scory-Engels, "Pressemitteilung Deutscher Buehnenverein, Werkstatistik 2011/2012." online, accessed 2016-12-27, August 2013. `http://www.buehnenverein.de/de/presse/pressemeldungen.html?det=359`.

[125] R. Wagner, Das Bühnenfestspielhaus zu Bayreuth: nebst einem Berichte über die Grundsteinlegung desselben. Leipzig: Fritzsch, 1873. excerpt online: http://www.bayreuther-festspiele.de/dokumente/das_b%$FChnenfestspielhaus_zu_bayreuth_112.html, accessed 2016-12-27.

[126] K. Müller, Das Richard Wagner Festspielhaus Bayreuth, ch. 4: the democratization of acoustics, pp. 174–177. Cologne: nettpress, 2007.

[127] C. L. Fog and R. M. Ballinger, "A new symphic hall, Musikhuset Aarhus, Denmark," in Proceedings of the Acoustics'08 Conference, Paris, pp. 363–368, 2008.

[128] J. H. Rindel, "Odeon application note - auralisation and how to calibrate the sound level for presentations," December 2015. http://www.odeon.dk/pdf/APPLICATION_NOTE_Auralisation_calibration.pdf, accessed 2016-12-27.

[129] W. Lachenmayr and J. Pätynen, "Influence of acoustics on emotional impact of music in Konzerthaus and Philharmonie Berlin," in Proceedings of the German Acoustical Association (DAGA), pp. 297–300, 2016. Aachen, Germany, 2016 March 14-17.

[130] M. Schroeder, D. Gottlob, and K. Siebrasse, "Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustic parameters," Journal of the Acoustical Society of America, vol. 56, pp. 1195–1201, 1974.

[131] W. Lachenmayr, C. Zamorano, E. Mommertz, and M. Kob, "Zum Einfluss von Reflexionen und Nachhall auf die Abstandswahrnehmung in Konzertsälen," in Proceedings of the German Acoustical Association (DAGA), pp. 339–342, 2015. Nuremberg, Germany, 2015 March 16-19.

[132] C. Zamorano, "Experimental investigations on proximity perception manipulating reflections in virtual and real room acoustical environments," Master's thesis, Hochschule für Musik Detmold, Germany, 2014.

[133] S. V. Amengual, W. Lachenmayr, and M. Kob, "Study on the influence of acoustics on organ playing using room enhancement," in Proceedings of the 3rd Vienna Talk on Music Acoustics, pp. 252–257, 2015. Vienna, Austria, 2015 September 16-19.

[134] S. V. Amengual, W. Lachenmayr, and E. Mommertz, "Spatial analysis and auralization for room acoustics using a tetrahedral microphone and SDM," Journal of the Acoustical Society of America Express Letters (conditionally accepted), 2017.

[135] S. V. Amengual, W. Lachenmayr, D. Kisic, and M. Kob, "Preliminary results of the effect of reverberation and sound delay on organ playing," in Seminar des DEGA-Fachausschusses Musikalische Akustik (FAMA), pp. 64–67, 2015. Hamburg, Germany, 2015 October 23-24.

List of publications

During the time of completion of this dissertation several articles were published, thus contributing to the scientific community. The following articles were written by the author:

- W. Lachenmayr, "Loudness-weighting of reverberation using electronic room enhancement," in the 60th International Conference of the Audio Engineering Society, 2016. paper no. 3-3, Leuven, Belgium, 2016 February 3-5

- W. Lachenmayr, A. Haapaniemi, and T. Lokki, "Direction of late reverberation and envelopment in two reproduced Berlin concert halls," in the 140th Convention of the Audio Engineering Society, 2016. paper no. 9503, Paris, France, 2016 June 4-7

- W. Lachenmayr and G. Engel, "Comparing room acoustics for the performance of Wagners Lohengrin," in the 138th Convention of the Audio Engineering Society, 2015. e-Brief 191, Warsaw, Poland, 2015 May 7-10

- W. Lachenmayr, C. Zamorano, E. Mommertz, and M. Kob, "Zum Einfluss von Reflexionen und Nachhall auf die Abstandswahrnehmung in Konzertsälen," in Proceedings of the German Acoustical Association (DAGA), pp. 339–342, 2015. Nuremberg, Germany, 2015 March 16-19

- W. Lachenmayr and J. Pätynen, "Influence of acoustics on emotional impact of music in Konzerthaus and Philharmonie Berlin," in Proceedings of the German Acoustical Association (DAGA), pp. 297–300, 2016. Aachen, Germany, 2016 March 14-17

The following publications were co-written and contributed to by the author:

- A. V. Osses, W. Lachenmayr, E. Mommertz, and A. Kohlrausch, "Predicting the perceived reverberation in different room acoustic environments using a binaural auditory model," Journal of the Acoustical Society of America Express Letters (conditionally accepted), 2017

- S. V. Amengual, W. Lachenmayr, and E. Mommertz, "Spatial analysis and auralization for room acoustics using a tetrahedral microphone and SDM," Journal of the Acoustical Society of America Express Letters (conditionally accepted), 2017

- S. V. Amengual, W. Lachenmayr, and M. Kob, "Study on the influence of acoustics on organ playing using room enhancement," in Proceedings of the 3rd Vienna Talk on Music Acoustics, pp. 252–257, 2015. Vienna, Austria, 2015 September 16-19

- S. V. Amengual, W. Lachenmayr, D. Kisic, and M. Kob, "Preliminary results of the effect of reverberation and sound delay on organ playing," in Seminar des DEGA-Fachausschusses Musikalische Akustik (FAMA), pp. 64–67, 2015. Hamburg, Germany, 2015 October 23-24

Curriculum vitae

April 17th, 1988	Born in Munich, Germany
1994-1998	Elementary school Oselschule, Munich Germany
1998-2007	Louise-Schroeder-Gymnasium in Munich, Germany, Abitur
October 2007	Beginning of the studies for Tonmeister at the University of Music and Performing Arts Vienna, Austria
July 2009/2011, Sept.- Oct. 2010	Student research assistant at the acoustical laboratory of the University of Applied Sciences Rosenheim, Germany
May - September 2012/ 2013	Audio engineering and production scholarships at The Banff Centre, Banff, Canada
May 2013	Graduation from the studies as Dipl.-Tonmeister, Mag. art. at the University of Music and Performing Arts Vienna, Austria
December 2013 - December 2016	Early-Stage-Researcher and fellow in Marie-Curie EU Training Network "BATWOMAN", hosted by Müller-BBM
April 2014 - February 2017	PhD candidate at the Erich-Thienhaus-Institute, Hochschule für Musik Detmold, Germany
since October 2016	Teaching position (Acoustics for sound-engineers) at the University of Music and Performing Arts Vienna, Austria
since January 2017	Acoustical consultant at Müller-BBM, Planegg/Munich, Germany