Listener Envelopment

Effects of changing the sidewall material in a model of an existing concert hall

Dan Nyberg
Listener Envelopment
Effects of changing the sidewall material in a model of an existing concert hall

Dan Nyberg

Arena Music Media and Technology
Luleå University of Technology
The Department of Music and Media
-2008-

Supervisor: Jan Berg
Luleå University of Technology, the Department of Music and Media
Abstract
This master thesis examines how the perception of listener envelopment is changed when altering the sidewall material in a computer model of a real concert hall. The research question is based on the lack of information on how the frequency content of the late arriving sound energy affects listener envelopment, but also on the conclusion that envelopment is an affect to large degree of late lateral reflective sounds. A computerized model is used for the purpose to take the research of listener envelopment into a more real acoustical situation than before. Alterations of the late lateral reflections are done by using four side wall materials attaining different sets of absorptions coefficients. A listening test is used for evaluating the change of listener envelopment and 21 listeners conducted the test. Results from earlier research were used to explain the attained results from the ANOVA tests. Sidewall materials which absorbs high frequencies and those that are nonabsorbent creates larger sensations of listener envelopment than materials in which absorbs low, low-mid and mid frequencies in this particular hall.
Acknowledgements
The author would like to thank, Björn Tunemalm and Rickard Ökvist for contributing with the computer model of Studio Acusticum concert hall and the absorption coefficients of the original sidewall material, Roger Johansson for helping with the CATT Acoustics software and all the other people involved in the completion of this work. Thank you.
Abbreviations

AH = Auralization with HRTF, Auralization is defined as the process of rendering the audible, by physical or mathematical modelling, sound field of a source in a space as to simulate the binaural listening experience at a given position in the modelled space.

ASW = Apparent source width or Auditory source width has same abbreviation

BE<sub>late</sub> = Late back energy ratio, Relative level of the late energy to the early one

BL<sub>late</sub> = Late back sound level

BRD = Binaural recording with a dummy head

BSI = Background spatial impression

C<sub>80</sub> = Early to late sound ratio see also C<sub>80</sub> (125-1K)

C<sub>80</sub> (125-1K) = Ten times the logarithm of the ratio of the sums of the early- and late-sound energies level averaged over octave bands 125 Hz to 1 kHz, the early energy is defined as the direct sound plus all reflections arriving within 80 ms

C-value = Difference in level between front to back energy

Co = Concrete material (original from real concert hall)

CSI = Continuous spatial impression

ESI = Early spatial impression

F = Fiber-rigid back (attained from CATT Acoustics library)

FE<sub>late</sub> = Late frontal energy ratio, relative level of the late energy to the early one

FL<sub>late</sub> = Late frontal sound level

Front to back energy ratio = $10 \log_{10} \frac{\text{Front}}{\text{Back}(dB)}$

H<sub>0</sub> = Null hypothesis

H<sub>1</sub> = Alternative hypothesis

G<sub>L</sub> (125-1K) = Measured late-arriving relative sound levels, averaged over octave bands 125 Hz to 1 kHz

G<sub>LL</sub> (125-1K) = Measured late lateral sound levels, averaged over octave bands 125 Hz to 1 kHz
HRTF = Head related transfer function (HRTF) is the acoustic filtering that free field sounds, arriving from different angles, undergoes because of the head, torso and pinna.

IID = Interaural intensity difference

ITD = Interaural time difference

Jnd = Just noticeable difference

LEV = Listener envelopment

LE_{late} = Late lateral energy ratio, Relative level of the late energy to the early one

\( LG_{x}^{\infty} = \) Sum of lateral energy averaged over octave bands 125 Hz – 1 kHz. Term used when searching for a lower integration time (referred to in the text as objective measurement).

\( LG_{80}^{\infty} = \) Sum of lateral energy after 80 ms average over octave bands 125 Hz – 1 kHz

\( LG_{105}^{\infty} = \) Sum of lateral energy after 105 ms average over octave bands 125 Hz – 1 kHz

LL_{late} = Late lateral sound level

RT = Reverberation time

TLL = The late to early energy ratio, Relative level of the late energy to the early energy

T-30 = Reverberation time derived from a straight-line least-square fit to the receiver curve at the interval -5 to -35 dB. This measurement is considered to be the best estimated of reverberation time.

T1 = Test material 1, sidewall material absorbing low frequencies

T2 = Test material 2, sidewall material absorbing high frequencies

VE_{late} = Late overhead energy ratio, Relative level of the late energy to the early one

VL_{late} = Late overhead sound level
Table of content

Abstract ..................................................................................................................... 1
Acknowledgements ................................................................................................. 2
Abbreviations ........................................................................................................... 3
1. Introduction ......................................................................................................... 6
2. Background .......................................................................................................... 7
   2.1 Definition of listener envelopment ................................................................. 7
   2.2 Properties of listener envelopment ................................................................. 8
      2.2.1 Late arriving sound energy impact on LEV ............................................. 8
      2.2.2 Relation between ASW and LEV in real environments ......................... 11
      2.2.3 Interaural functions and stimuli effect on spatial impression ................. 12
      2.2.4 Angle of arrival ................................................................................. 15
      2.2.5 Frequency and Reverberation influences on LEV ................................. 19
   2.3 Discussion and summary of the earlier work .................................................. 21
3. Research questions ............................................................................................. 24
4. Method ................................................................................................................ 25
   4.1 Selecting sidewall material .......................................................................... 25
   4.2 Creation of stimuli ...................................................................................... 27
   4.3 HRTF and auralization process ................................................................... 28
   4.4 Calibration of stimuli and listening level ...................................................... 28
   4.5 Creation of impulse responses and the auralization process ....................... 30
   4.6 Preliminary study ....................................................................................... 32
   4.7 Listening test ............................................................................................... 32
      4.7.1 Subjects ............................................................................................ 32
      4.7.2 Test method ...................................................................................... 32
      4.7.3 Data acquisition ............................................................................... 33
      4.7.4 Equipment and location of listening test .............................................. 33
      4.7.5 Equipment problems ...................................................................... 34
   4.8 Analysis ......................................................................................................... 34
      4.8.1 Post screening .................................................................................. 35
      4.8.2 Normal distribution ......................................................................... 35
      4.8.3 ANOVA analysis ............................................................................. 36
      4.8.4 Comparison in writing .................................................................... 36
5. Results ................................................................................................................. 36
   5.1 Normal distribution .................................................................................... 37
   5.2 ANOVA table ............................................................................................. 37
   5.3 Post-hoc test ............................................................................................. 38
   5.4 Comparison in writing ............................................................................. 38
6. Analysis .............................................................................................................. 40
7. Discussion ........................................................................................................... 42
   7.1 Results ....................................................................................................... 42
   7.2 Method ....................................................................................................... 44
8. Conclusions ........................................................................................................ 45
9. Further work ...................................................................................................... 45
References .............................................................................................................. 46

Appendix A ............................................................................................................ 48
Appendix B ............................................................................................................ 49
Appendix C ............................................................................................................ 51
Appendix D ............................................................................................................ 52
Appendix E ............................................................................................................ 53
1. Introduction

A variety of research has been conducted over the last three decades on the topic of spatial impression and how humans perceive listener envelopment in real environments. Much of the work show that the terms spaciousness and spatial impression can be divided into two factors [1,2,3].

- Auditory Source Width (ASW)
- Listener Envelopment (LEV)

There is however a confusion in the definitions of the terms spatial impression and spaciousness and weather or not they are different word for the same thing. It is indicated in [4][5] that spatial impression is related to the room impression, how large or small the room is perceived to be, and the term spaciousness is related to the degree to which the listener is in a large and enveloping space. When spaciousness is high spatial impression can also be high however the opposite can not perceived [4], so the two terms are related in this aspect.

Work by Berg [6] on the development of explanatory attributes for sound, shows that listener envelopment is an attribute in which gives the listener an impression of being surrounded by sound or being within the sound source. To be surrounded by the sound was a positive quality according to Berg [6]. This result was mostly due to listeners who were used to two channel stereo, got a positive experience when a multi-channel system gave an increase in envelopment [6]. Auditory source width however is related to the perceived broadness of the sound source and not to the perceived envelopment of the sound source [7].

A research review for this thesis has shown that listener envelopment is dependent on several physical factors; late arriving reflective sound energy from the initial sound, angular distribution of the late arriving sound energy, frequency content of the sound source, different reverberation times at different frequency bands, type of sound source and the amount of interaural time and intensity fluctuations at the ear of the listener.

Interesting indications emerged in front of the author when looking at the earlier research conducted on listener envelopment. These indications were that the frequency spectra of the late reflective sound energy could have an affect on listener envelopment and that the lateral reflections create the strongest sensation of all the arrival angles on envelopment. However these findings have not been investigated in the same test. It was also found that the majority of research has been conducted in controlled acoustical environments, such as anechoic chambers and damped listening rooms, with the reflections played over loudspeakers. This latter result raises the question if these results are valid and attained in real concert halls?

The author has incorporated these findings into a research question which are examined in this case-study. A computer model of a real concert hall Studio Acoustcum, Pitèå, Sweden is used and this approach takes the research one step closer to a real acoustical situation, where the reflections are closer to a real situation, then it would be in an anechoic chamber or damped listening room. The created research question is;

- How do different sidewall materials change the perception of listener envelopment in a concert hall, when the side wall materials absorb different frequencies?
The sidewall materials are changed, with different absorption coefficients, in the computer model, for altering the frequency spectra of the lateral reflections.

2. Background

This section presents previous research in the area of listener envelopment and explains listener envelopment for the reader. This section will summarise the results from the previous research, which have led to the physical properties of listener envelopment.

2.1 Definition of listener envelopment

As stated in the introduction, vast research on listener envelopment has been done. Here follows some definitions of LEV by other authors.

*Morimoto et al* defines LEV as a part of spatial impression as [7]:

“…spatial impression comprises at least of the following two components. One is auditory source width (ASW) which is defined as the width of the sound image fused temporally and spatially with a direct sound’s image and the other is listener envelopment (LEV) which is the degree of the fullness of sound images around the listener, excluding a sound image composing ASW.”

*Soulodre et al* defines LEV as a part of spatial impression as [8]:

“Spatial impression is known to be an important part of good rated concert hall acoustics, and it is now well established that spatial impression is composed of at least two components: apparent source width (ASW) and listener envelopment (LEV). ASW is defined as a broadening of the apparent source width of the sound source, while LEV refers to the listener’s sense of being surrounded or enveloped by sound.”

Definition of LEV by *Wakuda et al* [9]:

“…the listener’s sensation when the surrounding space is filled with sound images other then a sound image composing ASW…LEV is predominantly produced by late-arriving lateral reflections”

*Griesinger*’s definition [4] is a summary of other authors’ definitions of LEV and is defined as surround impression instead of LEV, but as explained in [4], it attains the same properties as LEV. To sum up Griesinger’s summary; Surround impression is one part of the term spaciousness and is defined by other authors as enveloping or listener envelopment. Surround impression creates a surrounding effect.

Based on two factors that were common for all the statements, the author has found two criteria’s for the definition of LEV. The first factor is that the sound shall surround the listener and the second is that LEV is excluded from the sound image width, which is defined as auditory source width. One of Morimoto’s figure 1 in [7] illustrates this very well, and has been re-drawn by the author in figure 1. The authors’ definition of LEV for this work is:

“Listener envelopment is the impression of the sound from the sound source surrounding the listener, not the sound source width, and the amount that listener feels inside/enveloped by the sound image.”
2.2 Properties of listener envelopment

The properties for creating listener envelopment include both the physical and perceptual (psycho-acoustical) parts. This section will explain how LEV is created, based on the review of the previous research. This section will include the late arriving sound energy’s impact on LEV, the relationship between ASW and LEV in real environments, interaural functions and stimuli effect on spatial impression, angle of arrival and frequency and reverberation influences on LEV.

2.2.1 Late arriving sound energy impact on LEV

Researchers have investigated the arrival time of reflections to see where LEV arisen. Soulodre et al [8] and Bradley et al [5] showed that the sum of lateral energy after 80 ms ($LG_{80}^-$) and the sum of lateral energy after 105 ms ($LG_{105}^-$) averaged over octave bands 125 Hz to 1 kHz gave an increase in the perceived envelopment of the sound as opposed to other tried measurements techniques.

Larger arrival times up to and over 120 ms were also investigated in [8] and it was indicated that an arrival time higher than 120 ms would be more enveloping. However a pilot test indicated that these large delay times produced a disturbing echo, instead of increased LEV [8]. It was also found when searching for an objective measurement of LEV ($LG_{x}^-$) that the correlation between the means from LEV scores attained from experiment 2 in [8] versus the $LG_{x}^-$ drops of after 115 ms, figure 2. These results shows that 105 ms attains a higher correlation (r=0.918) with this new measurement then $LG_{80}^-$. 

![Figure 2: correlation versus the lower integration limit, taken from [8].](image-url)
Both these measurements $LG_{80}^-$ and $LG_{105}^-$ were concluded to be not optimal as objective measurements of LEV [8]. $LG_{80}^-$ was found to provide a higher correlation to the subjective results of three conducted experiments, table 1 [8]. $LG_{105}^-$ gave a higher correlation to the results in experiment 1 [8]. Soulodre et al [6] concluded therefore that $LG_{105}^-$ is not consistently better than $LG_{80}^-$, and that neither measurement is optimal and that further investigations would be necessary.

Table 1: Correlation between LEV scores and $LG_{80}^-$ and $LG_{105}^-$ [8].

<table>
<thead>
<tr>
<th></th>
<th>Soulodre et al [6], Exp1</th>
<th>Soulodre et al [10], Exp1</th>
<th>Soulodre et al [10], Exp2</th>
<th>Soulodre et al [6], Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LG_{80}^-$</td>
<td>0.905</td>
<td>0.986</td>
<td>0.942</td>
<td>0.944</td>
</tr>
<tr>
<td>$LG_{105}^-$</td>
<td>0.918</td>
<td>0.977</td>
<td>0.974</td>
<td>0.956</td>
</tr>
</tbody>
</table>

Soulodre et al [8] also made an initial attempt to investigate the frequency dependent forward masking effect into an objective measurement, thus investigating different integration limits at different frequency bands. The integration limits for each octave band was selected by finding the values that gave the highest correlation to the subjective results of the study in [8]. The approach however limited the experiment because it ignored any integration across the frequency bands [8].

Table 2: Integration limits for new measure vs. octave band [8].

<table>
<thead>
<tr>
<th>Oct, band</th>
<th>63</th>
<th>125</th>
<th>250</th>
<th>500</th>
<th>1k</th>
<th>2k</th>
<th>4k</th>
<th>8k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int. limit (ms)</td>
<td>140</td>
<td>140</td>
<td>120</td>
<td>75</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>

Figure 3: Mean LEV scores for experiment 2 versus the new objective measure with frequency dependent integration limits, taken from [8].

The new measurement presented in on table 2, the straight line figure 3, are better than both $LG_{80}^-$ and $LG_{105}^-$ at predicting the LEV scores ($r= 0.949$) from experiment 2 in [8]. This measurement has however not undergone evaluations so no firm conclusions could be drawn about how effective the purposed measurement is [8].

Bradley et al [11] investigated the interaction of the effect of early to late sound energy ($C_{80}$ [125-1K]) on LEV for realistic conditions, see left plot in figure 5. $C_{80}$ [125-1K] is ten times the logarithm of the ratio of the sums of the early- and late-sound energies level averaged over
octave bands 125 Hz to 1 kHz, were the early energy is defined as the direct sound plus all reflections arriving within 80 ms. The late energy starts at 80 ms after the initial sound, thus a high level of $C_{80}$ has a low level of late energy after 80 ms and vice versa [11].

This experiment was also, as the previous experiments, a listening test. The stimuli were varied by changing the $C_{80}$ and overall listening level in 3 dB steps. For only investigating the change of $C_{80}$ and overall listening level the variables late sound energy, direction of the later and early arriving sound were kept constant. $C_{80}$ varied from -6 to +9 dB in 3 dB steps and the overall A-weighted levels of the musical content was 65, 68 and 71 dBA over a time of 18 s [11].

The result from the experiment shows that the LEV scores clearly decrease with increasing $C_{80}$ and decreasing overall level. An ANOVA analysis showed that there were highly significant main effects for both $C_{80}$ and the overall level however the interaction effect was not significant [11].

![Figure 4](image)

**Figure 4:** Mean LEV scores as a function of the later arriving sound levels $G_L[125-1k]$ taken from Bradley et al. [11]. The Author has excluded multiple linear regression analysis lines from original plot, for clarifying the presented results.

Figure 4 shows the mean LEV scores from Bradley et al’s experiment [11] but as a function of the later-arriving sound levels average over octave bands 125 Hz - 1 kHz ($G_L[125-1k]$), for the three overall levels. The comparison was used to better see the basic underlying effects of the results in term of early- and late-arriving sound levels. LEV is increasing at each data set when $G_L[125-1k]$ are increasing.

The graph also shows that the low and medium late arriving overall sound levels attain larger LEV scores than the high constant late arriving overall sound level. This latter decrease in LEV scores are due to, according to Bradley et al [11], a decrease in early level [11].

Griesinger showed a model based on his collected work [4,12,13] that envelopment is created after 150 ms if the late arriving reflections are spatially diffused. This is caused by an inhibition of the ear which starts after 50 ms and gradually decreases up to and over 150 ms. After these 150 ms the perception is becomes the background spatial impression (BSI) which is enveloping. The term BSI is discussed more in detail under the topic “Interaural functions and stimuli effect on spatial impression”.
Further more Griesinger’s [4,12,13] approach are not the same as Soulodre and Bradley’s [5,8,11] approach, where they are trying to find a time coefficient for envelopment, but examines the sound signal at the ear of the listener when envelopment is perceived.

2.2.2 Relation between ASW and LEV in real environments

Bradley et al [11] concluded as many others [1,2,3] that spatial impression is dependent on both ASW and LEV and both are present in empiric situations such as concert halls. Bradley et al [11] conducted an extended experiment, B, based on experiment A in [10]. This extended experiment attained wider range of $C_{80}(125-1K)$ values varying from -12 to +18dB, to better understand the perception of envelopment when either early- or late-arriving sound is dominant over the other.

The results, in figure 5, shows that between $C_{80}(125-1K)$ values ranging from -6 to +9 dB are very similar to the results from experiment A. For values between -4 and +6 dB there is again a linear relationship between the $C_{80}(125-1K)$ and the LEV responses. The most important results are the trends in either extreme of the $C_{80}(125-1K)$ at the right plot in figure 5 [11].

For very high $C_{80}(125-1K)$ values, the LEV scores approaches 1 in spite of the overall level and at for low $C_{80}(125-1K)$ values the LEV score attains a different height for each sound level. This shows the trade of between early- and late-arriving sound on the LEV scores, shown in experiment B and it is limited to the range where both have perceptible contributions. It is also shown when one component is 10 or 12 dB greater then the other, only the dominant component influences the judgment [11]. Thus early- and late-arriving energy is affecting each other perceptually. If late-arriving sound energy increases, then the impression of LEV increases. If early arriving sound energy (one parameter of creating ASW) is more dominant, then the impression of LEV is reduced [11].

![Figure 5: Both plots shows the mean LEV scores versus $C_{80}(125-1K)$ ratios at three overall sound levels, 65, 68 and 71 dBA. Left plot taken from experiment A and right plot taken from experiment B, taken from [11].](image)

Bradley et al [11] found also that the early arriving sound masks the late arriving sound more efficient, than the late arriving sound masks the early sound. Bradley et al [10] concluded that the balance between ASW and LEV are dependent on the balance between the early- and late-arriving lateral sounds.
2.2.3 Interaural functions and stimuli effect on spatial impression

Griesinger [4,12,13] have found and shown that the interaural functions, interaural intensity difference (IID) and interaural time difference (ITD), which are primary functions of determine localization is important for creation of envelopment in concert halls. This is confirmed to a certain degree by Mason et al [14]. In Mason et al’s study the ITD altered the perceived width and depth of the stimuli used.

Griesinger’s work [12,13] found that low frequencies less than 1000 Hz are dependent on both ITD and IID. The ITD is the principal cue for localizing the direction of a sound horizontally under 400 Hz, this function is very accurate when no reflections is present. This conclusion is supported by Blauert [15], his work shows that for sinusoid signals the localization for low frequencies is dependent on interaural time difference and for high frequencies it is dependent on the interaural intensity difference. However Blauert also shows that the intensity difference also affects the localization at the same time as the time difference and that both interaural functions are present and affecting the localization of a sound source. Thus makes the IID and ITD not independent in different frequency ranges [15]. These results from Griesinger [12,13] and Blauert [15] are interesting to pursue since there is an apparent difference in opinion, however it is not in the scope of this thesis, but to merely indicate that fluctuations in ITD and IID are affecting LEV as will be seen below.

Griesinger [12,13] shows that adding lateral- and late lateral-energy to the sound, interference is created between the direct to the lateral and late lateral arriving sound. This interference makes the ITD to shift. This shift becomes a fluctuation. When this fluctuation attains a speed of 3 Hz or more, then the perception of envelopment arises. Localization of sounds that have rapid attacks is also determined by ITD, this is due to the absence of reflections at the listener. The occurrence of envelopment in this situation is dependent on the late arriving reflections of sound for creating the fluctuations needed. At higher frequencies, > 1000 Hz, envelopment is dependent on the IID to create the interaural fluctuations needed. For music that is relatively transparent to reverberation, i.e. staccato played instrument, envelopment is created by the reverberant component. For continuous sound sources i.e. full orchestra, the total amount of fluctuations of ITD and IID creates the perceived envelopment of the perceived sound.

Mason et al [14] also found that fluctuations in the ITD create a perception of depth and width of the sound source. This perception was different depending on the angle of the arriving fluctuations. Fluctuations arriving from +/- 30 degrees towards the listener gave an impression of the sound to be in front and backwards the listener. With +/- 90 degrees the stimuli got an impression of being around the listener. By increasing the time difference fluctuation magnitude, at both +/- 30 and 90 degrees at different occasions respectably made the stimuli to wrap around the listener. The increase of depth was smaller than the increase in width, this can be seen in figure 6 were the amount of ITD fluctuations versus the perceived depth and width are shown from Mason et al’s [14] study.
Figure 6: Plot of mean values of the z-transformed Width and Depth measurements of the overall scene dimensions. The graphical responses are from all subjects and loudspeakers reproduction positions separated by time difference fluctuations magnitude, taken from [14].

Griesinger [4] concludes that spatial perception depends on the type of sound used. Different sounds are separated by the brain into different streams foreground streams and background streams. Foreground streams is related to sound event that are connected by timbre, pitch and note/meaning and early spatial impression, the latter is discussed later. The background stream is the late arriving sound from the initial sound which is not connected to the sound source as the foreground stream, it can also be called diffuseness of the sound, reverberance of sound and background spatial impression, the latter is discussed later [4].

Griesinger identified three spatial impressions, background spatial impression (BSI), early spatial impression (ESI) and continuous spatial impression (CSI). Only BSI and CSI can be enveloping.

Continuous spatial impression appears when lateral sound energy interferes with a continuous sound source. The delay time of the reflections needed for creating CSI is about 10 ms. This spatial impression is fully enveloping. It is also dependent on the ratio of the median sound to the lateral sound. Continuous spatial impression is also independent from the absolute sound level of the source [4,12,13]. CSI primarily depends on the degree of fluctuations in IID and ITD, which depends on the direct to the reverberant ratio and the azimuth of the reflections [4].
Early spatial impression results from lateral early energy which arrives within 50 ms of the end of an impulsive sound. This impression is not enveloping and is bound to the sound source (Haas effect) and to the foreground stream. ESI is also dependent on the ratio between the medial and lateral sound and does not depend on the source strength [4,12,13].

Background spatial impression appears when the source consists of a series of short notes or a sequence of phonemes. The brain organizes this initial sound into a foreground perceptual stream and if there is more than one sound source the brain organize the sources as separate streams. The energy arriving between these initial sounds is assigned to a single stream, which is called the sonic background [4,12,13]. The sound energy which is perceived between the initial sounds contains the reflections that cause the IID and ITD to fluctuate. These fluctuations change widely when reflected sound is present. The minimum delay needed for the reflections to be included in the background perception is 50 ms. When reverberation is present and the delay time increases to about 160 ms, the background impression becomes audible for the listener. But the effective delay time depends on the listener due to individual variations.

Griesinger [13] concluded from a listening test where the reverberation time (RT), the delay of the RT and source material was changed, that there is a neural inhibition to the perception of the background. This neural inhibition is lasting from a minimum of 50 ms after the end of the direct sound in the foreground event and is slowly decreasing. At 150 ms from the direct sound the inhibition has almost released which makes the background sound becomes strongly audible and if this sound is spatially diffused, envelopment is perceived, see figure 8.

![Figure 8: The separation process assigns the early fluctuations to the note itself. After a period of inhibition the later fluctuations are assigned to the sonic background, taken from [13].](image)

The BSI is absolute based on the amount of spatially diffused reverberant energy, which is dependent on the strength of the sound source. CSI perception depends on the ratio of the direct sound to the reverberant sound. This perception is smaller than the perception of BSI. This perceived difference between CSI and BSI is level dependent, and is approximately 6 dB. This means that solo music/lightly scored music (BSI) is perceived as more reverberant than thickly scored music (CSI) even if it’s louder in level [12].
2.2.4 Angle of arrival

Wakuda et al [9] and Furuya et al [16] investigated, at two separate occasions, the direction and the total amount of late sound energy on LEV, with listening tests.

Wakuda et al’s results attained from the listening test showed that the relative sound levels of each directional late energy (after 80 ms) to the direct sound influenced LEV. It was concluded that LEV increased as the late lateral sound level increased (LL\text{late}). The late sound levels from other directions than lateral, late back sound level (BL\text{late}) and late overhead sound level (VL\text{late}) also increased LEV. However late frontal sound level (FL\text{late}) was found to be not influencing LEV [9].

Based on the results a multiple regression analysis of the data was conducted, to see the contribution of each directional late sound level on LEV. It showed that the LL\text{late} had the most influence on LEV. It also showed that BL\text{late} and VL\text{late} influenced LEV to about 67 % and 48 % respectively of that for LL\text{late}, see table 3. Further more FL\text{late} was shown to influence LEV the least [9], see table 3.

Table 3: Results of multiple regression analysis between perceived LEV and the four directional late sound levels, significant at p < 0.05 [9].

<table>
<thead>
<tr>
<th>Multiple correlation coefficient</th>
<th>Standard registration coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL\text{late}</td>
<td>FL\text{late}</td>
</tr>
<tr>
<td>0.991</td>
<td>1.088</td>
</tr>
<tr>
<td></td>
<td>0.112</td>
</tr>
<tr>
<td></td>
<td>0.518</td>
</tr>
<tr>
<td></td>
<td>0.726</td>
</tr>
</tbody>
</table>

The second experiment conducted by Wakuda et al [9] investigated the relations between the directional late energy and LEV by keeping the total level of late sound constant. The directional late energy ratios LE\text{late} (late lateral energy ratio), VE\text{late} (late overhead energy ratio) and BE\text{late} (late back energy ratio) were changed for keeping the total level of the late sound constant. FE\text{late} (late frontal energy ratio) was fixed at 0.10 under the test. The experiment were divided into three different sets containing each one value of C\text{80}, -3, 0 and +3dB. The results showed that the perception of LEV is complicately related to the directional late energy ratios and it cannot be simplistically explained with one directional parameter. This was due to the unclear effects of the different directional late energies, [9]. A multiple regression analysis was therefore conducted on the three sets of C\text{80} to investigate the degrees of contribution of LE\text{late}, VE\text{late} and BE\text{late} to LEV [9]. The results showed that LE\text{late} attained the highest regression coefficient for all C\text{80} values. VE\text{late} is 31 % and 48 % of those of LE\text{late} and BE\text{late} is 46 and 49 % of those of LE\text{late} at C\text{80} values -3dB and 0 dB respectively, see table 4. The results shows according to Wakuda et al [9] that the contribution of late sound from above and behind the listener increases LEV at a rate of approximately 50 % of the effect of lateral angle when C\text{80} is -3dB.

Table 4: Results of multiple regression analysis between perceived LEV and the three directional late energy ratios, significant at p < 0.01 C\text{80}=-3dB, p < 0.005 for C\text{80}=0,+3dB. P levels based on analysis of variance [9].

<table>
<thead>
<tr>
<th>Experiment no.</th>
<th>C\text{80} (dB)</th>
<th>Multiple correlation coefficient</th>
<th>LE\text{late}</th>
<th>VE\text{late}</th>
<th>BE\text{late}</th>
</tr>
</thead>
<tbody>
<tr>
<td>2(a)</td>
<td>-3 dB</td>
<td>0.989</td>
<td>1.791</td>
<td>0.860</td>
<td>0.870</td>
</tr>
<tr>
<td>2(b)</td>
<td>0 dB</td>
<td>0.990</td>
<td>1.586</td>
<td>0.486</td>
<td>0.722</td>
</tr>
<tr>
<td>2(c)</td>
<td>+3 dB</td>
<td>0.997</td>
<td>1.086</td>
<td>0.200</td>
<td>-0.004</td>
</tr>
</tbody>
</table>

Furuya et al [16] examined the same properties for clarifying whether or not the degree of contribution of late energy with a single directional component is valid when the late sound
has plural directional components. As in [9] the directional late energies $LE_{late}$, $VE_{late}$ and $BE_{late}$ were varied but the late to early energy ratio (TLL) were kept constant. Also the $FE_{late}$ was fixed at 0,10 under the test because of the effect of late frontal sound was found to be much smaller than $LE_{late}$, $VE_{late}$ and $BE_{late}$. The multiple regression analysis on the data from this experiment used $LE_{late}$, $VE_{late}$ and $BE_{late}$ as explanatory variables and the psychological interval scale for LEV as the criterion variable as in [16]. The results from the multiple regression analysis show that $LE_{late}$ attains the highest coefficient of any case of TLL, see table 5. The coefficients for $VE_{late}$ and $BE_{late}$ are approximately 30-60 % of $LE_{late}$ which corresponds to the results attained by Wakuda et al, compared in [16].

Table 5: Results of multiple regression analysis between perceived LEV and the directional late energy ratios, significant at $p < 0.005$ [16].

<table>
<thead>
<tr>
<th>TLL (dB)</th>
<th>Multiple correlation coefficient</th>
<th>Standard regression coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$LE_{late}$</td>
<td>$VE_{late}$</td>
</tr>
<tr>
<td>3 dB</td>
<td>0.980</td>
<td>1.558</td>
</tr>
<tr>
<td>0 dB</td>
<td>0.910</td>
<td>1.686</td>
</tr>
</tbody>
</table>

The second experiment conducted by Furuya et al [16] investigated the degrees of contribution of the directional and total energy of the late sound on LEV. In this experiment both the directional late energies and TLL were varied. Also here a multiple regression analysis was implemented after attaining significant results below 5 % level. Here the TLL and late directional energies were the explanatory variables and the psychological interval scale for LEV was the criterion variable. As seen in table 6, the TLL regression coefficient is the largest, which means that the total amount of late energy contributes the most to LEV. For the three late directional energies the $LE_{late}$ is the largest and $VE_{late}$ and $BE_{late}$ are 35-62 % of that of $LE_{late}$, respectively, as can be seen in previous work [9] [16].

Table 6: Results of multiple regression analysis between perceived LEV and TLL and the three directional late energy ratios, significant at $p < 0.005$ [16].

<table>
<thead>
<tr>
<th>Multiple correlation coefficient</th>
<th>Standard regression coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLL</td>
<td>$LE_{late}$</td>
</tr>
<tr>
<td>0.932</td>
<td>0.967</td>
</tr>
</tbody>
</table>

Furuya et al [16] also compared the results from this study to Wakuda et al’s [9] experiment where the TLL was constant, with a correlation analysis, in order to verify the validity of the attained results, see table 7. A good correlation is obtained at TLL values of +3 and 0dB thus when late energy is not smaller then the early energy. Furuya et al [16] concludes that the ratio of late energy to the early has the largest effect on the perceived LEV. It is also reconfirmed that $LE_{late}$ strongly affects LEV. $VE_{late}$ and $BE_{late}$ are effective for LEV at the rate of approximately 30-60 % of the effect of $LE_{late}$

Table 7: Correlations of calculated LEV by the regression equation in experiment 2 [16] with measured LEV in experiment 2 [9] and experiment 2 [16]: r, correlation coefficient.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>3 dB</td>
<td>3 dB</td>
<td>0 dB</td>
</tr>
<tr>
<td>r</td>
<td>0.956</td>
<td>0.946</td>
</tr>
</tbody>
</table>
Bradley et al [11] investigated combined effects of the spatial distribution of the late energy and the sound level on LEV, with listening tests. Three sound levels where used in the experiment, 65, 68 and 71 dBA. The portion of late lateral was the lowest when all of the late energy was coming from the rear and front speakers towards the listener, which was expected by Bradley to give a minimum of perceived LEV. The late lateral energy was increased systematically by taking a portion of the late energy and adding them to the loudspeakers placed on the side of the listener. The total level of early- and late-arriving sound was maintained constant in the seven sound fields of each overall sound level set used in the experiment. The $C_{80}[125-1K]$ values for all stimuli used was -0,5 dB.

The results from this experiment shows that LEV are increasing with the increase of $G_{LL}$ (125-1K), which is the measured relative late lateral sound level. However the LEV values seem to reach a different plateau for each overall sound level set. There is also some $G_{LL}$ (125-1K) values which LEV will not vary, right hand side of figure 9. This residual level which Bradley et al talks about depends on the overall sound level [11].

The left most data points in figure 9 corresponds to the late energy directly in front of and behind the listener [11]. Bradley et al [11] concludes therefore with only late arriving sound from the back and from the front of the listener, LEV can vary with the level of the late-arriving sound. But the individual contribution of each direction on LEV could not be determined because they were not varied separately in the test.

![Figure 9: Mean LEV scores versus late lateral sound level, $G_{LL}(125-1K)$ for the three overall sound levels sets, taken from [11].](image)

Griesinger also showed that the narrowest angle of late arriving sound, for the creation of LEV, varies depending on the frequency content of the admitted sound [12,13]. A set of decorrelated loudspeakers, not playing the same signal, playing a stimulus with a large amount of frequencies with an upper cut of frequency of 1700 Hz creates envelopment at an typical loudspeaker placement of +/- 30 degrees. This angle is enough for creating envelopment when using this type of stimuli. An increase of upper cut of frequency of the sound makes the narrowest angle for creating LEV to move towards the median plane, which according to Griesinger coincided with results attained by Ando [13].

The use of narrow angles to produce envelopment can only be used when higher frequency content is available. That is because the amplitude of the fluctuations produced by the loudspeaker for creating envelopment depends on the sine of the angle of the loudspeaker.
For low frequencies the optimal angle is +/- 90 degrees and by moving the speaker towards the front, the fluctuations is decreased, as is the sensation of envelopment. Figure 10 shows the angular dependency of envelopment for a single reflection from Griesinger’s study [12].

Morimoto et al [7] showed that the front/back ratio is important for envelopment at three different C-values\(^1\). The experiment showed that by increasing the late backward energy the sensation of LEV is increased, as shown for in figure 11. These results was statistically significant at the level of 0,68 jnd\(^2\).

Griesinger [12,13] believes that this front/back ratio is primarily important for sound that includes a high portion of frequency energy above 3 kHz. This was due to the spectral perceptual cues for distinguishing between front and back localization. A small change in ITD caused by small movements of the head, which also is a perceptual cue of localizing front and rear sound, can not be used in the situation where the sound is spatially diffused. Griesinger also believes that the front/back discrimination is done during the separation of the sound into a foreground and background stream. This means that front/back ratio, is likely contributing to BSI and CSI where the BSI is more perceptually important [12,13].

\(^1\) Difference in level between front to back energy  
\(^2\) Jnd = Just noticeable difference
2.2.5 Frequency and Reverberation influences on LEV

Bradley and Soulodre et al's [5] investigation of trying to find an objective measurement of LEV concluded that the quantity $L_{\nu}G_{80}$ was most suitable for this purpose. They also concluded that the sum of octave bands 125 Hz - 1 kHz gave the highest correlation between a variety of sums between octave bands and other quantities, which indicated that low and mid frequencies are important for adequate LEV [5].

Morimoto et al [17] investigated the effect of reverberation time and frequency content on listener envelopment, with a listening test. The first experiment examined whether reverberation time affects LEV when changing the reverberation time (RT) and early-to-late ratio ($C_{80}$) independently, when keeping the relative sound pressure level of the reverberation sound constant.

Figure 12 shows the results from the first experiment with varied reverberation time and varied $C_{80}$ with the same sound pressure. Morimoto et al [17] concludes that RTs are statistically significantly different at all of the $C_{80}$ values. These results are shown to be statistically different from each other, and the probability of discriminating the difference is about 95 %, thus an increase in reverberation time increases the perception of LEV. The differences for all $C_{80}$ values are almost identical, which means that the effect of reverberation time on LEV is independent of $C_{80}$. The probability for discriminating the differences between the highest and lowest LEV is 96.6 % and 97.2% for RT = 1.0 and 2.0 s respectively, which indicates that $C_{80}$ also significantly affects LEV.

![Figure 12: Psychological scale of LEV as a function of $C_{80}$ and as a parameter of RT in experiment 1, taken from [16].](image)

In experiment 2 conducted by Morimoto et al [17] an investigation of reverberation time at different low frequencies bands was carried out. In the experiment the reverberation time at frequency bands 125, 250, 500, 1k, 2k, 4k and 8k were varied in this fashion, see appendix A. Stimuli 3 attained a relative flat RT over the frequency bands and stimuli 1 and 2 attained a shorter RT for the lower frequency bands than stimulus 3. Stimuli 4 to 6 had longer RT values then stimulus 3 in all frequency bands. $C_{80}$ where kept constant at 0 dB under the experiment and the relative sound pressure level of the reverberation sound at -25,6 +/- 0,5 dB and -27,6 +/- 0,8 dB for RT=1,0 and 2,0 respectively.
Figure 13: Psychological scale of LEV as a function of frequency characteristics of RT for (a) RT = 1.0 s and (b) RT = 2.0 s, taken from [17].

Figure 13 illustrates the perceived LEV on the different stimuli. The just noticeable difference at graph a) does not show a statistical difference at the 75% jnd level and graph b) show that there is a perceived difference at the used confidence level. By comparing the difference between stimuli 1 and 6, shortest RT and longest RT, respectively for both a) and b), shows that the subjects could discriminate the difference. Graph a) and b) thereby shows that lengthening the RT at low frequency bands increased the sensation of LEV as seen in figure 13.

Experiment 3 was implemented to investigate to see how shortening the reverberation time at low and high frequencies was perceived. The experiment used a $C_{80}$ at 0 dB and a reverberation time sound pressure level of $-26.3 \pm 0.3$ dB. The reference RT was 2.0 s at stimulus 3. Five other RT was used in the experiment. Stimuli 1 and 2 attained shorter RT at the low frequency bands and stimuli 4 to 5 attained shorter RT in the high frequency bands [17].

The results show that the decrease in RT at the low frequency bands, stimuli 1 and 2, are statistically discriminately when comparing to stimulus 3. Also the shortening of RT at the higher frequency band, stimuli 4 to 5 is also discriminated [17], see figure 14.

Summary of Morimoto et al.’s [17] finding; The results shows that $C_{80}$ and RT is individually influencing LEV and are not affecting each other. Changing the total amount of RT of a sound and the individual RTs for different frequency bands changes the perceived listener
envelopment. As the total reverberation time becomes longer, the perception of LEV also increases. By decreasing the RT of the low frequency bands, the perception of LEV decreases and the opposite happens when lengthening the RT at the same bands. This also happens for the high frequency bands [17].

2.3 Discussion and summary of the earlier work

It is apparent from the review of the earlier work, that listener envelopment is affected to a large degree of many physical parameters such as; late- to early-sound energy, the angle from which the late sound comes from, frequency content of the initial sound, loudness of a sound source, the degree of phase and amplitude fluctuations of the late sound, the reverberation times length at different frequency bands and the type of musical material used. This section will discuss the results attained from the review.

It can be concluded that the late arriving sound, i.e. late reflections affect listener envelopment. The time limit in which the late arriving sound can be classified as late is approximately between 80, 105 ms after the initial sound. Also an limit of 150 ms for background spatial impression was proposed. This quite late limit is due to a perceptual inhibition of the ear which give arise to the background spatial impression (BSI). There are also indications that different integration limits for the late sound is frequency dependent i.e. different octave bands have different integration limits, in ms, where LEV is created. A longer integration limit for low frequency bands and a shorter integration limit for higher frequency bands.

An increase in amplitude of the late arriving sound energy over the early sound energy increases the sensation of LEV. The relationship between ASW and LEV are based on the strength of early- and late- arriving sound energy respectively. When one component, late or early arriving energy, is about 10 -12 dB stronger, this component becomes dominant and influences the perception. It is also concluded when the late sound is stronger than early energy and the overall sound level is gained, an increase in LEV will appear. Also when the early arriving sound is stronger, then a sensation of LEV can still be perceived to a lesser extent, but now LEV is not dependent on the overall sound level. Both ASW and LEV can be found in real environments and are linked to each other.

ITD and IID fluctuations are important for the creation of LEV. These fluctuations affect all parameter that influences spatial impression.

- Angular direction of the late sound is depending on the frequency spectra of the sound for creating envelopment. Sounds attaining higher frequency content has a narrower angle of the late arriving sound for creating envelopment. Late sound energy arriving from +/- 90 degrees creates the most fluctuations in IID and ITD, which creates the largest sensation of envelopment.

- Perceived width and depth of a sound source varies with the change in ITD of a sound. If the sound comes from an angle of +/- 30 degrees with a low amount of fluctuations the sound is perceived as in front and behind the listener. Increasing the amount of the fluctuations moves the sound towards the side of the listener. This happens also when the fluctuations arises from late energy that comes from an angle of +/- 90 degrees, however the sound in this occasion is perceived to be already around the listener when a low amount of fluctuations is used. Increasing the fluctuations makes the sound to
wrap more around the listener. The width of the sound source is affected more than the perception of depth by the ITD fluctuations.

- Depending on the stimulus used as a sound source, different sensations of spatial impression are created and two of them are enveloping; continuous spatial impression (CSI) and background spatial impression (BSI). Late arriving sound \( > 10 \) ms is needed for creating CSI. This impression is dependent of the ratio of the medial sound to the lateral sound and the fluctuation of IID and ITD, but is not dependent on the absolute sound level of the source.

BSI arises when a sound source consist of a series of short notes or phonemes. The human brain separates the impulsive sounds and the sound energy between the sounds into separate streams. The energy arriving after the initial sound contains the reflections from the hall which caused the fluctuations in the ITD and IID to fluctuate widely. This creates the sensation of envelopment. The minimum of delay needed for the reflections to be separated into the background stream is 50 ms. As the delay time is changed up to 150 ms the background impression becomes more audible.

Further more BSI is dependent of the amount of spatially diffused reverberant energy, and is thereby dependent on the level of the sound source. The perceived difference between CSI and BSI is about 6 dB. This means that short sounds or notes in series are perceived as more reverberant than the continuous sounds, even though the latter is louder in level.

The angle of arrival of the late sound energy also affects the impression of LEV. The angle which affects LEV the most is the lateral angle. Sounds that come from this angle create a strong perception of envelopment.

The second most important angle is from the back. Late sound energy from this angle affect envelopment approximately 50 – 60 % then lateral angle would affect LEV. Late sounds coming from the back are also frequency dependent. This is primary important for sounds that contain large amounts of frequencies over 3 kHz. This is due to the perceptual cues for distinguishing the front and back locations are inhibited when IID and ITD are fluctuating widely. The third most important angle is the overhead angle. Late sound from this angle affects envelopment approximately 30 – 50 % of the effect of the lateral angle. And the least most enveloping angle is the frontal angle. The frequency content of the late arriving sound also affects the angular dependency of LEV and not only the backward angle. At higher frequencies the narrowest angle for creating listener envelopment moves towards the median plane, and vice versa happens with lower frequencies. This is due to ITD and IID fluctuations which changes with angle and frequency.

The reverberation time and frequency spectra are also influencing LEV. First of all the reverberation time and the early to late sound level is individually affecting LEV and is in that respect independent from each other as Morimoto et al. [17] showed. Increasing the late sound energy increases the LEV and increasing the total reverberation time also increases the sensation of LEV. By lengthening the reverberation times for low frequency bands 125 - 500 Hz increases the sensation of LEV. Decreasing the reverberation times at high and low frequency bands 125-500 Hz and 2-8 kHz decreases the sensation of LEV.

There is a consensus that late arriving sound energy are affecting LEV, however it’s not certain at which point the late arriving sound becomes late arriving. It is shown from the
results stated above that 80, 105 ms would be a good choice of setting a lower limit in which late arriving sound becomes late arriving. A longer integration limit for low frequencies would be good to use for creating the envelopment. These integration limits needs be longer than the 80, 105 ms, however it is not concluded if this would be better than 80, 105 ms integration time averaged over the octave bands 125 Hz -1 kHz. There are however more research on integration times 80, 105 ms than for 150 ms, which indicates that using 80, 105 ms as a lower integration time would be a “safer bet” when defining late arriving sound. When large amounts of late arriving sound is present then usually the reverberation is also large. The reverberation time is frequency dependent, as seen from the review. Long reverberation times at low and high frequency bands increases the sensation of LEV. However the low frequencies seem to be more important due to the objective measurements of LEV, $LG_l^\infty$, which uses an average value for the lower frequency bands to measure LEV.

The LEV is also level dependent as indicated in the results. The explanation to this is not particularly far-fetched, since both results that shows this, are based on the late-arriving sound which essentially is the reverberant sound. However it is interesting to see that LEV is level dependent, when late arriving sound is higher in level than the early sound energy makes the reverberation more perceived. The BSI which is primarily dependent on the reverberation to cause the fluctuations in ITD and IID, gives further insight into the level dependency of LEV. When sound sources are short and impulsive in nature, then the level of sound source is important for giving high reverberant late sound i.e. a sound source with a high sound level creates a larger perception of reverberation and is thereby level dependent.

Late sound energy that is arriving from lateral directions has been shown to affect the appearance of LEV the most. The cause of this can be explained by the interaural time and phase fluctuations, at the lateral angle these fluctuations is most widely. It is also shown that when these fluctuation increases, the sound wraps around the listener’s head and thereby creating envelopment. The perception of CSI is also dependent of the lateral arriving sound versus the direct. The rise of envelopment is dependent on the fluctuations caused by these lateral reflections. This indicates that there is a consensus in why late lateral arriving sound affects LEV with continuous sound sources. The backward arriving late energy is the second most important angle for the creation of envelopment. An explanation for this is because it is frequency dependent, however only for sounds that have large amount of frequencies over 3 kHz. This raises the question on how important the backward arriving sound in a real environment where the high frequency material is absorbed by the air, since the results are attained from tests with loudspeakers.

It can also be concluded that overhead and frontal arriving late energy seems to be level dependent and affect LEV when the total level of late arriving energy is raised.

The previous research has been conducted in anechoic environments and acoustically controlled rooms with reflections played over loudspeakers. This raises the question to which extent these results are valid in real concert hall?

These findings discussed here are essential points to have in mind and to investigate further when researching listener envelopment. One factor that will be investigated in this thesis is how the frequency content of the late arriving reflections changes the perception of listener envelopment in real environment. There has been research conducted on the topic as seen from the review, but it has not been investigated in a real natural environment, such as a concert hall. The major aim for this thesis is presented in the next chapter.
3. Research questions
In this section the research questions and aim for the conducted work is presented. The research question is based on the summary from the review of the earlier research.

To contribute and to investigate the perception of listener envelopment further, some questions needs to be addressed. To which extent are these results, stated in the earlier work, valid in real environments where listener envelopment is present? Here follows some issues that will be investigated which can shed light on this matter.

- What effect does the spectral content of the late lateral reflections have on LEV?
- How is envelopment perceived by changing the musical instrument?

It is shown from the review above that the late lateral energy has a large influence on perceived LEV, however these results are attained in acoustic controlled environments, such as anechoic chambers and listening rooms. It is also shown that the spectral content of the reflected late sound has large effects on LEV, primarily for low frequencies, but also for high frequencies. The high frequencies can however in a real concert hall be different then in a controlled environment with the reflections played over loudspeakers. In a real concert hall the air absorbs the high frequencies, this would be interesting to investigate in order to see if this has an effect on LEV. To investigate how LEV is affected by the spectral content of the late arriving sound, variations of sidewall materials, absorbing and reflecting different frequencies, can be used and applied in a real concert hall which attains a high sensation of LEV. By altering the sidewall material of the concert hall could create different late lateral reflection with varied spectral content which can be studied.

It is also shown that different musical content, short series of notes or continuous sounds could give arise to envelopment. These are dependent on different physical factors such as level of the late reverberation, angle of the late sound energy respectably. So by looking at these factors can give an insight on how the envelopment is changed by the different musical content.

The musical content could also give arise to different spectral differences on the late reflections, since the type of instrument or sound played in a concert hall contain certain frequency characteristics. So the use of different sound material with short notes and continuous sounds would therefore be a good choice for the investigation of the perceived LEV in a concert hall.
4. Method

This work will investigate to what degree different sidewall materials alter the spectral content of the late lateral reflections in a real concert hall and how these late lateral reflections affect perceived listener envelopment. Four side wall materials and the four music stimuli were chosen to cause alterations in LEV. A listening test was implemented with auralized stimuli including HRTF information. Auralization was done in a computerized model of the Studio Acusticum concert hall in Piteå, Sweden, which was modeled in CATT Acoustic Software.

This research method has its advantages and disadvantages. First it takes the question of perceived envelopment towards a real situation but not to the extent where it becomes uncontrollable, from a scientific point of view. The acoustical parameters in the model can be controllable to a certain extent then it would be possible in a real live situation. The model can however only estimate the reflections, since the model is based on octave band filters that are added up together [20].

4.1 Selecting Sidewall material

The sidewall materials include the original material in the concert hall (Co = Concrete), two computer modelled extremes (T1 = test material 1, T2 = test material 2) and one measured material from CATT Acoustic Library (F = fiber-rigid back). The modeled extremes where chosen to attain two set of spectral “extremes” were only low- and high- frequency were absorbed respectively.

The original material of the wall was chosen because of statements from listeners which have noticed a presence of LEV in the real concert hall. The fourth material was chosen to examine how a highly absorptive material in the mid range would change the perception of LEV.

The absorption- and scattering- coefficients of the materials used are showed in table 8A and 8B. The reverberation time in the concert hall with each wall material, measured with T-30, can be seen in figure 15. The C80 values, early to late sound ratio, from each model are shown in table 9.

---

3 “Auralization is defined as the process of rendering audible, by physical or mathematical modelling, the sound field of a source in a space as to simulate the binaural listening experience at a given position in the modelled space.” [18].

4 Head related transfer function (HRTF) is the acoustic filtering that free field sounds, arriving from different angles, undergoes because of the head, torso and pinna [19].

5 T-30 is derived from a straight-line least-square fit to the receiver curve at the interval -5 to -35 dB. This measurement is considered to be the best estimated of reverberation time [20].
Table 8A: Absorption coefficients, and scattering coefficients of the chosen sidewall materials, specified from 0-100 %.

<table>
<thead>
<tr>
<th>Material</th>
<th>125 Hz</th>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1 k</th>
<th>2 k</th>
<th>4 k</th>
<th>8 k</th>
<th>16 k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concrete, original material (Co)</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fiber-rigid back (F)</td>
<td>8</td>
<td>32</td>
<td>79</td>
<td>93</td>
<td>87</td>
<td>80</td>
<td>73</td>
<td>66</td>
</tr>
<tr>
<td>Test material 1 (T1)</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>20</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Test material 2 (T2)</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>70</td>
<td>80</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 8B: Scattering coefficients, test material 1, 2 and Fiber-rigid back are set at default in CATT Acoustic

<table>
<thead>
<tr>
<th>Material</th>
<th>125 Hz</th>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1 k</th>
<th>2 k</th>
<th>4 k</th>
<th>8 k</th>
<th>16 k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concrete, original material (Co)</td>
<td>10</td>
<td>12</td>
<td>13</td>
<td>15</td>
<td>16</td>
<td>18</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>Fiber-rigid back (F)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Test material 1 (T1)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Test material 2 (T2)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>
4.2 Creation of stimuli

The types of music stimuli were chosen on the premises that a variety of different musical stimuli can give excite different frequencies in the concert hall which can affect the perceived LEV. The music stimuli that were chosen was a violin playing legato, crash cymbal, male vocals singing in staccato and a percussive nylon acoustic guitar playing a staccato. The frequency plots of each music stimulus can be seen in appendix B. The purpose of the experiment is not to investigate the stimuli/musical contents affect on LEV, but merely to excite the concert hall in different ways.

The stimuli were recorded in a recording studio at Luleå University of Technology. The studio attained variable acoustics and was set at maximum absorption for minimizing reflections under the recording, for complete information about recording equipment see appendix C. After the recording of the stimuli, calibration was carried out. The stimuli were convolved with HRTF impulse responses, attained from CATT Acoustics, one convolution for each wall material. The HRTF model was chosen for creating the sensation of being in the...
real environment and for creating externalization of the sound around the head, instead of the in head localization which is present at normal listening with headphones [19].

4.3 HRTF and Auralization problems
Because of the choice of this method the author would like to address some difficulties with using HRTF and auralization of a computer model of a concert hall, and how these difficulties have been treated prior to the listening test.

Problems with HRTF can occur in practice when listening with headphones. The subjects could perceive that there is a lack of “presence” in spatially simulated sounds. Sounds spatialized near the median plane (0°) could be perceived as “inside” the head instead as “outside” the head. Sounds which are processed to arrive from the front of the listener could be perceived as arriving from the back. To simulate sounds with nonzero elevation are difficult, since the HRTFs are different from person to person. Also by using a generalized set of HRTFs could give the listener a perception of the sound that was not intended for that spatial location [19]. To eliminate the “inside the head sensation” the position of the receiver in the model are placed of axis from the median plane. The other practical problems are hard to get round and could be addressed as errors in the test.

The difference between binaural recording with a dummy head (BRD) and auralization with HRTF (AH) is that the BRD contains more information and fluctuations in the frequency range. The BRD are sounding more “natural” than the AH [21]. Saher et al [21] concluded that AH could recreate an acoustical environment at a level between “rather different” and “slightly different” to a real environment. Auralization could be used a strong tool for assessing an acoustical environment before it is built. The computer model was chosen in spite of these problems for the simple matter of replacing the materials of sidewalls more easily than it would be in the real environment [21].

4.4 Calibration of stimuli and listening level
Calibrations of the stimuli were conducted in two steps; before and after convolving the stimuli. The first calibration was done after recording the stimuli to get a perceived equal sound level of the recordings. This was done by creating a session in a digital audio workstation where one stimulus represented one channel fader each, at the zero position. The five listeners then adjusted each fader so the stimuli were perceived as equal in strength. Each fader position was logged after each session. The averages were then calculated for each stimulus and used for calibrating the stimuli before the convolution.

The second calibration was done after the convolutions, to attain equal true pressure signal at the receiver position. CATT Acoustics used the scale factor for calibrating the true pressure to attain the same value of each stimulus. The first level calibration could have been affected due
to this, however this latter calibration was necessary due to the fact that the CATT Acoustics auto scales each stimulus to attain maximum dynamic range at each convolution. The calibration could only calibrate stimuli attaining the same length and musical content. A total of four calibrations were done, one calibration for each set of stimuli containing the same musical content [20]. The limitation of only calibrating stimuli with the same musical content is not largely important, due to the listening test design. The subjects are only evaluating envelopment between the computerized models and not between the musical content.

Table 11: Gain calibration of the different stimuli completed by CATT Acoustics.

<table>
<thead>
<tr>
<th>Calibration level CATT</th>
<th>Guitar</th>
<th>Violin</th>
<th>Cymbal</th>
<th>Vox</th>
</tr>
</thead>
<tbody>
<tr>
<td>+22.2dB</td>
<td>+22dB</td>
<td>+11.6dB</td>
<td>+6.7dB</td>
<td></td>
</tr>
</tbody>
</table>

The listening level for the listening test was attained at the first calibration. The same five listeners got the freedom after the calibration to listen to one stimulus and adjust the headphone amplifier to their preferred listening level, this setting was logged by the author. Because of the lack of markers on the volume control it had to be logged by drawing the volume control setting on a paper for each listener. It showed that most of the listeners choose a similar volume setting. The midpoint of extremes chosen was used for the test, see figure 16. The soundcard had to be changed from the first calibration session to the listening test. To compensate for this change the author compared the output level from each soundcard, by listening with the same headphone volume setting and matched these perceived output levels.

Figure 16: a = lowest chosen listening level, b = highest chosen listening level, C = midpoint used under the listening test.
4.5 Creation of impulse responses and the auralization process

The auralization process and impulse responses creation were attained in a three step process. First the sidewall materials were changed in the Planes.geo file code [20]. The listener position, head direction and sound source position were changed from the original position, attained from the original model, to new ones, figure 17, table 12. These positions were chosen to avoid effects such as front and back confusion and “inside the head” sensation of the sound, as discussed earlier in the method section.

![Figure 17: Source and Receiver positions in the computer model; 05 = receiver, head direction stage; A0 = Source, direction upward.](image)

<table>
<thead>
<tr>
<th>Coordinates for positions</th>
<th>Axis</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receiver position [05]</td>
<td>-0.5</td>
<td>17</td>
<td>3</td>
<td>Meter [m]</td>
</tr>
<tr>
<td>Source position [A0]</td>
<td>0.0</td>
<td>4.5</td>
<td>1.5</td>
<td>Meter [m]</td>
</tr>
</tbody>
</table>

The second step was to set room temperature and other physical settings in which the full detailed calculation was taken into account, see figure 18 for more information on the full detailed calculation settings. After applying these settings then the full detailed calculation was carried out.
Figure 18: Full detailed calculation settings from CATT acoustics, step 1-3 is settings done prior to step 4 which is the full detailed calculation menu.

The third step was to create an impulse response from the data from the full detailed calculation and convolve the recorded stimuli with the impulse response. This was due to attain the different impressions of the concert hall properties on the stimuli. This step is called Post processing in the software. Full resolution HRTF model, compensation of headphones, length of impulse response was selected before creating the impulse response. Convolution was then conducted with CATT Acoustics own convolution apparatus. For a full set of settings for the post processing see appendix D. The compensation for the headphones could be a source of error in the test. The headphone model used for compensation in CATT Acoustic (STAX Lambda Pro) is not the same as used in the listening test (STAX Lambda Nova). This was however the only way to go since the need for high fidelity headphones in the listening test was critical, and the only compensation for STAX headphones was for Lambda Pro.
4.6 Preliminary study
A preliminary evaluation of the different types of stimuli created was carried out for the purpose of seeing if the difference in wall material does change the perception of LEV. A listener was used in the preliminary study and the evaluation was carried out verbally between the author and the listener. The evaluation showed that the stimuli contained tendencies of differences in envelopment, such as reverberation time and frequency characteristics, however this is only indications. These results gave further hope that there are influences of the sidewall material on LEV.

4.7 Listening test
The purpose of the test was to investigate if there was a change in the perception of listener envelopment caused by the varied sidewall material. The listeners were told to evaluate the amount that they felt inside/enveloped by the sound image.

4.7.1 Subjects
The listening test included 21 subjects, all with some sort of education in sound evaluation.

4.7.2 Test method
The test was divided into two and three steps for the listeners. The inconsistency of steps was because of problems on formulating questions for the form. For 11 of the listeners (the three step test) a form was given at the beginning of the test. This form instructed the listener to write down all the perceived differences between one pair of stimuli. These pairs were chosen by the author and contained the original wall material and one of the three other materials used in the test. Each listener got a set of combinations A versus B, C or D containing the same musical content; vocal or violin. A total of 11 comparisons were attained. This step of the test was to see if any other differences were perceived by the listeners other than envelopment. The question stated was the following;

- Write down the perceived differences between X and X. Example; frequency content, long/short reverberation time and so forth.

The second step in which all the 21 listeners participated in was a training session. This session enabled the listeners to listen to each one of the instruments/musical contents before conducting the primary test.

Step three was the primary listening test. Eight trials of evaluations were conducted by each listener. Each trial contained four versions of one instrument/music content and each stimulus contained one version of the sidewall materials. The fifth to eight evaluations was repetitions of the first four evaluation trials. The task presented to the listeners were the following; “Grade the Perceived amount of Envelopment in the presented stimuli”. The trials were in random order to eliminate and avoid systematic errors in the test, see table 13 for an example of the evaluations.
Table 13: Example of evaluation matrix for four evaluations.

<table>
<thead>
<tr>
<th>Material</th>
<th>Original</th>
<th>Test material 1</th>
<th>Test material 2</th>
<th>Fiber-rigid back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus 1</td>
<td>Cymbal</td>
<td>Cymbal</td>
<td>Cymbal</td>
<td>Cymbal</td>
</tr>
<tr>
<td>Stimulus 2</td>
<td>Male singing (Staccato)</td>
<td>Male singing (Staccato)</td>
<td>Male singing (Staccato)</td>
<td>Male singing (Staccato)</td>
</tr>
<tr>
<td>Stimulus 3</td>
<td>Nylon Acoustic Guitar</td>
<td>Nylon Acoustic Guitar</td>
<td>Nylon Acoustic Guitar</td>
<td>Nylon Acoustic Guitar</td>
</tr>
<tr>
<td>Stimulus 4</td>
<td>Violin (legato)</td>
<td>Violin (legato)</td>
<td>Violin (legato)</td>
<td>Violin (legato)</td>
</tr>
</tbody>
</table>

4.7.3 Data acquisition
The test had a modified MUSHRA test design, the grading was similar to the MUSHRA [22] test however a reference stimulus were not used. The listener could listen to all four versions of the stimuli in one set continuously, and grade each version on a scale from 0-100. The listeners were instructed verbally as well as in writing, what they were evaluating and how the evaluation procedure worked. They were given instructions that one of the four stimuli had to be 100 which corresponded to maximum envelopment in that particular set, see appendix E for instructions to the listeners. This grading was chosen to see how each musical content give arise to envelopment and how each sidewall material affected the envelopment with that particular musical content. However the text beside the grading scale in the software could not be eliminated, seen in figure 19, this text was explained to the listener to not be used when performing the grading.

4.7.4 Equipment and location of listening test
The listening test was conducted in a damped room for eliminating outside noise which could interfere with the listener’s evaluation. A wall of foam was placed in front of the listener’s position to further eliminate noises from the neighboring room.
The evaluation was carried through with circum-aural\textsuperscript{6} headphones, STAX Lambda Nova, and the software interface STEP was used which is designed for evaluating sound, figure 19. The equipment setup for the listening test contained two computers one laptop- and one stationary-computer, one USB SoundBlaster soundcard, and a headphone amplifier with headphones. In figure 20 and table 14 the setup and equipment can be seen.

\textbf{4.7.5 Equipment problems}

There occurred some complications during the experiment. The second day of running the experiment the laptop started to interfere with the listening, by “looping” the starting point of the sound. After a time of problem solving it was corrected by changing computer to the stationary computer, this correction made it possible to carry out the experiment. The foam wall came to a second use when the computer was changed. The stationary computer was placed behind the wall to eliminate the excess fan noise induced by the computer. The listeners who were exposed to the problem said after the test that it provoked an irritation, however in spite of that they could perform the evaluation.

\textbf{4.8 Analysis}

The analysis of the listening test data was conducted in several steps. The first step was to investigate if there were any outliers in the data. This was done by taking the difference from each listeners grading from the first trials (trial 1-4) and the repetitions (trial 5-8), and looking at the standard deviation of each listeners difference score between trial 1-4 and trial 5-8. The standard deviation of the difference can be seen in figure 21. There are several large differences in the standard deviation of the difference between the listeners. This could be because the listeners used different sets of scales at each trial, thus increasing or decreasing

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Model & Type of equipment & Operating system \\
\hline
Acer aspire 3500 & Laptop computer & Windows Xp \\
Dell & Stationary computer & Windows 2000 \\
USB SoundBlaster & USB Soundcard & \\
STAX Lambda Nova & Headphones & \\
STAX & Headphone amplifier & \\
\hline
\end{tabular}
\caption{Equipment setup.}
\end{table}

\textsuperscript{6} A circum-aural model encloses the pinna of the ear and seats upon the surface of the head surrounding the pinna \cite{23}. 

Figure 20: Listening test setup.

Table 14: Equipment setup.
the standard deviation between the rounds. It could be due to inconsistency in the perception between the first and second rounds, which causes the differences.

4.8.1 Post screening
No outliers were excluded, based on the premise, if one listener varying the most were excluded the next listener who have the largest standard deviation would also be excluded and so forth. This could be risky to do, since eliminating listeners in this way could give misleading results in the ANOVA analysis. But also this large difference could be a first indication of inconsistency between the two rounds, round 1(trial 1-4) and round 2 (trial 5-8).

![Std differance](image)

*Figure 21: The standard deviation of the difference between trial 1-4 and 5-8 of each listeners score.*

Since each listener was instructed to grade one stimuli as 100 in the test, different scales were attained at each trial.

4.8.2 Normal distribution
Since there were 4 stimuli (different musical content) and 4 sidewall materials, a set of 16 Shapiro-Wilks tests were preformed with a confidence limit of 99 % for each round, respectively. The normal distribution test is for only analysing if the data is normal distributed enough to be used in the ANOVA analysis.
4.8.3 ANOVA analysis
To determine if there was an inconsistency between the rounds 1 and 2, the rounds were implemented as a factor in the ANOVA analysis. For eliminating the effect of using different scales, the listeners were implemented as a factor in the ANOVA analysis. The major aim of the ANOVA analysis was to see if statistical differences existed in the perception of envelopment, for each one of the four sidewall materials depending musical content and not the perceived difference of the instruments. Both the sidewall materials and instrument were implemented as factors in the analysis.

Four factors in the ANOVA analysis
- Material (Concrete, fibre-rigid back, test material 1 and test material 2)
- Instrument (Acoustic guitar, cymbal, male vocal, violin)
- Round (Round 1 and 2)
- Listener (21 subjects)

The null- ($H_0$) and alternative hypotheses ($H_1$) for the ANOVA analysis were formed as:

\[ H_0: \text{There is no change in the perception of listener envelopment caused by the different sidewall material.} \]
\[ H_1: \text{There is a change in the perception of listener envelopment caused by the different sidewall material.} \]

If the ANOVA test shows statistical significant differences, interactions will be searched for. After these two investigations a post-hoc test, Tukey HSD, will be employed on the sidewall materials to see between which materials there are a statistically significant different at the 95% confidence level [24].

4.8.4 Comparison in writing
The last analysis was a compilation of the perceived differences attained from written comparison. This analysis was done to give a further insight in how the listeners perceived the different stimulus. The written answers were translated into English from Swedish and inserted into a table, see table 18. The answers were analysed by comparing each answer from one listener to another, which have the same comparison combination i.e. answer concrete to test material 1 compared with answer concrete to test material 1 and so forth.

5. Results
In this section are the ANOVA analysis, post-hoc test and the written comparison presented. The ANOVA table shows the significant difference for each factor and the statistical significant interactions between them. The post-hoc test shows what the significant differences are. The written comparison section presents the translated answers. The results from the alternative analysis are presented in the similar fashion.
5.1 Normal distribution

Table 15: Normal distribution test, Shapiro-Wilks' test at the 99 % confidence level

<table>
<thead>
<tr>
<th></th>
<th>Concrete</th>
<th>Fibre-rigid back</th>
<th>Test material 1</th>
<th>Test material 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Round 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cym</td>
<td>0,0000293401</td>
<td>0,0166598</td>
<td>0,0336059</td>
<td>0,0151649</td>
</tr>
<tr>
<td>Git</td>
<td>0,00177468</td>
<td>0,622468</td>
<td>0,0457611</td>
<td>0,0000000839</td>
</tr>
<tr>
<td>Vio</td>
<td>0,00002,78637</td>
<td>0,0313112</td>
<td>0,0019604</td>
<td>0,152556</td>
</tr>
<tr>
<td>Vox</td>
<td>0,00319793</td>
<td>0,0368026</td>
<td>0,413831</td>
<td>0,00000198389</td>
</tr>
</tbody>
</table>

Can be rejected at 99 % confidence limit as not normal distributed

Table 15 shows that 43,75 % can be rejected as not normal distributed and 56,25 % can not be rejected as not normal distributed in round 1, 43,75 % be rejected as not normal distributed in round 2 and 56,25 % can not be rejected as normal distributed at the 95 % confidence level respectively.

The concrete material can be rejected as not normal distributed for all combinations. This result could be an effect of the listeners similar grading of that particular sidewall material.

5.2 ANOVA table

Table 16: Analysis of Variance for LEV - Type III Sums of Squares.

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F-Ratio</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MAIN EFFECTS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: Instrument</td>
<td>1504,5</td>
<td>3</td>
<td>501,501</td>
<td>1,40</td>
<td>0,2411</td>
</tr>
<tr>
<td>B: Material</td>
<td>39128,1</td>
<td>3</td>
<td>13042,7</td>
<td>36,49</td>
<td>0,0000*</td>
</tr>
<tr>
<td>C: Round</td>
<td>0,18006</td>
<td>1</td>
<td>0,18006</td>
<td>0,00</td>
<td>0,9821</td>
</tr>
<tr>
<td>D: Listener</td>
<td>81336,0</td>
<td>20</td>
<td>4066,8</td>
<td>11,38</td>
<td>0,0000*</td>
</tr>
<tr>
<td><strong>INTERACTIONS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td>2527,74</td>
<td>9</td>
<td>280,86</td>
<td>0,79</td>
<td>0,6296</td>
</tr>
<tr>
<td>AC</td>
<td>368,814</td>
<td>3</td>
<td>122,938</td>
<td>0,34</td>
<td>0,7935</td>
</tr>
<tr>
<td>AD</td>
<td>13104,2</td>
<td>60</td>
<td>218,404</td>
<td>0,61</td>
<td>0,9902</td>
</tr>
<tr>
<td>BC</td>
<td>1744,08</td>
<td>3</td>
<td>581,359</td>
<td>1,63</td>
<td>0,1823</td>
</tr>
<tr>
<td>BD</td>
<td>31814,9</td>
<td>60</td>
<td>530,249</td>
<td>1,48</td>
<td>0,0141*</td>
</tr>
<tr>
<td>CD</td>
<td>5271,35</td>
<td>20</td>
<td>263,568</td>
<td>0,74</td>
<td>0,7879</td>
</tr>
<tr>
<td><strong>RESIDUAL</strong></td>
<td>174765,489</td>
<td>357,393</td>
<td>0,74</td>
<td>0,7879</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL (CORRECTED)</strong></td>
<td>351565,671</td>
<td>671</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All F-ratios are based on the residual mean square error.
* denotes a statistically significant difference.

Table 16 shows statistical significant differences at 99 % confidence level for factors B and D, it also shows interactions at the 99 % confidence level between factors B and D. Interactions between listeners and sidewall material were found. This interaction was expected since the listeners were evaluating the different sidewall materials, so further analysis were not conducted on the interaction between listener and material.
5.3 Post-hoc test

Table 17: 95.0 percent Tukey HSD

<table>
<thead>
<tr>
<th>Material</th>
<th>Count</th>
<th>LS Mean</th>
<th>LS Sigma</th>
<th>Homogeneous Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>168</td>
<td>67.6667</td>
<td>1.45854</td>
<td>X</td>
</tr>
<tr>
<td>T1</td>
<td>168</td>
<td>72.3393</td>
<td>1.45854</td>
<td>X</td>
</tr>
<tr>
<td>T2</td>
<td>168</td>
<td>81.8155</td>
<td>1.45854</td>
<td>X</td>
</tr>
<tr>
<td>Co</td>
<td>168</td>
<td>87.0536</td>
<td>1.45854</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 18: 95.0 percent Tukey HSD

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Sig.</th>
<th>Difference</th>
<th>+/- Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co - F</td>
<td>*</td>
<td>19,3869</td>
<td>5,31704</td>
</tr>
<tr>
<td>Co – T1</td>
<td>*</td>
<td>14,7143</td>
<td>5,31704</td>
</tr>
<tr>
<td>Co – T2</td>
<td></td>
<td>5,2381</td>
<td>5,31704</td>
</tr>
<tr>
<td>F – T1</td>
<td></td>
<td>-4,67262</td>
<td>5,31704</td>
</tr>
<tr>
<td>F – T2</td>
<td>*</td>
<td>-14,1488</td>
<td>5,31704</td>
</tr>
<tr>
<td>T1 – T2</td>
<td>*</td>
<td>-9,47619</td>
<td>5,31704</td>
</tr>
</tbody>
</table>

* denotes a statistically significant difference.

Figure 22: Plot of means and 95 % Tukey HSD intervals, for each sidewall material versus the LEV scores. Co= concrete; F = Fibre-rigid back; T1= Test material 1; T2 = Test material 2. For all stimuli/musical contents.

Table 17, 18 and figure 22 shows the sidewall materials that has statistical significant differences, these results are attained form the post hoc test. The statistical significant differences are between concrete material and fibre-rigid back, concrete material and test material 1, test material 2 and test material 1, fibre-rigid back and test material 2.

5.4 Comparison in writing

Table 19: Comparison in writing, the notation (example A to B) in the table is the stimuli name presented for the subject and the name below the notation is the sidewall material compared in the test, respectively.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Vox</th>
<th>Violin</th>
</tr>
</thead>
<tbody>
<tr>
<td>A to B</td>
<td>A, gives a broader spatial impression then B. A blossoms laterally, and B has a more focus in the depth information.</td>
<td>A is more compressed. Feels as a “traffic jam” in the middle of the stereo spectra. Harder sound, unpleasant. B feels more even and more enveloping.</td>
</tr>
</tbody>
</table>
A is wider but flatter, B is deeper but narrower.

<table>
<thead>
<tr>
<th>A to C</th>
<th>Concrete vs. Test material 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A; shorter reverberation time then C, +50 ms.</td>
<td>C is more sharper in the treble and upper midrange then A.</td>
</tr>
<tr>
<td>A is wetter reverb then C.</td>
<td></td>
</tr>
<tr>
<td>A; longer pre delay, +25 ms.</td>
<td></td>
</tr>
<tr>
<td>A; sounds as a bigger room then C.</td>
<td></td>
</tr>
<tr>
<td>C; some ….not interpretable… frequency “hump” at approximately 6 kHz.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A to D</th>
<th>Concrete vs. Test material 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I experience a very small difference between the sounds. Sound D feels a bit sharp and metallic with a slight distortion in the strong parts. Sound D feels bigger and at the same time more assertive.</td>
<td>It was very hard to hear any difference at all, but I thought after a while that the reverberation time did differ.</td>
</tr>
<tr>
<td>D = Warmer and little darker then A, thus not as much of midrange.</td>
<td>D = Higher resolution</td>
</tr>
<tr>
<td>D = appealing and a bit longer reverberation time</td>
<td>I prefer D!</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A to B</th>
<th>Concrete vs. Fiber-rigid back</th>
</tr>
</thead>
<tbody>
<tr>
<td>A can be perceived as harder on the vocals. A is a bigger room with more bonuses, but the difference that I perceive can be imaginary due to the repeated listening. They are very similar.</td>
<td>I perceive B as much more clear because of the reverb dose not “smear out” the attack; less early reflections and/or more damping in the room. A; swells out the direct sound and gives a smeared almost unpleasant impression. At several times I’ve perceive that certain frequencies “comes thru”- they amplifies strongly by the reverberation in a way that I feel undesirable. This phenomenon is clearly less percent in stimulus B.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A to C</th>
<th>Concrete vs. Test material 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus A has a wider reverb and gives a feeling of bigger envelopment. I get the illusion that the reverb is longer then in stimulus C, however it can not be established. After listening several times are the perceived the reverberation time not different between the stimuli. I get the feeling of being in a room with larger distance between left and right wall in stimuli A and the distance between the walls in stimuli C is smaller.</td>
<td>I perceive that C has a little bit of nasal sound, which made it according to me easier to judge the distance of the direct sound. C felt more natural. If I closed my eyes and listen, then it felt easier to visualise that I was present in a concert hall and listen to the sound, then it were with example A, where the reverberation was to big and dominant.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A to D</th>
<th>Concrete vs. Test material 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. seems to be sharper then D. D feels more mellow and has a more pleasant reverberation. The differences are however small.</td>
<td></td>
</tr>
</tbody>
</table>

There were several perceived differences attained from the use of different musical contents for the same sidewall material but also similarities. What these are is presented in the analysis of the written comparison.
6. Analysis

The normal distribution tests for round 1 and 2 indicates that 56.25 % and 43.75 % could be normal distributed for round 1 and 2 at a 99 % confidence level, respectively, see table 15.

The ANOVA table showed that two of the four factors have significant effects on LEV (<P=0.0001), material and listener respectively. It is also shown that there is a statistical significant interaction between the listener and material, no other interaction was found.

The Tukey HSD test at 95 % confidence level performed on the factor B (material) showed that there is a significant difference between four combinations of sidewall materials. No other statistical significant difference was found.

- The significant differences were between concrete and fiber-rigid back, concrete and test material 1, fiber-rigid back and test material 2, test material 1 and test material 2. The non significant differences were concrete and test material 2 and fiber-rigid back and test material 2.

The comparison in writing between the different sidewall materials (fiber-rigid back, test material 1 and test material 2) versus the original sidewall material (concrete) showed some differences in the perceptions depending on what musical content used. The term impression used in the analysis is referred to the perceived judgment of the hall.

- Concrete versus Fiber-rigid back
  Vocal as a stimulus: The listeners used words to describe the impression of the concrete as bigger, blossoms laterally, wider, more spatial impression, flatter, hard to the vocals. For the fiber-rigid back the listeners used words as, focused depth and narrower.

  Violin as a stimulus: The impression of the concrete were describe with words as hard sound, compressed, unpleasant, swells out, smears out (unpleasant), amplifying certain frequencies, a lot of frequencies in the middle of the stereo spectra. The impression with the fiber-rigid back were described as, clearer, does not smear out the attack, less early reflections (more damped), more enveloping, more even in frequencies.

Here there are clear differences between the stimuli used for vocal stimulus. The concrete are perceived wider and has a bigger sensation of spatial impression where the fiber has a more focused depth but are narrower. With the violin as a stimulus made the impression of the concrete; hard, unpleasant, largely diffuse and amplifies certain frequencies. The fiber-rigid back gives an impression of the sound; more clearly, more enveloping containing less early frequencies and more damped than the concrete material.

- Concrete versus Test material 1
  Vocal as a stimulus: The listeners described the impression of the hall with concrete through words as, shorter reverberation time, more sensation of the reverb, longer pre delay, bigger room, wider reverb, bigger envelopment, distance from the sidewall are larger then for test material 1. The impressions of the hall with test material 1 were described as, a smaller distance for the sidewalls and frequency “hump” at 6 kHz.
Violin as a stimulus: The impressions of the concrete were perceived to contain a big and dominant reverb. The impression with fiber were described as more natural, easier to perceived a concert hall, had sharper upper midrange and treble, has a nasal sound which makes it easier to perceive the distance of the direct sound.

Here there is not an obvious perceived difference between the uses of musical content as it was in concrete versus fiber-rigid back. There are similarities in the perceptions of the concrete sidewall material in spite of the use of different stimuli, which is related to the amount of reverberation. However it can not be concluded if this is a positive of a negative quality. For the test material 1 the similarities between musical content are based on the frequency content of the impression, where the high frequencies were perceived larger in both comparisons versus the concrete material.

- **Concrete versus Test material 2**
  Vocal as a stimulus: The listeners indicated that there were small differences between the concrete and test material 2. But described the impression of the test material 2 as, sharper, more metallic sounding with distortions in strong parts and bigger but more assertive then the impression of the concrete material.

  Violin as a stimulus: There is also indicated when using a violin as stimulus that there is a small difference between the sidewall materials. The concrete material was described as sharper than test material 2. The impression of the test material 2 were perceived, warmer, containing a low amount of midrange, have longer reverberation time, has a high resolution, are appealing, is more mellow and pleasant, than the impression of the concrete material.

The impression between the concrete and test material 2 are shown to be small, however it is shown that there is a difference in impression between musical content used. Using the vocal stimulus makes the impression of the test material 2 to be perceived hard, metallic at the same time as it is big and assertive. The impression of the violin also becomes warmer and appealing. The concrete material makes the impression of the violin stimuli sharper than the test material 2.
7. Discussion

7.1 Results

When comparing the results from the plot of means, figure 22, attained from the first ANOVA analysis with the reverberation times in figure 15, it can be seen several similarities. The concrete material has the largest reverberation time of the four materials, and is perceived as most enveloping.

The similarity between the reverberation time of test material 2 and the results in the plot of means is that the reverberation time and perceived LEV are placed second respectively. However there are not a statistical difference between the concrete material and test material 2. This could be due to the similarity between the reverberation times of these two wall materials.

Another interesting finding when comparing reverberation time and the plot of means is that test material 1 and fiber-rigid back material have short reverberation times at high frequencies bands than the concrete material. This could be why these two materials are significantly different in perceived LEV in the plot of means. This is also the case for the test material 2 versus test material 1 and the fiber-rigid back material. Also here there are significant differences. However the test material 2 has a shorter reverberation time than test material 1 at the 4000 Hz band. The latter finding could be an answer to why the test material 2 is perceived lower in LEV than the concrete material. But since there is no statistical difference between them this are just speculations.

The length of the reverberation times from each hall model shows that fiber-rigid back and test material have over all shorter reverberation times than the concrete and test material 2. This could be an explanation of the results from the ANOVA analysis. The earlier work shows that an increase of the total amount of reverberation time or increase in separate frequency bands increases the LEV. The results from this experiment can also confirm this.

The C\textsubscript{80} values in table 9 gives also further knowledge to the results attained from the ANOVA analysis. The concrete material has the lowest amount of C\textsubscript{80} values for all frequency bands, followed by test material 2. For the high frequency band, 4000 Hz, the C\textsubscript{80} value increases for all materials. The interesting thing here is the C\textsubscript{80} value for the low frequency bands.

The concrete and test material 2 have the lowest amounts of C\textsubscript{80} values of the four sidewall materials at frequency bands 125-500 Hz. This could be an explanation for the attained results from the ANOVA and why both the concrete and test material 2 are perceived as more enveloping. When relating these results to the earlier work it can be seen that the C\textsubscript{80} values for low frequency bands are also important on the perceptions of LEV in this test.

In the written comparison several differences between instruments/musical content and sidewalk materials were perceived. First there was a perceived difference between the materials concrete and the fiber-rigid back, with the vocal stimuli. The concrete were perceived as wider and attained a bigger sensation of spatial impression and the fiber-rigid back had a more focused depth but was perceived narrower. Also a difference was perceived between the sidewalk materials when the violin was used as a sound source. The concrete was perceived as hard, unpleasant, and largely
diffuse. The violin in the fiber-rigid back model was perceived more clearly and more enveloping.

The difference between materials with the vocal stimulus can be described by the physical measurement attained from each computer model, where the $C_{80}$ is lower for the concrete and higher for the fiber-rigid back material. The interesting result is attained when the violin is used. Here it is hard and unpleasant for the concrete material and more clear and enveloping for the fiber-rigid back material. A proposed answer to this is that the violin’s frequency spectra between 500-1000 Hz, B 2 in appendix B, are strong in amplitude and the $C_{80}$ values for the concrete material are low in these particular frequency bands which makes the sound more diffuse and not well perceived. For the fiber-rigid back the $C_{80}$ values for the frequency bands 500-1000 Hz are much higher, which is due to the high absorption in the midrange which diminish the late arriving sound. The hardness of the sound could be due to the reverberation times at the higher frequencies. These reverberation times are much longer for the concrete material than for the fiber-rigid back.

There is no obvious difference between concrete material and test material 1 in spite of the change of instrument/musical content. For the concrete material the amount of reverberation was perceived as large for both instruments/musical contents. An explanation for this can be because of the larger reverberation time attained in the concrete model, when comparing it to the model of test material 1. The similarities between the violin and vocal stimuli on test material 1 were related to the large amount of high frequencies in the model versus the concrete model. This result is not hard to understand since test material 1 is designed to absorb only low frequencies and thereby increasing the perception of reflective high frequencies. These results indicate that reflected high frequencies can be perceived.

The perceived difference between the concrete material and the test material 2 are small, even when the instrument/musical content is changed indicated by the written comparison. This indicates that there is a small difference between concrete and test material 2 even though the statistical data indicates no difference at the 95 % confidence level.

There are however perceived differences between the sidewall materials depending on the instrument/musical content used. For the test material 2 the perception of the vocal becomes hard and metallic at the same time big and assertive. The violin becomes however warmer and appealing. For the concrete material the perceived of the violin becomes sharper than for test material 2.

The hardness, sharpness and metallic impression of the vocal and violin can not be easily explained here. However a proposed answer could be that both sidewall materials have large $C_{80}$ values at 4000 Hz in which both instruments/musical contents have large amounts of energies in, especially the vocal, figure 28. The perceived warmth of the violin can not be explained for the test material 2 other than that the test material 2 absorbs high frequencies in which the listener perceives as warmer.

The results from the written comparison are interesting since many of the differences can be explained by the physical factors used for determining the perceived envelopment. What is also interesting is that there are perceived differences between the different instruments/musical contents, but few used the term envelopment when describing the differences, this could be because envelopment is not the first thing that the subjects perceives. No statistical interaction between musical content and LEV were found in the
ANOVA analysis, this could be because there are no differences between them, and another test method could give a clearer view. However, the major concern of this work was on the perceived difference in listener envelopment caused by the sidewall material.

The normal distribution tests for round 1 and 2 showed that 17 of the combinations are normally distributed at a 99% confidence level. This small amount of normally distributed data could create a bias in the analysis, where the statistical difference that is attained could be due to the difference between normal distributed and not normally distributed data.

7.2 Method
The test method for attaining perception data on LEV was a listening test based on the MUSHRA method but without the hidden reference. One of the disadvantages of using this method was that one stimulus needed to be 100 at each trial and the listener would rate the other stimuli based on this. This creates a different set of scales at each trial and listener. This problem needs to be addressed with normalization or implementing the listener as a factor in the ANOVA. A different test design where the listener only could choose between four grades of LEV could have given different sets of data, which don’t need to be normalized or be implemented as a factor in the ANOVA. Another disadvantage of the test was the layout of the listening test computer program. Beside the slider in each trial was a set of scales: excellent, good, fair, poor, and bad. These could not be removed from the layout. This could have confused the listener negatively since different scales were used in the listening test, then those presented. The room in which the test was conducted had leakage in terms of outside noises, regardless of the foam wall present in the room, and this could have affected the listeners’ evaluation.

The test method used has its advantages since the approach with auralization of impulse responses attained from computer models of concert halls lies in the grey zone between the previous researches conducted in anechoic chambers and real concert halls. This approach gives more control of what is happening to the sound than it would be possible in a real concert hall. It gives though the opportunity to test previous finding to see if they hold in a more realistic situation. The approach has some disadvantages as presented under the section method. Such as, that spatially simulated sound only can give a sensation between “rather different” and “slightly different” than a real concert hall could give. The use of HRTF has also restraints such as; front to back confusion, “in the head” localization and lack of presence of spatially simulated sounds. Since a majority of the listener didn’t commented that the stimuli sounded artificial, the author hopes that these results can still shed light on perception of LEV.

These results could be verified or falsified by conducting a similar test in other models of concert halls or even changing the sidewall materials in a real concert hall and conducting listening test in it. However, results from previous research coincide well with these results attained here, which gives validity to the test. It would be good to conducting the same test for several different concert halls for checking if similar results can be attained as in this study.
8. Conclusions

The study shows that there are statistical perceived differences on listener envelopment when changing the sidewall material in this particular concert hall. The concrete material is shown to create the largest sense of listener envelopment versus the fiber-rigid back material and test material 1, which absorbs low mid frequencies and low frequencies, respectively. This pattern is also shown for test material 2 which absorb high frequencies. The test material 2 was perceived as highly enveloping versus the fiber-rigid back material and test material 1. These results can be explained with earlier research on listener envelopment where early to late sound ratio, C\textsubscript{80}, and different reverberation times at different frequency bands were shown of affect the perceived listener envelopment. These results indicates that high frequency absorbing and non-absorbing materials has an advantage over materials absorbing low and low-mid and mid-frequencies, when creating listener envelopment in this particular concert hall.

The use of different instruments/musical content could have an effect on listener envelopment. The written comparison showed perceived differences between the concrete material and fiber-rigid back depending on the instruments/musical content used however few of the subjects used the term envelopment when describing the perceived differences. These differences could be explained also by the earlier research where the type of stimuli gives different reflective patterns, depending on spectral content of the stimulus and how the instruments/musical content are played.

9. Further work

Since this work are only conducted in one model of a concert hall, a proposed further work would be to conduct this experiment in a different models of concert hall’s to see whether or not the results can be replicated. This work gives also indications that the instrument/musical content affects listener envelopment. This factor needs to be addressed in a separate experiment, since it was difficult to attain clear results from the presented work. Also few subjects mentioned envelopment when they described the perceived difference between the sidewall materials, and this is a interesting result which further work could be conducted on, it can be that envelopment is not the first thing that the listeners is influenced by or even aware of when listening on music in a concert hall.
References


ISSN: 1402-1544; ISRN: LTU-DT--02/17--SE


[9] Wakuda, Akiko; Furuya, Hiroshi; Fujimoto, Kazutoshi; Isogai, Kenji; Anai, Ken. Effects of arrival direction of late sound on listener envelopment.

[10] Soulodre, Gilbert A.; Lavoie, Michel C.; Norcross, Scott G. Investigation of Listener Envelopment in Multi channel Surround System”.
*113th AES Convention*, Pre print 5676, Los Angles, 2002


[16] Furuya, Hiroshi; Fujimoto, Kazutoshi; Wakuda, Akiko; Nakano, Yusuke. The influence of total and directional energy of late sound on listener envelopment. Acoustical Science and Technology, Vol. 26, No 2, pp. 208-211. 2005


Appendix A


<table>
<thead>
<tr>
<th>Stimulus</th>
<th>125 Hz</th>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1 k</th>
<th>2 k</th>
<th>4 k</th>
<th>8 k</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.50</td>
<td>0.63</td>
<td>0.75</td>
<td>0.85</td>
<td>0.93</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>0.66</td>
<td>0.77</td>
<td>0.77</td>
<td>0.91</td>
<td>0.94</td>
<td>1.0</td>
<td>1.01</td>
</tr>
<tr>
<td>3</td>
<td>1.13</td>
<td>1.00</td>
<td>1.04</td>
<td>1.02</td>
<td>0.96</td>
<td>1.0</td>
<td>1.01</td>
</tr>
<tr>
<td>4</td>
<td>1.82</td>
<td>1.28</td>
<td>1.19</td>
<td>1.06</td>
<td>1.01</td>
<td>1.06</td>
<td>1.99</td>
</tr>
<tr>
<td>5</td>
<td>2.11</td>
<td>1.58</td>
<td>1.30</td>
<td>1.10</td>
<td>1.01</td>
<td>1.02</td>
<td>1.00</td>
</tr>
<tr>
<td>6</td>
<td>2.34</td>
<td>1.82</td>
<td>1.41</td>
<td>1.10</td>
<td>1.03</td>
<td>1.02</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Experiment 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.49</td>
<td>0.92</td>
<td>1.21</td>
<td>1.61</td>
<td>1.81</td>
<td>1.91</td>
<td>1.93</td>
</tr>
<tr>
<td>2</td>
<td>1.16</td>
<td>1.18</td>
<td>1.46</td>
<td>1.87</td>
<td>1.90</td>
<td>1.91</td>
<td>1.93</td>
</tr>
<tr>
<td>3</td>
<td>2.20</td>
<td>1.93</td>
<td>1.91</td>
<td>2.03</td>
<td>1.96</td>
<td>1.92</td>
<td>1.94</td>
</tr>
<tr>
<td>4</td>
<td>2.92</td>
<td>2.67</td>
<td>2.18</td>
<td>2.12</td>
<td>1.99</td>
<td>1.93</td>
<td>1.94</td>
</tr>
<tr>
<td>5</td>
<td>3.33</td>
<td>3.22</td>
<td>2.45</td>
<td>2.17</td>
<td>1.98</td>
<td>1.91</td>
<td>1.94</td>
</tr>
<tr>
<td>6</td>
<td>3.82</td>
<td>3.33</td>
<td>2.63</td>
<td>2.20</td>
<td>2.03</td>
<td>1.94</td>
<td>1.94</td>
</tr>
</tbody>
</table>
Appendix B

B 1: Frequency plot of mono recorded cymbal stimuli before convolution.

B 2: Frequency plot of stereo recorded guitar stimuli before convolution.

B 3: Frequency plot of stereo recorded violin stimuli before convolution.
B 4: Frequency plot of mono recorded vocal stimuli before convolution.
Appendix C

C 1: Recording of stimuli.

<table>
<thead>
<tr>
<th>Information Recording</th>
<th>Microphone</th>
<th>Distance from Mic</th>
<th>Room type</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cymbal</td>
<td>Line Audio CM3</td>
<td>15 cm</td>
<td>Damped recording room</td>
<td>Cardioid</td>
</tr>
<tr>
<td>Male singing (staccato)</td>
<td>AKG c414</td>
<td>25 cm</td>
<td>Anechoic chamber</td>
<td>Cardioid</td>
</tr>
<tr>
<td>Nylon Acoustic Guitar</td>
<td>Line Audio CM3, Stereo pair</td>
<td>12 cm</td>
<td>Damped recording room</td>
<td>Cardioid</td>
</tr>
<tr>
<td>Violin (Legato)</td>
<td>Line Audio CM3, Stereo pair</td>
<td>45-50 cm</td>
<td>Damped recording room</td>
<td>Cardioid</td>
</tr>
</tbody>
</table>
Appendix D

**STEP 1**

<table>
<thead>
<tr>
<th>General settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impulse response (IR)</td>
</tr>
<tr>
<td>Length: 3000 ms</td>
</tr>
<tr>
<td>Sample rate: 44100 Hz</td>
</tr>
<tr>
<td>Size: 12200</td>
</tr>
<tr>
<td>Reflections:</td>
</tr>
<tr>
<td>Transform size: 4096</td>
</tr>
<tr>
<td>Apply source exp/gain</td>
</tr>
<tr>
<td>Fractional delay</td>
</tr>
<tr>
<td>Stopband suppression:</td>
</tr>
<tr>
<td>Level: -40.0 dB</td>
</tr>
<tr>
<td>Suppress below 88 Hz</td>
</tr>
</tbody>
</table>

**STEP 2**

<table>
<thead>
<tr>
<th>Preferences, user: Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projects folder:</td>
</tr>
<tr>
<td>Surface properties: ABLIBOC.DAT</td>
</tr>
<tr>
<td>Default anechoic file: Rb_44_en.wav</td>
</tr>
<tr>
<td>CATT-Edit font/size: COURIER NEW 13</td>
</tr>
<tr>
<td>Sample rate, Hz</td>
</tr>
<tr>
<td>HRTF library: TAI_plain_44.DAT</td>
</tr>
<tr>
<td>Headphone: TAI_LAMBDAPRO_plain_44.DAT</td>
</tr>
<tr>
<td>Display warnings:</td>
</tr>
<tr>
<td>Use small Plot-file Control</td>
</tr>
<tr>
<td>Show Plot-file Control</td>
</tr>
<tr>
<td>Auto minimize inactive View-module</td>
</tr>
<tr>
<td>Disable screen saver while calculating</td>
</tr>
<tr>
<td>2DAIS History plot (instead of 2DAI)</td>
</tr>
<tr>
<td>Calculate CW/DC (instead of D-50/%)</td>
</tr>
<tr>
<td>Create 32-bit SIM-file (not for Lake)</td>
</tr>
<tr>
<td>8 and 16 kHz options:</td>
</tr>
<tr>
<td>Include in overview text-files:</td>
</tr>
<tr>
<td>Include in overview plot-files:</td>
</tr>
<tr>
<td>Include in map 'sum':</td>
</tr>
<tr>
<td>SPL range for map statistics: 20 dB</td>
</tr>
<tr>
<td>Initial 3D-balloon view angles:</td>
</tr>
<tr>
<td>Hor: 30 °, Ver: 30 °</td>
</tr>
</tbody>
</table>

D1: Post-processing steps conducted in CATT Acoustic, for creating the stimuli.
Appendix E

Instruction to listening test

Task: Grade the Perceived amount of Envelopment in the presented stimuli

You will be met by eight trials in the test. Your task at each trial is to grade the perceived amount of envelopment on four stimuli. This is done by adjusting faders on a scale 0-100, 0 is non-enveloping and 100 is most-enveloping.

One of the stimuli at each trial must be set as 100 = most-enveloping, before the trial is completed and you can go to the next.

Envelopment is defined as the amount that you feel inside/enveloped by the sound image, see figure below. The dots represent envelopment of the guitar sound and it is this that you are going to grade.