THE SECRET OF THE MUSIKVEREIN AND OTHER SHOEBOX CONCERT HALLS

T Lokki Aalto University School of Science, Department of Computer Science, Finland
J Pätynen Aalto University School of Science, Department of Computer Science, Finland
S Tervo Aalto University School of Science, Department of Computer Science, Finland
A Kuusinen Aalto University School of Science, Department of Computer Science, Finland
H Tahvanainen Aalto University School of Science, Department of Computer Science, Finland
A Haapaniemi Aalto University School of Science, Department of Computer Science, Finland

1 INTRODUCTION AND MOTIVATION

The Musikvereinssaal in Vienna is indisputably one of the most admired concert halls in the world. Even those who prefer defined and clear sound to overwhelming enveloping reverberation praise the acoustics of the Musikverein. Researchers have studied the acoustics of this golden hall, mostly by the objective measures, but still the superior acoustics remains unexplained. General consensus among acousticians seems to be that numerous strong early lateral reflections, high ceiling, and bright enveloping reverberation are the keys to the success. Here, we aim to explain the acoustics of shoebox concert halls, including the Musikverein, with the measurements of spatial impulse responses. Moreover, we propose to shift the research focus in concert hall acoustics from traditional energy decay based parameters towards aspects on how acoustics can support music and make it more enjoyable.

Although the Musikverein provides strong sensations for any listeners not all people like its overwhelming reverberation in blind listening. We have recently conducted a large listening test[8] that included preference judgements between six concert halls on three different seats with two different signals. The details of the measurements of halls and the multichannel sound renderings for the listening tests are explained in another paper of this conference[14]. The preference test results for 28 subjects, who compared each hall against each other twice with a paired comparison method, are presented in Fig. 1. The results reveal that listeners can be categorized into three preference groups, which have contrasting preference profiles. The largest group includes listeners who like shoebox halls, in particular with more tranquil music (Beethoven, 7th symphony, introduction). With full orchestra *forte* (Bruckner, 8th symphony, scherzo) this group still prefers shoebox halls, but the Musikverein (as measured unoccupied) is too reverberant for them, especially at balcony. The second preference group likes less reverberant and enveloping halls such as the Philharmonies in Cologne and Berlin. The last group generally prefers Konzerthaus in Berlin although the preference depends quite a lot on the seat. Essentially, these results clearly show the well-known fact that the preferred acoustics is a matter of individual taste and that the preferred order of halls might change according to the listened music.

Considering the reputation of the golden hall, the listening test results are surprising, as the Musikverein does not stand out in the indicated preferences. The experimental setup may not necessarily represent the entire acoustic experience. For one, the two selected sound samples did not have large variations in music dynamics, although it is one of the principal aspects in musical expression. Beethoven passage was steady *mezzo piano* of strings and woodwinds while Bruckner was a *forte* passage, dominated by brass instruments. Based on our previous research, we find it possible that there are several relations between musical expression and room acoustics. In order to understand the possible reasons behind such relations, we present the sound field typical to shoebox and non-shoebox halls in temporal, spectral, and spatial domains. Subsequently, we propose hypotheses that could explain the acoustical success of the Musikverein.

A few articles show measurement data or detailed analysis on the acoustics of the Musikverein. Most of them concentrate on analyzing objective parameters, and one presented in Acoustics'08[3] shows the main room acoustical parameters at octave bands both in concert and in ballet (chairs removed) configurations. Naturally, the values of measured room acoustical parameters can also be found in the book by Beranek[1], which presents objective parameters measured in both occupied and unoccupied conditions. However, the book does not present any further analysis of the acoustics of this famous hall. One more paper[9] reporting measured room acoustical parameters does not explain why the Musikverein would be in someway extraordinary or different from other halls of the same size. Some subjective analysis of the Musikverein is presented by Clements[2], accompanied by an interesting review of its renovations during the years.

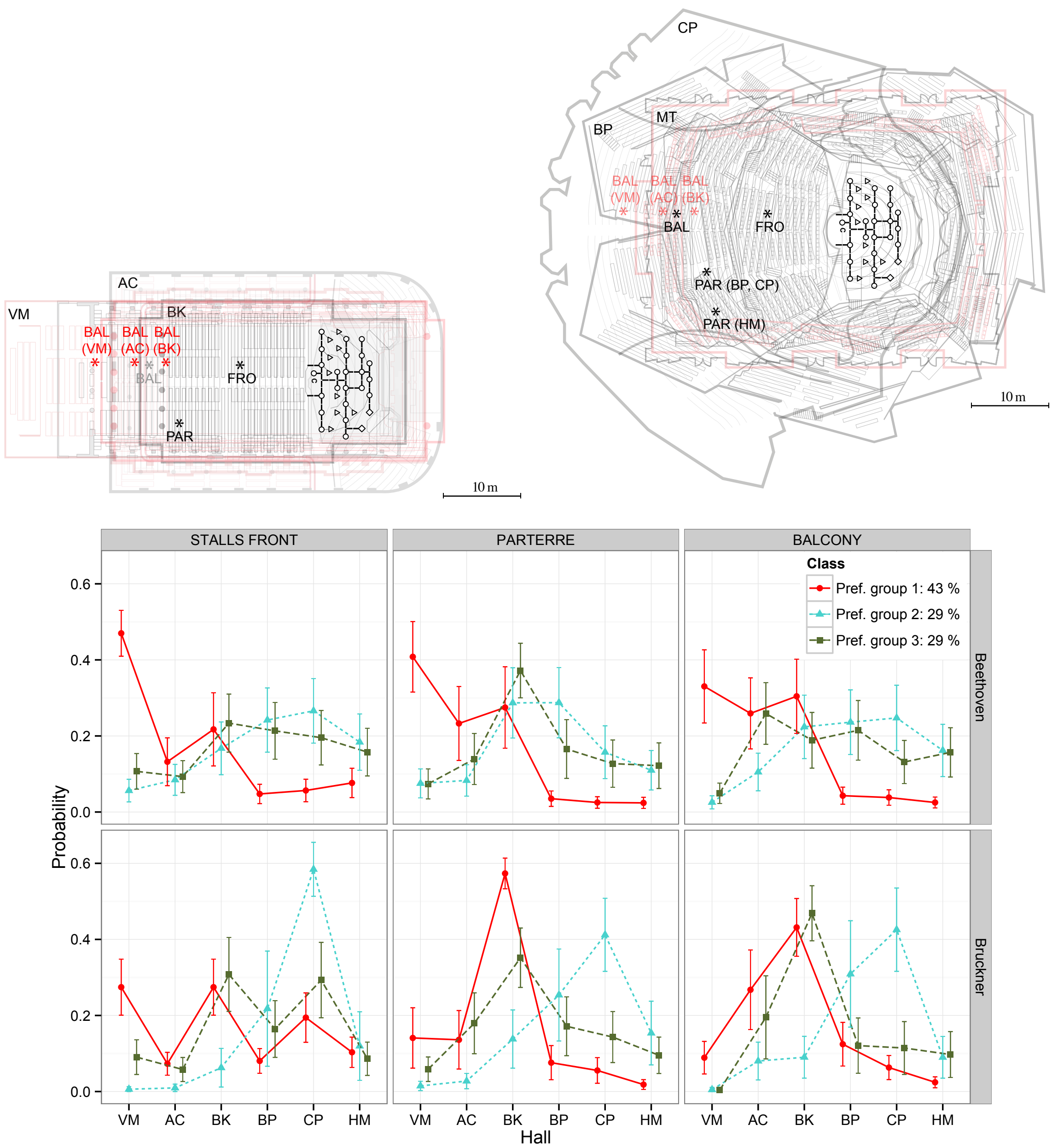

Figure 1 Preferences from the listening tests, in which 28 subjects compared six halls at three listening positions with two signals twice. The listening positions were at Front Stalls (FRO), Parterre (PAR), and the first row on the Balcony (BAL). The halls were Vienna Musikverein (VM), Amsterdam Concertgebouw (AC), Berlin Konzerthaus (BK), Berlin Philharmonie (BP), Cologne Philharmonie (CP), and Helsinki Music Centre (HM).

2 ANALYSIS OF MEASURED SPATIAL IMPULSE RESPONSES IN THE TIME-FREQUENCY-SPACE DOMAIN

In 2012 we made a European tour[18] measuring ten concert halls with the loudspeaker orchestra[12]. Here, we present the measured data in the time-frequency-space domain[13] of four concert halls on the seats that were used in the listening tests[8] (see Fig. 1). Data is averaged over 24 source positions. The purpose of the analysis is to describe typical acoustical conditions of shoebox and vineyard halls. The analysis shows clearly the differences across hall types are larger than in one hall across seats, except at balcony, where sound fields are more similar in all hall types.

Figures 2-4 illustrate the cumulative sound energy in the time-frequency and in the spatiotemporal domains. Two left columns are traditional shoebox halls (VM and BK) and two right columns are modern vineyard halls (BP and HM). The first row shows the average cumulative frequency response at 5, 30, 120, and 3000 ms after the initial direct sounds.

Direct sounds, i.e., the first 5 ms of the sound obviously comes from the frontal directions. In shoebox halls the stages are higher, and thus the receiver positions do not receive the stage floor reflection, in contrast to vineyards. Interestingly in shoebox halls the direct sounds are quite weak, except at balconies although most literature on room acoustics discusses the importance of strong direct sound. For example, in shoebox halls parterre (Fig. 3) the low frequencies within the first 5 ms are low in magnitude compared with high frequencies. When the floor reflections are present (as in vineyards with raked audience area) the frequency responses of the first 5 ms are quite different. The mid frequencies are quite strong, but low frequencies below 100 Hz are strongly attenuated similarly to shoeboxes.

First early reflections until 30 ms are visualized with dark blue color and it is known that perceptually they are integrated into the direct sound. Two main differences between shoebox and vineyard halls are the shape of the frequency response and the spatial distribution of these early reflections. The shoebox halls provide lateral reflections very early, see the triangular shape of the dark blue area in the visualizations. Note also that in vineyards, in particular on seats further from the stage, ceiling reflections are stronger and arrive earlier than side reflections although only a few side reflections are present, resulting in an oval-shaped energy distribution in the lateral plane. The other major difference lies in the frequency responses. In shoeboxes, the early reflections strengthen the low frequencies below 125 Hz substantially, yet the mid frequencies up to 1 kHz remain at quite a low level. Such a time-dependent filtering is caused by the seat-dip effect in halls with flat floor and open seats[17]. In contrast, when the sound cannot pass under the seats and floor is raked the low frequencies below 120 Hz remain attenuated and sound energy increases only slightly between 5 and 30 ms. Moreover, in mid frequencies in vineyard halls there is only a small amount of energy increase during the first early reflections. Our explanation for this phenomenon is that when the direct sound is accompanied with the reflection from the stage floor the mid frequencies are much stronger immediately, and relatively weaker early reflections cannot overpower the strong direct sound. In other words, the direct sound accompanied with the floor reflection masks the perception of early reflections. Similar ideas on masking were presented by Marshall already almost 50 years ago[10,11], but hardly any research has been done on this topic.

Later reflections between 30 and 120 ms (or even 200ms, not visualized here) increase the overall sound energy. In shoebox halls, the increase is particularly strong in mid frequencies equalizing the frequency response to be more or less flat approx. 100 ms after the direct sound. Moreover, the energy in this time window reaches the measurement position almost evenly from all directions and the sound fields in all three measured positions have a round shape in all visualization planes. In vineyard halls, the energy distribution is not as smooth, and bundles of energy from ceiling (or canopy) and side walls are seen on all measured seats. Moreover, the frequency responses keep the same shape with their peaks and dips as earlier although the cumulative energy increases the overall level.

Finally, the **reverberation** after 120 ms augments the cumulative energy to its final form. In the

Musikverein the increase is the largest resulting in the highest final level. The noticeable differences between shoeboxes and vineyards are in the smoothness of frequency responses, level of low frequencies, and spatial distribution of sound energy.

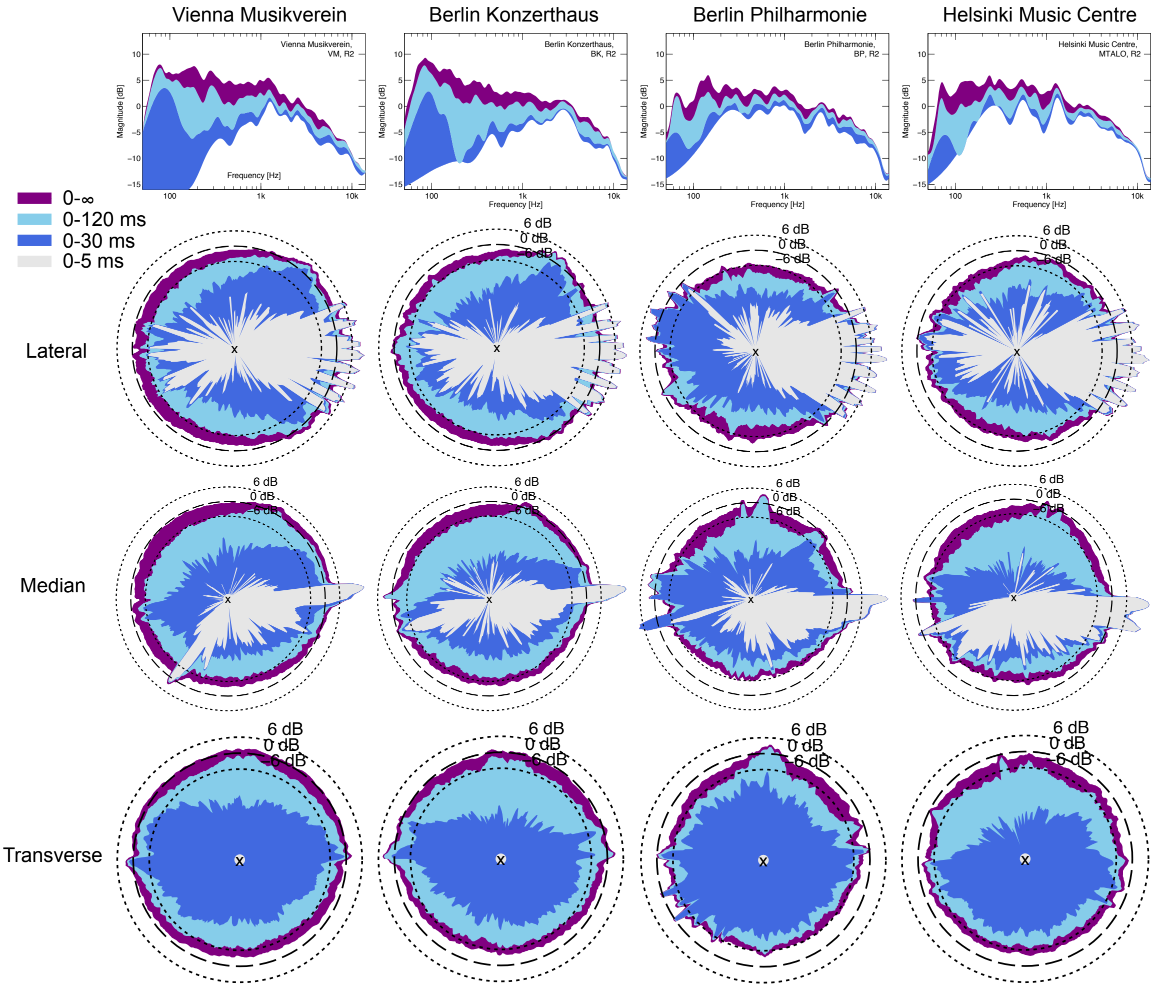

Figure 2 Spatiotemporal visualizations of cumulative sound energy at close to the stage (stalls front). The colors indicate the level of cumulative energy at 5, 30, 120, 3000 ms after the direct sound and the data is average of 24 source channels on stage. From top: frequency responses, sound energy in plane (2^{nd} row), in section (3^{rd} row), in transverse plane (4^{th} row). In spatiotemporal plots -6, 0, and 6 dB curves indicate the level re. level in free field at 10 m (i.e. the same level adjustment than with strength *G*).

3 HOW ACOUSTICS COULD SUPPORT THE ORCHESTRA AND MUSICAL INTERPRETATION AS PERCEIVED BY LISTENERS

In music research, individual musical elements, such as phrasing, articulation, dynamics, and microtiming, are often used in evaluating expressiveness of musical performances. The musicians use these aspects in their interpretation to convey the ideas and emotions that the composer has wrote to the form of art: music. Room acoustics literature does not discuss widely these musical elements. Maybe acousticians have considered that the acoustics hardly influences music interpretation, as the impulse response is linear and time-invariant, and thus cannot change dynamics or microtiming. Only the reverberation time is known to change interpretation; the longer the reverberation is, the slower the orchestra has to play.

The loudspeaker orchestra can reproduce music identically in each hall allowing rigorous perceptual studies. Our recent results suggest that even though a concert hall is a linear conveyor of sound, it can render sound in a non-linear manner, due to the non-linear excitation (the instruments of an orchestra) and the non-linear receiver (human spatial hearing)[7,15]. Therefore, we hope that research on concert hall acoustics would redirect the focus to explain the aspects in music a concert hall can influence, i.e., how a concert hall can affect music to be as expressive and impressive as possible. That would require that research should more often take in the account the combination of musical interpretation and room acoustics. Here, we reflect our recent research results in the light of the analysis of temporal development of spatial impulse responses presented above.

Expressive and touching music performance can evoke strong emotions for the listener. This requires that the listener can concentrate on the music and that the concert hall conveys the music from the stage to the listener by retaining or even enhancing the minute musical expressions. Next, we make some remarks on the musicians' expression in their interpretation, and how room acoustics could influence them. Note that individual expressions are often linked together although they are here discussed one by one.

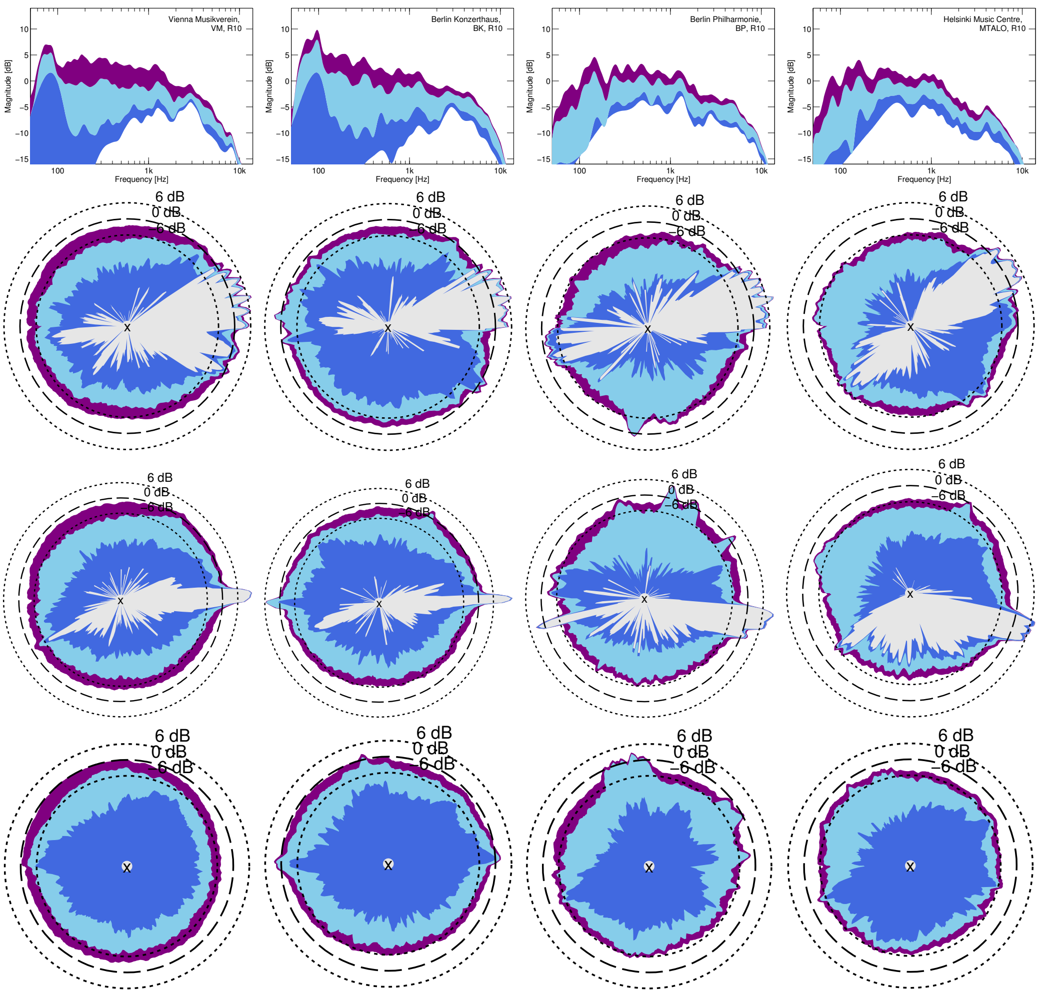

Figure 3 Spatiotemporal visualizations of cumulative sound energy at parterre.

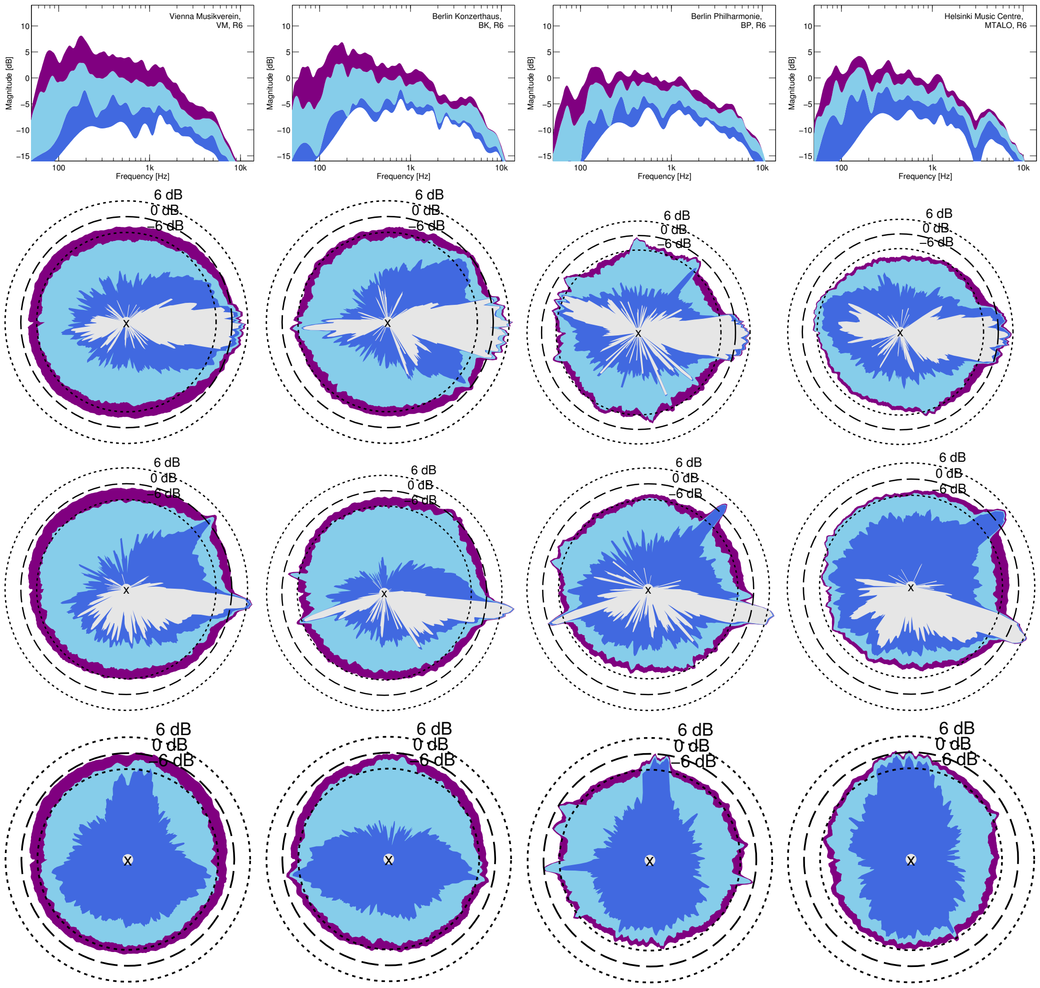

Figure 4   Spatiotemporal visualizations of cumulative sound energy at first row of the balcony.

Dynamics stand for the variations in music level, which can be instantaneous or gradual (e.g. *crescendo* or *diminuendo*). With dynamic variation musicians aim to form the motifs and render music more expressive. Traditionally, dynamics have been associated only with the skill of the orchestra and the musicians, and room acoustics has not been considered having any effect on dynamics. This is not true, as room acoustics can pertain to dynamics [15] and have variable impact on listeners [14]. In the Musikverein such impact is particularly strong, rendering the music exceptionally expressive. This is most probably due to the low level of the direct sound, which leaves more room for early lateral reflections and late reverberation to be more audible, in particular with stronger playing style, when higher harmonics are much more prominent than with soft playing [15].

Phrasing is a combination of dynamics and tempo changes to emphasize the structures of music. Phrasing can be seen as concepts and practices related to grouping consecutive melodic notes, and a musical work is often made up of a melody that consists of numerous consecutive phrases. Reverberation is crucial for phrasing and it supports the music and helps musicians to tie consecutive phrases together. When reverberation reaches the listener from lateral directions and from the upper hemisphere, spatial human hearing can separate reverberation to its own perceptual "stream" that Kahle calls *room presence* [4]. Consequently, such enveloping reverberation leaves room for the following notes without reducing the definition, articulation, and clarity.

Articulation is a common term to how individual notes are started, emphasized, and played. In vineyard halls, the articulation is often distinct due to strong direct sound accompanied with the stage floor reflection and moderate reverberation. If the reverberation is long as in the Helsinki Music Center (T60 = 2.5 s) the articulation is lost when an orchestra plays in *fortissimo*. This hall has primarily frontal reverberation, and even if the direct sound is relatively strong, the frontal reverberation reduces the clarity because new notes in music are not separated perceptually so well from the reverberation of previous notes. In shoebox halls articulation might be blurred, in particular on more distant seats due to the high amount of reverberation. However, in the Musikverein the articulation is surprisingly clear and translucent close to the stage until the halfway of the hall, even in an unoccupied condition. Maybe the spatial separation of the direct sound, early reflections and enveloping late reverberation preserves articulation lucid and reduces masking the onsets of notes. Definitively, the temporal development of the frequency response (Figs. 2 and 3) also plays a role here, but more rigorous studies are needed to understand masking in relation to the spatial and temporal development of the sound field.

Intimacy is discussed in another article[6], where it is viewed as a perceptual construct which connects us simultaneously to the musical performance and the surrounding space. Moreover, intimacy is viewed as a dynamic aspect, and thus is affected by the dynamic and spatial responsiveness of the hall. In the light of the presented visualizations, intimacy is affected by the manner how musical expressions are enhanced by the spatiotemporal and spectral evolution of the sound field. In our current opinion mid frequencies may not be so important, but that really low frequencies ($<$100 Hz) and high frequencies ($>$2 kHz) should be pronounced.

Spaciousness is also a function of played dynamics. The level dependency has already been found by Marshall[10] and Keet[5] although Keet used the same music sample with different levels, instead of the natural spectral change of an orchestra. When a hall has strong lateral early reflections, the widening of source in large *crescendo* could be really strong as in the Musikverein. In addition, in shoebox halls the space often "wakes up" in *fortissimo* enveloping the listener inside the music. A combined sensation of spacious, yet close sound, may also contribute to the perceived intimacy.

Microtiming is related to phrasing, i.e., in small temporal changes that musicians make within the nominal tempo. Another interesting timing issue is the coincident notes and how they are played by different instrument groups. One such playing style is that low frequency instruments play slightly before other instruments[*]. Recently, we organized a brief listening test to find out whether the hall could emphasize such microtiming[16]. The results showed that subjects preferred the chords in which bass instruments were played 40 ms earlier than high violins and woodwinds. The preference was even stronger in a shoebox hall (Amsterdam Concertgebouw), which has quite similar temporal development of the frequency response as in the Musikverein. A possible explanation is that because the mid-high frequency response is delayed, the higher partials of the bass-register instruments are less masked by the treble-register instruments, when the bass-register instruments start to play earlier.

4 CONCLUSIONS AND FUTURE DIRECTIONS

To summarize, great acoustics is often defined to encompass good articulation, definition, and warm enveloping reverberation. Such acoustics is achieved when the spatial and temporal masking do not hinder the listener to hear continuous music, while enveloping reverberation completes the sound - in other words the hearing can use all possible information available. The direct sound does not need to be very strong, as early reflections from the sides and up strengthen and enrich the direct sound. Avoiding the comb-filter effects of stage floor or low-hanging canopy reflections retain the direct sound spectrally unimpaired. The seat-dip effect could be beneficial, as it leaves headroom, both in spatiotemporal and time-frequency domains, for the early reflections to increase spaciousness. When reflections arrive from large lateral angles or from elevated lateral directions, they still increase spaciousness without reducing the definition and clarity of new direct sounds (with continuous music). Finally, spatially enveloping late reverberation blends the music naturally and augments the perceived

[*]An example: "'Die Kunst des Dirigierens"' with Herbert von Karajan and Vienna Philharmonic Orchestra between 1:00- 1:37 – `http://www.youtube.com/watch?v=Shc-4AZVaNk`

frequency response as it correlates weakly with the direct sounds of the consecutive notes. Based on the geometry and presented analyses, we find it likely that these effects are materialized in the acoustics of the Musikverein.

We also promote that concert hall acoustics research should focus more on the combination of the orchestra sound and room acoustics. The spectra of the musical instruments vary a lot according to the played dynamics, thus the orchestra excites a concert hall differently at different dynamics. In addition to the instruments' physical properties, the widening of the excited frequency spectrum is also evident in many orchestral works from a compositional point of view. Quite often in *piano* only a few instrument sections are in voice, but during *crescendos* the remaining sections join into the texture. At the same time, the pitch of the leading voice increases. Furthermore, the classic counterpoint dictates that the bass line moves opposite to a rising melody -- toward lower frequencies. In such cases, not only the nominal dynamics change, but the excited frequencies span a much wider band. In our opinion, one secret of the Musikverein is that it optimally emphasizes these low and high frequencies both in time and space, resulting in large variations in dynamics and spaciousness and thus rendering the music highly expressive.

Acknowledgments: The research leading to these results has received funding from the Academy of Finland, project no. [257099].

5 REFERENCES

1. L. Beranek. *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*. Springer-Verlag (New York), 2nd edition, 2004. 664 pages.
2. P. Clements. Reflections on an ideal: Tradition and change at the grosser musikvereinssaal, vienna. *Acoustics Bulletin*, 21:5--16, November/December 1999.
3. A. Farina, D. E. Commins, and N. Prodi. Experimental analysis of the acoustical behaviour of Musikverein in concert and ballet configurations. In *Proceedings of Acoustics '08*, page 6 pages, 2008. Abstact in J. Acoust. Soc. Am. , 123(5), 2973 (2008).
4. E. Kahle. Room acoustical quality of concert halls: perceptual factors and acoustic criteria - return from experience. In *International Symposium on Room Acoustics (ISRA 2013)*, International Symposium on Room Acoustics 2013, Toronto, Canada, Jun. 9-11 2013.
5. W. V. Keet. The influence of early lateral reflections on the spatial impression. In *Proc. 6th International Congress on Acoustics*, volume 3, pages E53--E56, Tokyo, Japan, 1968.
6. A. Kuusinen and T. Lokki. Auditory distance perception in concert halls and the origins of acoustic intimacy. In *IOA Auditorium Acoustics*, Paris, France, Oct. 29-31 2015.
7. T. Lokki and J. Pätynen. The acoustics of a concert hall as a linear problem. *Europhysics News*, 46(1):13--17, 2015.
8. T. Lokki, J. Pätynen, A. Kuusinen, and S. Tervo. Music, listening position, and individual taste of the listeners influence the qualitative attributes and preferences of concert hall acoustics. *Submitted manuscript*, June 2015.
9. K.-H. Lorentz-Kierakiewitz and M. Vercammen. Acoustical survey of 25 European concert halls. In *NAG/DAGA 2009 Int. Conf. on Acoustics*, pages 957--961, Rotterdam, the Netherlands, Mar. 2009.
10. A. H. Marshall. A note on the importance of room cross-section in concert halls. *Journal of Sound and Vibration*, 5(1):100 -- 112, 1967.
11. A. H. Marshall. Levels of reflection masking in concert halls. *Journal of Sound and Vibration*, 7(1):116--118, 1968.
12. J. Pätynen. *A Virtual Symphony Orchestra for Studies on Concert Hall Acoustics*. PhD thesis, Aalto University School of Science, 2011.
13. J. Pätynen, S. Tervo, and T. Lokki. Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses. *Journal of the Acoustical Society of America*, 133(2):842--857, February 2013.
14. J. Pätynen, S. Tervo, and T. Lokki. Subjective impact of concert hall acoustics. In *IOA Auditorium Acoustics*, Paris, France, Oct. 29-31 2015.
15. J. Pätynen, S. Tervo, P. W. Robinson, and T. Lokki. Concert halls with strong lateral reflections enhance musical dynamics. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 111(12):4409--4414, 2014.
16. H. Tahvanainen, A. Haapaniemi, J. Pätynen, and T. Lokki. Perceptual relevance of asynchrony between orchestral instrument groups in two concert halls. In *Third Vienna Talk on Music Acoustics*, Vienna, Austria, September 16-19 2015.
17. H. Tahvanainen, J. Pätynen, and T. Lokki. Analysis of the seat-dip effect in twelve european concert halls. *Acta Acustica united with Acustica*, 101(4):731--742, 2015.
18. S. Tervo, J. Pätynen, and T. Lokki. Spatio-temporal energy measurements in renowned concert halls with a loudspeaker orchestra. In *the 21st International Congress on Acoustics (ICA'2013)*, Montreal, Canada, June 2-7 2013. Invited paper.