Pätynen, Jukka; Lokki, Tapio

Perception of music dynamics in concert hall acoustics

Published in:
Journal of the Acoustical Society of America

DOI:
10.1121/1.4967157

Published: 01/11/2016

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Pätynen, J., & Lokki, T. (2016). Perception of music dynamics in concert hall acoustics. *Journal of the Acoustical Society of America*, *140*(5), 3787-3798. https://doi.org/10.1121/1.4967157

Perception of music dynamics in concert hall acoustics

Jukka Pätynen[a)] and Tapio Lokki
Department of Computer Science, Aalto University School of Science, FI-00076 Aalto, Finland

(Received 1 July 2015; revised 16 August 2016; accepted 19 October 2016; published online 17 November 2016)

Dynamics is one of the principal means of expressivity in Western classical music. Still, preceding research on room acoustics has mostly neglected the contribution of music dynamics to the acoustic perception. This study investigates how the different concert hall acoustics influence the perception of varying music dynamics. An anechoic orchestra signal, containing a step in music dynamics, was rendered in the measured acoustics of six concert halls at three seats in each. Spatial sound was reproduced through a loudspeaker array. By paired comparison, naive subjects selected the stimuli that they considered to change more during the music. Furthermore, the subjects described their foremost perceptual criteria for each selection. The most distinct perceptual factors differentiating the rendering of music dynamics between halls include the dynamic range, and varying width of sound and reverberance. The results confirm the hypothesis that the concert halls render the performed music dynamics differently, and with various perceptual aspects. The analysis against objective room acoustic parameters suggests that the perceived dynamic contrasts are pronounced by acoustics that provide stronger sound and more binaural incoherence by a lateral sound field. Concert halls that enhance the dynamics have been found earlier to elicit high subjective preference. © *2016 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4967157]

[MV] Pages: 3787–3798

I. INTRODUCTION

In live performances of symphonic music, the orchestra sound is inseparably combined with the acoustics of the concert hall. For many listeners, music represents aesthetic enjoyment and emotional expression.[1–3] In European-influenced classical music, the composer has a variety of key elements for conveying the musical expressions, such as pitch, note duration, timbre, and dynamics (Ref. 4, p. 6). Without the intended variations in the music dynamics and tone color, the expressivity is often diminished. This, in turn, would impede the listeners' experience and the possible emotional impact sought from the performance.

The room acoustics defines how the sound is conveyed from the instruments to the listeners, and the room impulse response is commonly considered a linear time-invariant system. However, the listener's ears and auditory system are highly dependent on the direction, spectrum, and level of the incident sound. The directions of the incident sound are closely associated with the room geometry, and depend also on the source directivity. In a music performance, the sound level and spectrum of the instruments depend on music dynamics, since the harmonic overtones become disproportionally stronger with the increased music dynamics. For instance, paths of sound propagation that are particularly sensitive to higher frequencies due to the directional hearing become perceptually more significant with the change of signal spectrum.[5] Together the source, path, and receiver form a non-linear system,[6] where the room acoustics may affect the perception of music dynamics in many aspects.

In this paper, we show with a listening experiment that the perceptual effects of varying orchestra dynamics depend on the acoustics of the concert hall, even if the contrasts in performed music dynamics are identical. Also, we investigate the perceptual attributes in which the sound changes as the interaction of music dynamics and concert hall acoustics. The general results suggest that a strong and lateral sound field enhances the perceptual dynamic effects, which provides evidence also for earlier studies on the musical dynamic range of concert halls.[5] The outcome of the experiment offers also an explanation for the subjective rank-ordering of acoustic quality,[7] as certain halls with a reputation of outstanding acoustics appear to provide strong perceptual responsiveness to varying music dynamics, thus, rendering a more expressive listening experience.

II. BACKGROUND

The research of room acoustics reaches back to the turn of the 20th century with discoveries on interaural differences by Lord Rayleigh, and room reverberation by Sabine.[8] Until then, new halls were designed after existing buildings, which in turn were constructed within the practical and technological limitations. Over the following decades, the statistical decay time of late reverberation became the dominant design criterion in concert halls.[9] Subsequently, independent research groups have identified by listening experiments[10] a common set of perceptual attributes describing the sound in concert halls: reverberance, loudness, spaciousness, clarity, and intimacy.[9,11–13] The need for predicting perceptual qualities from room impulse responses has resulted in a multitude of objective metrics[14,15] such as strength, interaural

[a)]Electronic mail: Jukka.Patynen@aalto.fi

 0001-4966/2016/140(5)/3787/12/$30.00

cross-correlation coefficient, clarity, and lateral energy measures.

The direction of sound reflections is one of the most significant single concepts in concert hall acoustics, and it was not recognized until the late 1960s in the seminal works by Marshall.[16,17] Importantly, the early lateral reflections were identified as the principal origin for the broadening of the sound image. For measuring the perceived source broadening due to early lateral reflections, Keet proposed an incoherence metric with a stereophonic microphone setup.[18] Later, the lateral energy fraction (LEF)[19] has been widely adopted for measuring the proportion of early energy reflected from the sides. As the lateral sound causes interaural differences, it reduces the interaural coherence[19,20] between the signals entering the ears.[21,22]

In the early studies, Keet also observed with music that an elevated sound level increases the impression of apparent source width in concert halls.[18] Therefore, early reflections arriving from the sides would effectively be more audible than those arriving from the back or front when the overall sound level is increased. This effect was confirmed with synthetic sound fields for early and late reflections (Ref. 23, p. 33), as well as in real concert halls[24] for early reflections. Together these findings imply that a desired spatial spreading of sound image can be achieved either by increasing the amount of early lateral energy,[18] or by increasing the overall sound level.[25] In essence, the perceptual effect by the sound level differences is not limited only to overall loudness, but it also influences the spatial impression and, potentially, to other perceptual aspects as well.

Although the aforementioned research has discovered many central concepts in room acoustics, previous studies share a notable shortcoming. That is, practically all listening tests have considered continuous music to be stationary with respect to dynamic levels. For instance, prominent studies[19,20,26,27] have used passages from Mozart's Symphony No. 41 as anechoic material. While the music dynamics vary during such a music signal, the experiments have not taken into account how the music dynamics might influence the perception of the room acoustics.

The potency of the combination of music dynamics and room acoustics becomes evident when we consider the nonlinear frequency spectrum in music dynamics instead of an artificially modified presentation sound level. The research on orchestra instruments[28,29] shows how the high-frequency overtones are emphasized disproportionally along increasing playing dynamics. Therefore, the music dynamics affects both the level and spectrum of the signal exciting the room, while the frequency-dependency of spatial hearing and spectral masking[30] gives significance to the reflection directions and the room geometry.[5,23,24]

Authors such as Meyer and Beranek, have pointed out the importance of the concert halls' response to music dynamics. They have stated that, first, while quiet dynamics can be acceptable in poor halls, the high quality of forte is an indication of an acoustically good hall (Ref. 29, p. 199), and second, that "the dynamic response of the concert hall enhances music listening immeasurably" (Ref. 22, p. 509). In addition, Kahle suggested that the sound level-dependent perception of lateral reflections "make the concert hall to wake up" with increasing music dynamics (Ref. 10, p. 29). Pätynen *et al.* showed an objective study where rooms with lateral early reflections enhance the transmission of the dynamic range in music to the listeners' ears.[5] These observations propose that the concert halls may transmit music dynamics differently, and also that the rendering of music dynamics could be a substantial factor in the overall room acoustic quality. Still, the subjective perception of the dynamic response has remained unexplored. For clarity, dynamic response stands for the perceptual effects of changing sound image along varied music dynamics in room acoustics.

The following experiment focuses on resolving whether concert hall acoustics vary in their dynamic responses, and in which perceptual aspects the dynamic responses in different concert halls are manifested.

III. METHODS

The dynamic response and its perceptual factors were investigated with a listening test. The presented orchestra music signal contained a prominent contrast in music dynamics, and the identical signal was rendered with various measured concert hall acoustics. The spatial sound was presented to the subjects with loudspeaker reproduction (see Fig. 1). Details regarding the listening test setup are described in Secs. III A–III C.

The listening test methodology was paired comparison with a simultaneous free attribute elicitation. The subjects' task was to compare two stimuli at a time, and choose the one that appeared to change more prominently on the whole during the music. They could also indicate a tied comparison if both stimuli appeared to change identically, or they could not decide on one stimuli having a more prominent change. The subjects were instructed to concentrate on the general observations about the sound rather than focusing on smaller details, such as individual notes or instruments.

At this point, it is essential to note that the perceived difference between stimuli, i.e., dynamic response, was not necessarily the apparent dynamic range of music. Instead, the underlying dynamic step could have produced other perceptual contrasts between the stimuli. For example, the varied music dynamics may have yielded a pronounced impression of an extending width of sound in certain room acoustics. In order to explore the perceptual factors in which the compared acoustics manifest the changes in music dynamics, the subjective criteria for each selection was collected during the test. After each pair, the subjects wrote on a paper form a short description of the most prominent difference between the perceived changes during the stimuli. In order to avoid biasing the subjects' judgments, the authors refrained from providing details about the stimuli, or direct cues for possible perceptual differences in the test instructions. Despite the unusual experiment setting, the subjects understood the given tasks immediately.

After the experiment, the answer sheet was talked through with each subject. The purpose of the discussion was to resolve possible unclear descriptions or definitions and to ensure the correct interpretation. In particular, the

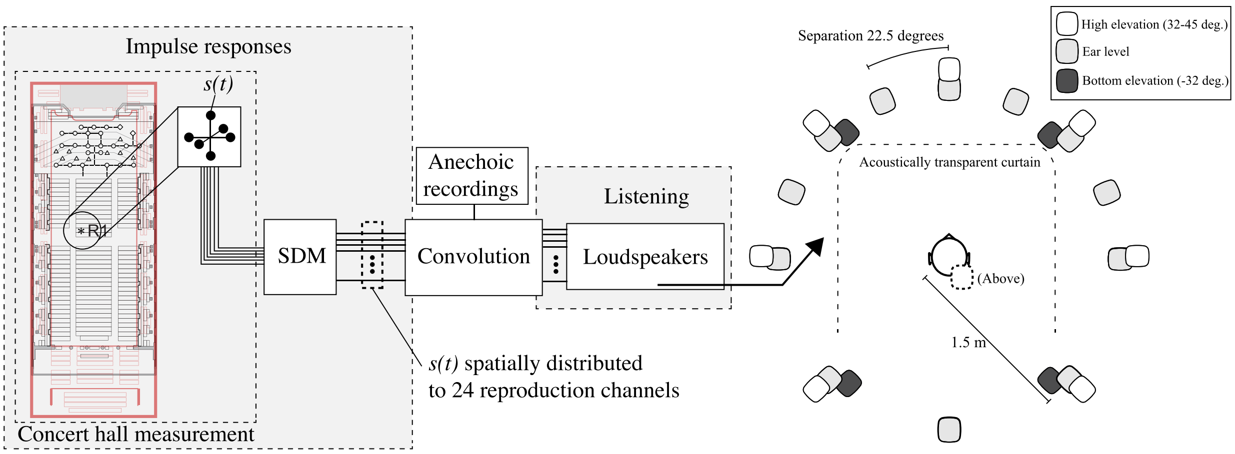

FIG. 1. (Color online) Block diagram of the room-acoustic measurement and auralization process for one measurement source channel. The left side of the figure shows the positions of the reproduction loudspeakers in the listening room.

subjects confirmed the polarities of certain terms prone to ambiguity, such as distance. Then, the subjects condensed the descriptions into concise perceptual attributes for the subsequent analysis of the paired comparison results. The subjects compared 6 concert halls (15 pairs) in three respective receiver positions each. Hence, one subject produced 45 comparisons accompanied by the respective attributes.

Paired comparisons were completed using a touch screen with buttons for switching seamlessly between concert halls and their respective answer selections. The subjects could listen to the stimuli as long as necessary to make their judgments, but modifying the playback loop was not allowed. The screen was positioned in front of the subject at a height suitable for keeping the subjects looking more forward than down, as in a concert situation. A neutral field of view was provided by an acoustically transparent curtain obscuring all equipment apart from the user interface.

Twenty-eight subjects participated in the test (14 male, ages 22–64 yr, average 40 yr). Their musical backgrounds were heterogeneous, ranging from ordinary music consumers to music professionals. Especially, we included subjects from the age group that typically attends classical concerts, that is, over 50 yr. Audiometry reported normal hearing for all subjects considering their age and occupation. All had a minimum of five hours of experience in critical listening and describing perceptual differences, as they had participated, on different days before the present experiment, in four sessions of listening tests with individual vocabulary profiling.[31,32] The experimenters did not disclose that the underlying orchestra signal was identical in each stimulus, or that the stimuli represented different concert halls. The music signal applied here was used uniquely for the present experiment.

A. Music signal

Human hearing adapts quickly to gradual changes in sound or acoustics. At the same time, recollecting a preceding acoustic condition accurately is difficult.[33] Since our hypothesis is that dynamics in music alter the perceived acoustic impression, the presented signals should not allow the listeners time to adapt during slow gradual dynamic changes.

Although many orchestral works contain sharp increases in music dynamics, we created a signal with sudden, but musically feasible, dynamic steps from anechoic orchestra recordings.[34] Anton Bruckner's Symphony No. 8, II movement, includes a long crescendo with the full orchestra at the end of the first theme. Hence, a signal with a strong contrast in playing dynamics was constructed from segments before and after the crescendo, omitting the gradual increase in between. An uninterrupted combination of bars 41–43 and 53–55 is suited for studying the effect of orchestra dynamics while other musical factors including the tonality, instrumentation, texture, and playing registers remain nearly constant (see Fig. 2). The instrumentation during the select excerpt includes three flutes, oboes, clarinets, and bassoons, eight French horns, three trumpets and trombones, a tuba, timpani, and strings. The characteristic sound of a string section was simulated to the anechoic recordings with a specific method.[35] It recreates the natural differences of individual players with variations in pitch and timbre, as well as time-variant delays and amplitude. The resulting complement of 16 first and second violins, 10 violas, 8 cellos, and 6 double basses corresponds to a typical orchestra for this composition.

A short-time frequency analysis [see Fig. 3(a)] shows the spectral content of the music signal. The spectral contrast during the constructed passage is shown in Fig. 3(b). The low-frequency increase results mainly from the entrance of timpani and tuba, as they play the fundamentals of 77 and 154 Hz (E♭). The sound level change at the middle frequencies is more subtle in comparison to the high-frequency region, which includes only the harmonic overtones.[5]

For the listening experiment, the segmenting of the signal was performed after convolving the uncut passage. Hence, this approach includes also cutting the reverberation, i.e., skipping forward ten bars when listening to a performance. This approach is analogous to the switching back and forth between two room-acoustic conditions, as enabled in the listening test. The reverberation tail was not included in the end of the reproduced stimuli.

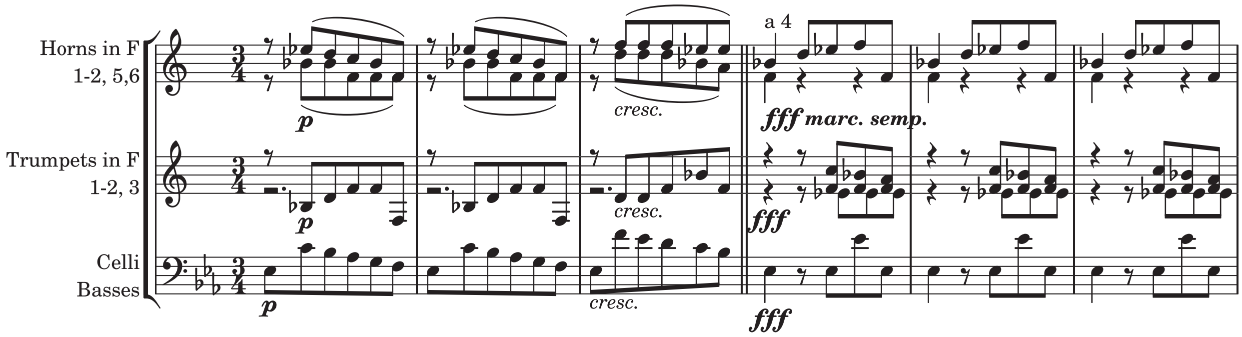

FIG. 2. Reduced orchestra score of the combined two passages in the music stimuli.

B. Concert hall measurements and auralization

The six European concert halls included in the experiment are listed in Table I with their abbreviations and physical and acoustic properties. Three of the halls are of the classical shoebox-type [Fig. 4(a)] with parallel side walls, flat floor, elevated stage, and relatively high ceiling. The other three represent more modern surrounding designs [Fig. 4(b)] with terraced seating blocks, or semi-circular amphitheatre style. The hall measurements were conducted in unoccupied halls. The absence of the audience absorption is assumed to influence mostly the reverberant sound, and less the direct sound and early reflections. These matters are considered further in Sec. V.

Three receiver positions were selected from each hall. The position at the front stalls (R1) was at a distance of 11 m from the orchestra, 2 m off-center to the left. The second position further back in the parterre (R2) was 19 m along the midline from the orchestra and 6 m off-center to the right [8 m in Helsinki Music Centre (HM) due to a sound engineer booth]. The receiver distance at the first row of the balcony (R3) depended on the hall geometry, and varied between 21.8 [Berlin Konzerthaus (BK)] and 29.2 m [Vienna Musikverein (VM)]. In the halls without balconies [Berlin Philharmonie (BP) and Cologne Philharmonie (CP)], a position at the rear parterre was chosen at a corresponding distance (see Fig. 4).

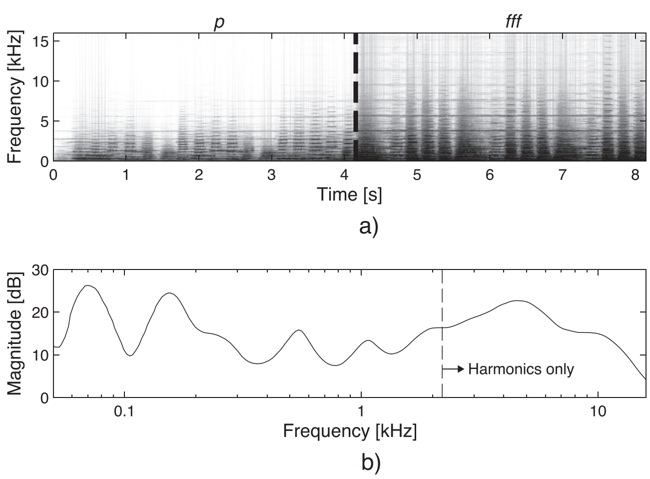

FIG. 3. Spectrum analysis of the music signal. (a) Frequency content of anechoic source signal segments. (b) Octave-smoothed difference curve between the average spectra of the first (in piano) and last three bars (in fortissimo).

The measurement source was a loudspeaker orchestra,[36] which consisted of 33 loudspeakers connected in 24 independent channels.[37] The dimensions of the loudspeaker layout, shown in Fig. 4, follow a typical orchestra arranged by the American seating. Nine loudspeaker channels representing the string instruments shared the signal with nine auxiliary loudspeakers on the floor, facing upward, for a combined directivity pattern that resembles more closely that of the actual instruments.[36] Equal source calibration was confirmed by using a 200–1000 Hz band limited noise measured at 1 m distance on-axis. The receiver was a G.R.A.S. type 50-VI 3D vector intensity probe (Holte, Denmark), which consists of three co-centric pairs of omnidirectional microphones arranged about the x-, y-, and z-axes. The distance between opposite microphones was 10 cm. The spatial room impulse response was separately measured from each source channel at a 48 kHz sample rate. We analyzed the responses using the spatial decomposition method (SDM).[38] In essence, SDM estimates the direction-of-arrival for each discrete sample in the impulse response by analyzing time-difference of arrivals between the six microphone capsules in short time-windows. The topmost capsule in the array also represents the omnidirectional pressure signal. In the spatial sound synthesis, the instantaneous pressure in the omnidirectional impulse response is assigned to the nearest reproduction loudspeakers[39] according to the direction estimates from the SDM. The result is a spatial convolution reverb from one measurement channel to 24 reproduction loudspeakers in the listening room (see block diagram in Fig. 1). Convolution with respective anechoic recordings produces an impression of few instruments being played on stage, and the combination of the same processing through all measurement sources results in the multi-channel output for the entire orchestra sound.

The same monaural room impulse responses and spatial information from SDM analysis were used for estimating various objective room-acoustic parameters. Instead of using a separate figure-of-eight microphone, we calculated lateral early (LEF) and late [lateral energy (LJ)] energy parameters from the room impulse responses from the probe with the octave-band filtered omnidirectional pressure and SDM direction estimation. The zero direction for figure-of-eight weighting was aligned toward the center of the stage.

TABLE I. List of European concert halls included in the listening experiment. V, N, G, and EDT denote concert hall volume [m^3], number of seats, average strength, and average early decay time, respectively. Measured values for G and EDT are averages from 500 and 1000 Hz octave bands over 24 source channels and all receiver positions. (†, estimated.)

Identification	Hall	Shape	V (m^3)	N	G (dB)	EDT (s)
VM	Vienna Musikverein	Rectangular	15000	1680	4.1	3.1
AC	Amsterdam Concertgebouw	Rectangular	18780	2040	2.8	2.4
BK	Berlin Konzerthaus	Rectangular	15000	1575	2.7	2.1
BP	Berlin Philharmonie	Vineyard	21000	2220	2.1	1.9
HM	Helsinki Music Centre	Vineyard	24000	1700	1.4	2.0
CP	Cologne Philharmonie	Fan	19000†	2000	1.9	1.6

Anechoic reference measurements have shown the SDM-based figure-of-eight to produce the desired directivity pattern at a wide range of octave bands. Additionally, we measured binaural room impulse responses with a Bruel and Kjær H.A.T.S. dummy head (Nærum, Denmark) at five equally spaced positions down an off-center line beginning 7 m from the stage. The binaural measurement position second closest to the stage coincides with position R1. Binaural measurements corresponding to the distance of R2 are behind position R1 instead of off-center to the right. Interaural cross-correlation-based measure was estimated from single positions nearest to the respective receiver positions. Binaural dynamic responsiveness (BDR) is a relative room-acoustic measure[5] that is obtained from the difference in the auditory excitations by full-orchestra spectra of opposing dynamics. The applied values were calculated as averages over three nearest distances (see example in Ref. 5). Binaural measurements are absent from the balconies and Helsinki Music Centre.

C. Listening room

The listening tests were conducted in a rectangular semi-anechoic room. The walls and the ceiling are treated with at least 5-cm thick sound-absorbing materials and a varying air space behind the absorption. The rigid floor is covered with a carpet around the listening position. The average reverberation time of the room is 0.11 s at the mid-frequencies, and the average peak-to-peak level difference between the direct sound and the strongest reflection at the 1–8 kHz frequency band is 12.8 dB. These values comply with the peak-to-peak difference of at least 10 dB recommended for subjective multichannel audio evaluation (ITU-R BS.1116-1). The short reverberation in the room is considered not to impair the listening accuracy or the relative differences between stimuli. The A-weighted background noise level from a moderate ventilation and the internal noise from idle loudspeakers is approximately 24 dB.

The spatial sound reproduction system comprised of 24 loudspeakers surrounding the listening position in three dimensions (3-D). Figure 1 illustrates the loudspeaker positions. Most loudspeakers (16) were in the frontal hemisphere since the majority of the sound energy in concert halls arrives from that region. The spatial resolution of human hearing is also the most accurate in frontal directions. Nominal distance from the listening position to the loudspeakers was 1.5 m, and variations up to 0.2 m in the actual

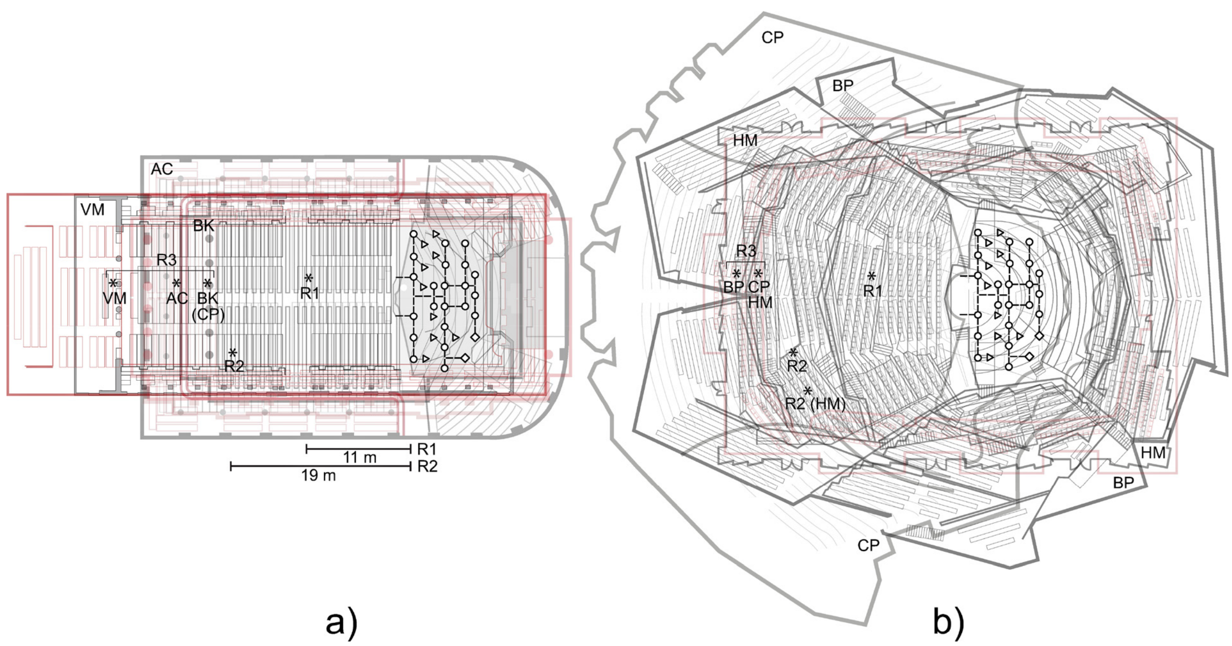

FIG. 4. (Color online) Overlaid floor plans of the (a) rectangular and (b) non-rectangular concert halls included in the listening experiment. Hall-dependent receiver positions R3 on the balcony front row are denoted with the hall abbreviations in parentheses (see Table I). Parts drawn in different shade indicate balconies above the main audience area.

distances were compensated for by delaying respective loudspeaker signals. The calibrated A-weighted sound levels from individual loudspeakers to the listening position were within ±0.5 dB.

The fidelity of the chain of spatial analysis and reproduction was assessed by a comparison of the room-acoustic parameters before and after the reproduction. For this purpose, the monaural and spatial objective parameters were estimated, first, from the impulse responses measured in the concert halls and, second, from the respective impulse responses reproduced in the listening room. The correlation coefficients between the measured and reproduced parameters over the included halls and receiver positions were high $r > 0.86$ (parameters averaged over 125 Hz–8 kHz octave bands) and $r > 0.96$ (250 Hz–2 kHz).

D. Analysis

Paired comparisons between six halls by a single subject yielded choice matrices of size 6×6 for each receiver position. An element of the initial zero matrix is increased by one when a hall represented by the row index is chosen over a hall represented by the column index.[40] The choice matrices were aggregated across individual subjects for the overall main results. We formed separate aggregate matrices for analyzing the effects of the hall typology and the receiver distance by combining the respective matrix rows and columns. To analyze the perceptual factors of dynamic response, the authors assigned the obtained descriptions manually into groups of similar attributes. The individual paired comparison answers were then distributed to more elaborate choice matrices based on respective attributes.

Several approaches exist for analyzing a choice matrix. We calculated the probabilities of choosing a certain concert hall over other equal alternatives with the Bradley-Terry-Luce (BTL) model.[41,42] It estimates the scale values that underlie the observed choice frequencies. The analysis model by Courcoux and Semenou suggests against segmenting the subjects into groups based on their answers, and provides the statistical significance of the differences between halls.[40] This approach also enables testing of hypotheses about perceived magnitudes in the framework of standard statistical theory. Our analysis is based on incomplete design[40] since the sets of complete paired comparisons are partitioned into aggregate matrices according to the associated perceptual attributes. A comprehensive overview on comparison models and their application is discussed by Choisel and Wickelmaier.[43]

The relations between the perceptual effects and objective room acoustic parameters were investigated with a correlation analysis using the Pearson's correlation coefficient. The large number of statistical tests between listening test data and objective parameters increases the possibility of false positive discoveries. The Bonferroni-type corrections to p-values can be conservative, and thus prone to false negatives (type II error). Furthermore, the objective room-acoustic parameters are not entirely independent, particularly over octave bands. For these reasons, we applied the Benjamini-Hochberg step-up procedure[44] to the correlation analysis for controlling the false discovery rate.

IV. RESULTS

The classification of choice criteria from 1260 individual comparisons resulted in a total of 57 unique attributes, and they were combined into seven attribute groups. 277 comparisons (22.0%) were tied. That is, the subjects could not indicate which stimuli changed more prominently, or they perceived that the stimuli changed identically. The perceptual criteria for 25 choices could not be described. Ten most frequently appearing attributes accounted for 75.4% of all non-tied comparisons, i.e., pairs where one hall was indicated to have a more prominent dynamic response.

The overall results indicate that the variation in music dynamics yields perceptual effects of different magnitudes depending on the concert hall and listening position. With the tied comparisons omitted, VM and BK were found to provide the most pronounced dynamic response, as seen in Fig. 5(a). In position R1, the difference between halls is more subtle, as four halls nearly reach on or over the indifference threshold. Further back (R2) and on the balcony (R3), the overall dynamic response is relatively enhanced in

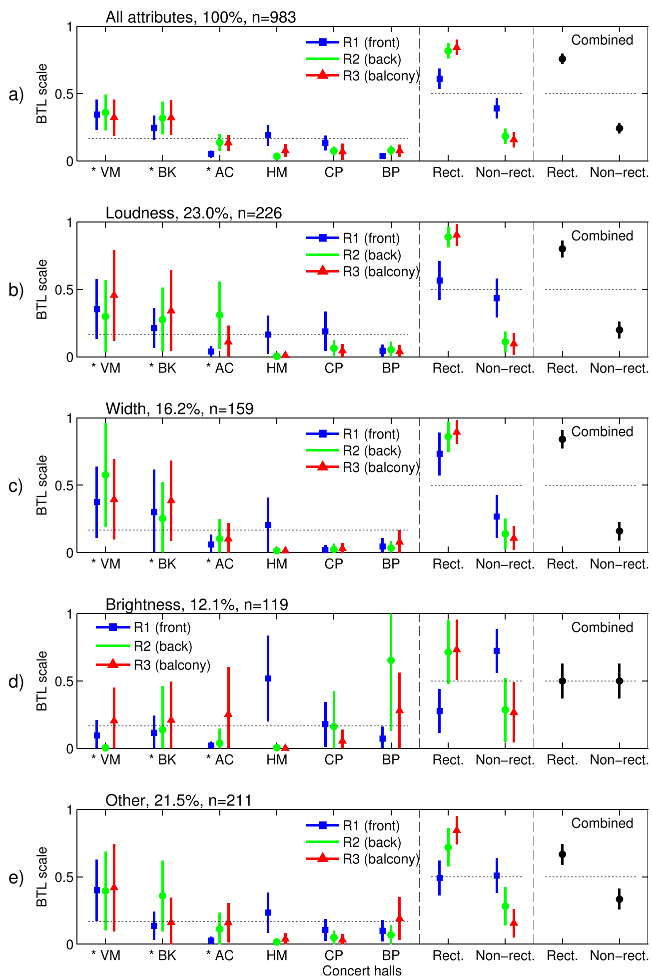

FIG. 5. (Color online) Results from the paired comparisons using the BTL analysis by all non-tied comparison and selected attribute groups. Vertical axis indicates the probability of choosing a hall over others for eliciting a larger perceptual contrast by the underlying music dynamics. The whiskers represent the 95% confidence intervals around the mean values. Rectangular halls are denoted with an asterisk. Horizontal dashed lines indicate the threshold of fully random answers.

Amsterdam Concertgebouw (AC) and remains high in VM and BK. Instead in HM and CP, perceptual effects from music dynamics are reduced. The effect of the receiver position to the overall dynamic response is pronounced in the aggregate results by hall type in Fig. 5(a). Near the orchestra, the difference between hall types is noticeable, and becomes evident at the middle stalls and balcony. The overall difference with the three positions combined together is positive for the shoebox-type in the included set of concert halls. Here it should be noted that the hall typology comparisons include only the applicable paired comparisons, thus, excluding pairs where two halls of the same type are compared together.

Aggregation of the paired comparisons by the attribute groups provides a detailed analysis on the different perceptual aspects of the dynamic response. The subjects' descriptions could be arranged into seven specific categories. Groups labeled as loudness, reverberance, bass, and brightness contain only descriptions that refer directly to the respective terms. Width group includes all attributes referring to spatial aspects, such as envelopment, spaciousness, width, and size. Clarity group includes individual attributes associated with articulation. All remaining and undetermined attributes were collected to a single group.

The most common attributes (23.0%) described a straightforward difference in the loudness variations, that is, the perceived dynamic range of music [see Fig. 5(b)]. Only 3 subjects out of 28 did not report dynamic range difference as the selection criteria for any compared pair. The relations between halls have a pattern similar to the overall results. AC gains relatively more apparent dynamic range in position R2. Analysis by hall type suggests that the contrast between rectangular and non-rectangular halls is marginal in the front position. However, at further receiver distances, the difference in the perceived dynamic range becomes substantial.

The rest of the attribute groups include smaller subsets of comparisons, which yield increasingly wider confidence intervals. Due to the less conclusive differences between individual halls, we present these results as provisional. The second largest (16.2%) attribute group for dynamic width and spaciousness [see Fig. 5(c)] follows the trend of the loudness attributes. Position R1 in HM has a relatively high degree of dynamic spatial responsiveness in comparison to other positions or other non-shoebox halls. A corresponding effect is observed with the dynamic-dependent spectral brightness [see Fig. 5(d)]. The dynamic brightness difference between hall typologies favors non-shoebox halls in position R1, while the order is reversed for R2 and R3.

The smallest attribute groups are compared only by hall typologies in Fig. 6. Results for the reverberance attribute suggest that the perception of hall reverberance grows with the music dynamics more in the included shoebox halls regardless of the receiver position. The apparent increase of bassiness shows a behavior similar to the width attribute group, while the clarity group results are inconclusive.

The remaining group consists of 40 mixed or unexplainable attributes. They account for 21.5% of all non-tied comparisons, and some attributes appear only a few times in total. This multifaceted group includes terms such as "openness," "fullness," and "richness." Seventy-four comparisons refer to dynamics-dependent perception of distance and proximity. These results follow the pattern of the loudness attribute group, and the differences between hall typologies grow wider at increasing receiver distances [see Fig. 5(e)].

In general, the analysis indicates that the perceptual differences elicited by varied music dynamics are smaller in the front positions between hall typologies. The differences between receiver positions suggests that the physical distance in the hall influences strongly the dynamic range and dynamic brightness.

Comparison matrices in Fig. 7 visualize the proportions and distributions of tied comparisons and the loudness attributes between halls and receiver positions. The highest percentages of tied comparisons occurred between two halls of the same type. For instance, the perceptual differences caused by the dynamics were difficult to distinguish between non-rectangular halls in R1. In R2 and R3 the contrasts between pairs AC-BK, and AC-VM were challenging to discriminate. Conversely, the lowest values in Figs. 7(a)–7(c) also indicate the pairs where differences in the dynamic response were observed the most often with any attribute. The comparison matrices in Figs. 7(d)–7(f) reveal that the difference in the perceived dynamic ranges was a substantial deciding factor especially in R2 and R3 between HM or CP and rectangular halls.

A. Correlation between perceived dynamic response and objective room acoustic parameters

Many perceptual aspects in room acoustics are estimated or quantified by the standardized objective parameters. To investigate the possible connections between the objective parameters and the perceptual effects by music dynamics, we calculated the correlation coefficients between the mean BTL probabilities in the three receiver positions and the respective objective parameters in octave-bands

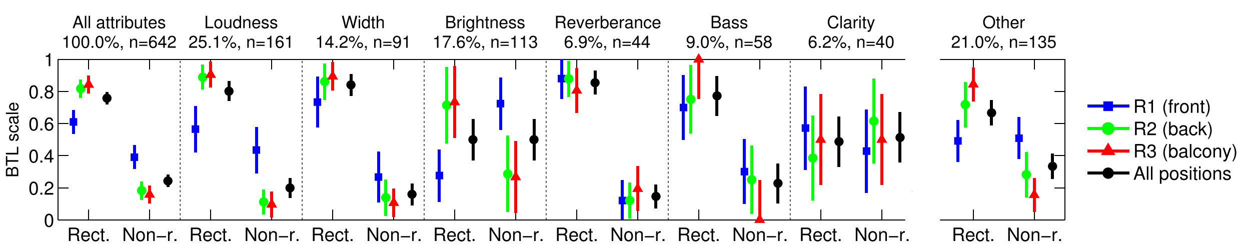

FIG. 6. (Color online) Results by attribute groups and hall typologies.

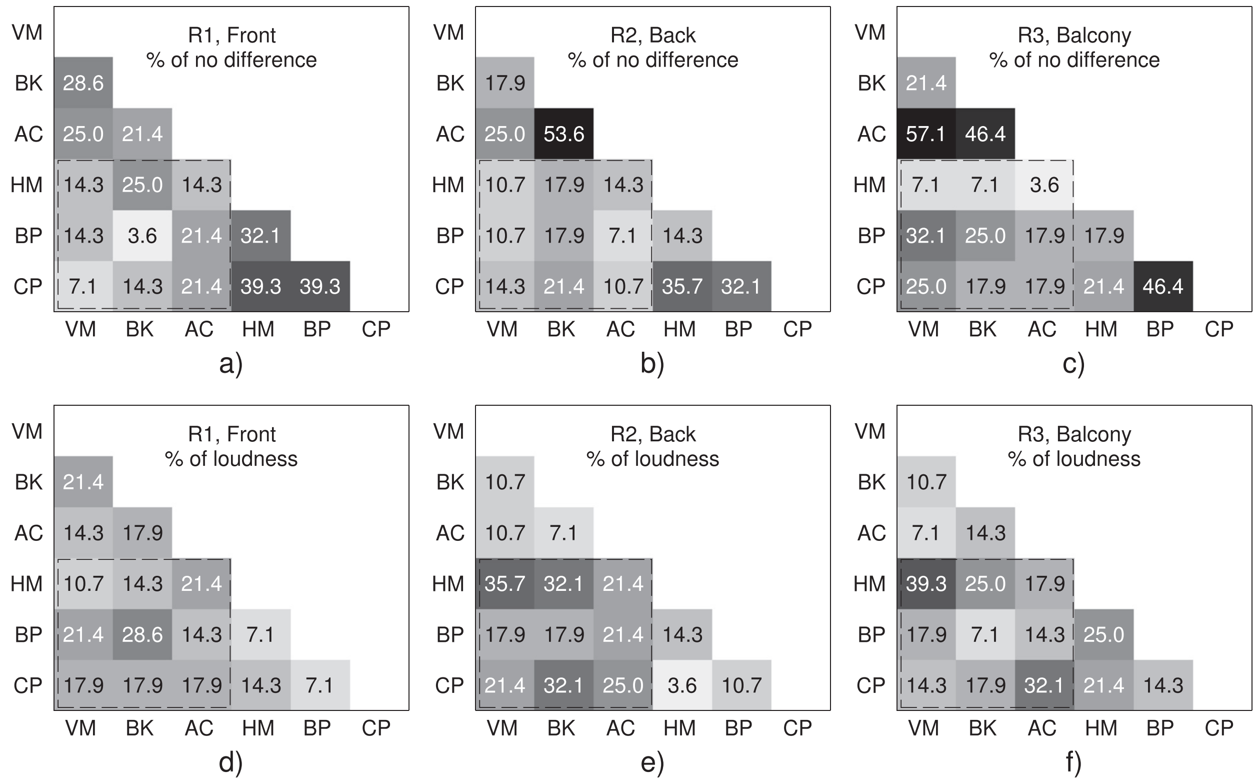

FIG. 7. Comparison matrix of the tied comparisons between concert halls in positions R1-R3 (a)–(c) and loudness attributes (d)–(f). The numbers inside the matrices indicate the percentage of all comparisons associated with the respective halls and attributes. The concert hall pairs inside the dashed line rectangle in the lower left corners are comparisons between two halls of different types.

averaged over all measurement source channels. Binaural parameters at low frequencies are omitted due to random variations.

The overall results for all non-tied comparisons suggest that the most prominent objective parameters contributing to the perceived contrast in music dynamics are the strength of sound (G) and late lateral energy (LJ) (see Table II). Binaural quality index (BQI) is the inverse of interaural cross-correlation coefficient in the early sound, and at high frequencies it correlates significantly with the dynamic effects near the orchestra. However, in R2 and R3 the correlations remain high but do not reach statistical significance.

The correlation results for the comparisons by attributes referring to the perceived dynamic range are shown in Table III. In position R1, none of the parameters correlate significantly with the perceived dynamic range, although G and BQI have the highest correlations at several frequency bands. In R2 and R3, several parameters (G, LJ, BQI) correlate with the dynamic range at a significant level. BDR, which is a relative parameter describing the dynamic range, correlates strongly at high frequencies ($r = 0.98$) in R2.

The highest correlation coefficients are encoded without adjusted significance in Table IV for all attribute groups. The dynamic width follows mostly BQI, EDT, G, and LJ parameters. Brightness attribute group correlates with LEF and LJ only in the balcony position. Dynamic-dependent reverberance is enhanced by longer EDT as well as higher LJ, and the reverberation time T correlates only in the balcony position. The group for other attributes follows loosely the same correlating parameters as the loudness attributes.

TABLE II. Correlation coefficients between overall mean BTL probabilities in paired comparison and selected room acoustic parameters as averages of two adjacent octave bands. Single (*) and double (**) asterisks indicate correlations at the ($p < 0.05$) and ($p < 0.01$) levels, respectively. Correlations which are significant with the Benjamini-Hochberg correction on an $\alpha = 0.2$ false discovery rate are emphasized in bold. In R2 and R3, none of the correlations are significant with the Benjamini-Hochberg correction.

Hz	G	C80	EDT	LJ	T	LEF	BQI	BDR
Position R1, 6 halls								
63–125	0.65	0.30	0.16	0.65	****0.97**	0.47	—	—
125–250	*0.88	0.03	0.66	0.80	*0.81	0.74	0.62	—
250–500	*0.87	−0.36	0.63	0.67	0.66	0.66	0.67	—
500–1k	0.72	−0.48	0.51	0.60	0.53	0.59	0.75	−0.29
1k–2k	*0.83	−0.44	0.49	0.62	0.51	0.54	*0.91	0.33
2k–4k	****0.93**	−0.28	0.38	0.52	0.37	0.43	****0.98**	0.37
4k–8k	0.04	−−0.21	0.24	0.34	0.13	0.49	***0.95**	0.35
Position R2, 6 halls								
63–125	0.70	0.49	−0.18	0.80	0.32	−0.34	—	—
125–250	*0.87	−0.59	0.71	*0.90	0.72	−0.01	0.68	—
250–500	*0.88	*−0.83	0.76	*0.85	0.63	0.41	0.73	—
500–1k	0.75	*−0.82	0.74	0.79	0.57	0.42	0.83	−−0.01
1k–2k	0.78	*−0.86	0.76	0.80	0.61	0.40	*0.90	0.73
2k–4k	0.75	−0.61	0.66	0.73	0.52	0.36	0.83	0.73
4k–8k	0.46	−0.36	0.45	0.51	0.28	0.39	0.73	0.73
Position R3, 6 halls								
63–125	0.57	−0.73	0.78	*0.87	0.30	−0.25	—	—
125–250	*0.81	−0.80	*0.86	*0.90	0.74	0.15	—	—
250–500	*0.86	−0.80	0.74	*0.90	0.64	0.40	—	—
500–1k	*0.85	−0.79	0.69	*0.85	0.56	0.44	—	—
1k–2k	*0.88	−0.78	0.73	0.81	0.60	0.38	—	—
2k–4k	*0.82	−0.76	0.73	0.77	0.54	0.47	—	—
4k–8k	0.22	−0.71	0.70	0.61	0.44	0.52	—	—

TABLE III. Correlation coefficients between mean BTL probabilities for paired comparisons by loudness attributes, i.e., perceived dynamic range, and room acoustic parameters similar to Table IV.

Hz	G	C80	EDT	LJ	T	LEF	BQI	BDR
Position R1, 6 halls								
63–125	0.71	0.51	−0.04	0.59	**0.92	0.38	—	—
125–250	*0.91	0.24	0.54	0.69	0.72	0.64	0.47	—
250–500	*0.86	−0.22	0.54	0.53	0.55	0.52	0.50	—
500–1k	0.74	−0.31	0.42	0.46	0.42	0.43	0.58	−0.37
1k–2k	*0.82	−0.28	0.42	0.48	0.41	0.38	0.78	0.21
2k–4k	*0.87	−0.15	0.30	0.37	0.28	0.35	*0.95	0.28
4k–8k	0.11	−0.05	0.10	0.16	0.01	0.41	0.85	0.27
Position R2, 6 halls								
63–125	0.57	0.41	−0.22	0.68	0.49	−0.25	—	—
125–250	*0.82	*−0.91	0.66	**0.93	0.61	0.28	**0.99	—
250–500	**0.92	**−0.98	0.75	**0.97	0.59	0.62	**0.96	—
500–1k	*0.84	**−0.96	0.80	**0.92	0.58	0.57	*0.92	0.16
1k–2k	**0.93	*−0.89	0.77	*0.92	0.60	0.55	*0.89	0.57
2k–4k	*0.91	−0.68	0.62	*0.89	0.50	0.64	*0.89	0.69
4k–8k	0.57	−0.65	0.57	0.74	0.28	0.68	*0.92	**0.98
Position R3, 6 halls								
63–125	0.61	−0.72	0.71	*0.91	0.31	−0.26	—	—
125–250	*0.81	*−0.87	*0.91	*0.91	0.78	0.09	—	—
250–500	*0.86	*−0.88	*0.81	**0.93	0.70	0.38	—	—
500–1k	*0.88	*−0.85	0.76	*0.89	0.62	0.41	—	—
1k–2k	*0.91	*−0.84	0.80	*0.85	0.67	0.37	—	—
2k–4k	0.79	*−0.83	0.80	0.80	0.61	0.50	—	—
4k–8k	0.07	−0.77	0.77	0.57	0.49	0.56	—	—

The proportion of early lateral energy (LEF) correlates with the perceptual dynamic attributes on a statistically significant level only in a few occasions. The correlation coefficients for clarity (C80) are negative throughout the analysis.

Since the scales of the objective parameters and BTL probabilities are not necessarily comparable, we repeated the correlation analysis with logarithms of BTL values. Despite the variations to the correlation coefficients due to the logarithmic transformation, the parameters showing the most significant correlations remain unchanged. In addition, the correlation analysis was repeated also with the nonparameteric Spearman's rank correlation. The results with the Spearman's correlation correspond also to the presented results with small variations.

V. DISCUSSION

The main results suggested that the overall dynamic response is indeed a recognizable and differentiating perceptual factor between concert hall acoustics. The multitude of attributes provides solid evidence to the assumptions by Beranek[22] and Meyer[29] on the enhancing effect by music dynamics in certain halls. Earlier research has proposed the apparent source width and spaciousness as the foremost sound level-dependent effect.[18,24] However, here the most frequent attributes were related to the different loudness contrasts. That is, the apparent dynamic range may reveal to be one of the most prominent music dynamics-related contributors in concert hall acoustics. The analysis showed that the perceptual dynamic ranges in halls are more homogeneous near to the orchestra. In such positions, the sound field is typically dominated by the direct sound, and the angular spread of the sound sources on stage is wider. The differences between halls, as well as hall typologies, were pronounced at longer physical distances where the proportion of reflected sound to the overall energy is higher.[45] These findings are in agreement with a recent model on the dynamic range variation between concert halls.[5] In position R2, the relative measure (BDR) for the expansion of the dynamic contrasts showed strong correlation with the perceived dynamic range at the high frequencies, as proposed earlier.[5]

The effect of music dynamics on the varying perceived spatial aspects arose as the second notable attribute group. Along with the dynamic range, this effect was mostly observed in the included halls that have relatively narrow, parallel walls and a flat floor. According to preceding studies, the increasing spaciousness occurs more in halls that provide early reflections outside the median plane[24] and, hence, have lower early interaural coherence.[18] The literature often associates these properties to the shoebox-type halls.[22,46,47]

The correlation analysis suggested that, despite of the high correlation coefficients, none of the included objective parameters could conclusively predict the degree of overall perceptual dynamic effects by the included room-acoustic conditions. Early interaural incoherence (BQI) had a high correlation with the ratings by dynamic loudness and width attributes depending on the receiver position. This measure has also been proposed to classify halls by their overall acoustic quality.[22] Contrary to expectations, LEF values showed only moderate correlations, although the local LEF and BQI values correlated mutually at the mid-frequencies ($r > 0.8$). Also, LEF correlations could be more susceptible to small LEF deviations, as the ranges of LEF values were relatively smaller than with BQI. With these observations, the discussion is briefly extended into the relation of these spatial parameters. LEF indicates the proportion of early lateral energy in the sound field and its values are increased the most by reflections arriving from ±90 degree angles azimuth. BQI, in turn, describes the binaural incoherence that builds up by lateral energy. Due to the differences between figure-of-eight and binaural receivers, directions that maximize LEF are not necessarily the most beneficial in increasing BQI and the perceived spatial effect.[21] Instead, a wider range of angles (approximately 30–90 degrees azimuth) have been shown optimal for BQI.[21,49] Such lateral angles also coincide with the directional regions of increased binaural loudness,[5] which is related to the perceived dynamic range.

The overall strength of the concert hall has also been identified as a critical factor for the room-acoustic quality.[20] Here, the positive correlations between the G parameter and the perceptual dynamic effects may result from the following reasons. First, the thresholds for the audibility of early and late lateral reflections are reduced when the overall sound level is higher.[23] Hence, in louder halls the lateral reflections become perceptually more efficient. This may also suggest a connection to the presented correlations with LJ. Second, a higher amplification by room acoustics shifts the dynamic contrast to sensation levels where the difference could be

TABLE IV. Correlations between mean BTL results by separate attribute groups (see Fig. 6) and objective room acoustic parameters in frequency bands. Symbols indicate a statistically significant (unadjusted $p < 0.05$) positive (+) or negative (−) correlation where the correlation coefficient $|r| \geq 0.8$. Objective parameters are denoted with single letters (G: strength; C: clarity (C80); E: early decay time (EDT); J: late lateral energy (LJ); T: reverberation time; L: lateral energy fraction (LEF); I: binaural quality index (BQI): B: binaural dynamic responsiveness (BDR).

Position R1

	All (100.0%)								Loudness (20.9%)								Width (13.9%)								Brigthness (16.4%)								Reverb. (9.7%)								Bass (11.2%)								Clarity (5.8%)								Other (22.1%)							
Hz	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B
63-125					+								+								+																							+																	+			
125-250	+				+				+											+	+															+	+	+	+																									
250-500	+								+																									-	+	+		+	+																		+							
500-1k																																		-	+	+		+	+																									
1k-2k	+						+		+								+						+										+	-	+	+		+	+																		+							
2k-4k	+						+		+						+		+						+											-	+	+		+																									+	
4k-8k							+																+													+		+																										

Position R2

	All (100.0%)								Loudness (23.6%)								Width (16.4%)								Brigthness (9.3%)								Reverb. (12.2%)								Bass (11.9%)								Clarity (7.2%)								Other (19.4%)							
Hz	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B	G	C	E	J	T	L	I	B
63-125																													-												+			+																				
125-250	+			+					+	-		+			+				+		+														+		+				+			+													+			+				
250-500	+	-		+					+	-		+			+		+		+	+															+																						+			+				
500-1k		-							+	-		+			+		+		+																+																						+			+				
1k-2k		-					+		+	-		+			+			-	+																+																						-					+		
2k-4k									+			+			+				+																+																													
4k-8k															+	+																																	+															

Position R3

	All (100.0%)						Loudness (24.5%)						Width (18.2%)						Brigthness (10.7%)						Reverb. (12.6%)						Bass (5.3%)						Clarity (5.7%)						Other (21.8%)					
Hz	G	C	E	J	T	L	G	C	E	J	T	L	G	C	E	J	T	L	G	C	E	J	T	L	G	C	E	J	T	L	G	C	E	J	T	L	G	C	E	J	T	L	G	C	E	J	T	L
63-125				+						+						+																	+					+						-				
125-250	+		+	+			+	-	+	+				-	+	+								+	+			+	+				+		+									-	+			
250-500	+			+			+	-	+	+				-		+									+		+	+	+													-		-	+	+		
500-1k	+			+			+	-		+			+	-		+									+		+	+	+													-		-	+	+		
1k-2k	+						+	-		+			+			+								+	+	-	+		+													-	+	-	+	+	+	
2k-4k	+							-					+			+						+		+			+		+															-	+	+	+	+
4k-8k																				-		+		+																				-	+		+	+

perceived larger due to the unequal intensity difference limen in the equal loudness contours.

From one perspective, the stronger dynamic effects found in the shoebox halls could result solely from their smaller room volumes. Here, the presented shoebox halls have a combination of higher strength and more lateral sound field with respect to other included halls. From another point of view, BQI only considers the early sound field, and is assumed to depend more on the room geometry[46] than on the total volume.[20,50] On this basis, isolating the contributions of room-acoustic strength and spatial properties to the perceptual dynamic effects would require experiments where the parameters could be controlled independently either by synthetic sound fields or by different sets of existing halls.

The reproduced acoustics were measured in unoccupied concert halls. In some halls the audience constitutes a major part of the total absorption, while in other halls the unoccupied seats are designed to simulate the absorbing effect by a seated listener. Therefore the audience's effect on the reverberant sound is not equal in different halls. With regard to the presented results, the hypothetical correction to the reverberation and strength could reduce the differences between hall types by a small amount. Yet, we expect the mutual order to remain similar to the present results. This assumption is supported also by the high correlations with the BQI parameter, which has been shown to vary only little between unoccupied and occupied conditions.[7]

The nature of this experiment required the listeners to compare two differences, while conventional paired comparison listening tests settle for evaluating one difference. The subjects found the experiment challenging at first, but none of the participants reported that the test would have been overwhelmingly difficult. The duration of the experiment was comparable with a typical paired comparison experiments with an equal number of comparisons.

The artificially created difference in the dynamics may, on one hand, exaggerate the perceived differences compared to a gradual *crescendo*. On the other hand, the instrumentation and texture remained practically constant over the dynamic step. This condition is more conservative in contrast to the numerous examples in the symphonic repertoire, where the composers often accentuate sudden increases in dynamics with the expansion in the instrumentation and range of pitches. Such effects are expected to render the contrasts in the sound level and spectrum of the music signal higher than in the present experiment.

The attribute groups derived from the responses resemble the perceptual descriptors discovered in earlier research utilizing less dynamic music signals. This is not a surprising result, as several studies have independently arrived at overlapping collections of perceptual attributes. On this basis, we propose that the comprehensive impression of concert hall acoustics could consist of two layers of perceptual factors. That is, one layer describes the general sound in the hall with static attributes, and another layer describes the variation of the acoustic impression along with varying music through dynamic attributes. The attribute sets are not necessarily identical, nor do the dynamic attributes alter the perception of their static counterparts linearly.

The present study does not consider preference ratings of room acoustics as such. Yet, it is likely that a high degree of dynamic response contributes to a stronger subjective preference by enhancing the music expressivity. The well-known concert hall ratings by Beranek[7] include the currently compared halls in the order of VM-BK-AC-BP. The averages of BTL probabilities corresponds to this order, thus, providing evidence for the relation between the halls' responsiveness to music dynamics and their overall acoustic quality.

VI. CONCLUSIONS

We studied the perception of orchestra dynamics in acoustics of different concert halls. The result of the listening test indicates that the perceptual effects by variation in music dynamics depends on the concert hall acoustics. According to the paired comparisons between three rectangular and three non-rectangular halls, the subjects assessed that the difference in the halls' response to varying orchestra dynamics is frequently described as increased dynamic range and more prominently expanding width and envelopment. These aspects were rated the highest in Vienna Musikverein and Berlin Konzerthaus. The analysis by the hall typology demonstrated that the included shoebox-type rooms emphasize most of the reported perceptual factors more than the compared halls. Furthermore, these perceptual contrasts between hall typologies are pronounced when the distance to the orchestra is increased. In essence, we can conclude that the perceived room-acoustic responsiveness to music dynamics is more prominent with stronger and more lateral sound field impinging the listening position.

Current findings agree with the results reported from earlier listening experiments and observations regarding sound-level-dependent width, differing hall response to dynamics, and the influence of room acoustics to dynamic range. Although Marshall and Barron concluded in 2001 that the relation between early reflections and spatial impression would have been resolved,[48] the effects by reflected sound continue to reveal yet more different perceptual effects.

Performing musicians, in particular, often express that concert halls are musical instruments that the musicians need to play—just like their personal instruments. The results presented in this paper support this claim, as certain halls are perceived to respond to the variations in the played music more than others. The outcome of this study indicates that the acoustics of concert halls are abundant with non-linear perceptual effects. Therefore, instead of how a room impulse response modifies sound, the research should focus also on how concert hall acoustics contribute to music.

ACKNOWLEDGMENTS

This research has received funding from the Academy of Finland, Projects Nos. 257099 and 289300.

[1]E. Schubert, "The fundamental function of music," Musicae Scientiae **13**(2 suppl), 63–81 (2009).

[2]J. A. Sloboda, "Music structure and emotional response: Some empirical findings," Psychol. Music **19**, 110–120 (1991).
[3]A. J. Blood and R. J. Zatorre, "Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion," Proc. Natl. Acad. Sci. U.S.A. **98**(20), 11818–11823 (2001).
[4]H. Owen, *Music Theory Resource Book* (Oxford University Press, New York, NY, 2000).
[5]J. Pätynen, S. Tervo, P. W. Robinson, and T. Lokki, "Concert halls with strong lateral reflections enhance musical dynamics," Proc. Natl. Acad. Sci. U.S.A. **111**(12), 4409–4414 (2014).
[6]T. Lokki and J. Pätynen, "The acoustics of a concert hall as a linear problem," Europhys. News **46**(1), 15–18 (2015).
[7]L. L. Beranek, "Subjective rank-orderings and acoustical measurements for fifty-eight concert halls," Acta Acust. Acust. **89**(3), 494–508 (2003).
[8]W. C. Sabine, *Collected Papers on Acoustics* (Harvard University Press, Cambridge, MA, 1922).
[9]L. L. Beranek, "Concert hall acoustics," J. Acoust. Soc. Am. **92**(1), 1–39 (1992).
[10]E. Kahle, "Validation d'un modèle objectif de la perception de la qualité acoustique dans un ensemble de salles de concerts et d'opéras" ("Validation of the objective model for the perception of acoustic quality of concert halls and operas"), Ph.D. thesis IRCAM, Paris, France, 1995, pp. 1–248.
[11]R. J. Hawkes and H. Douglas, "Subjective acoustic experience in concert auditoria," Acta Acust Acust **24**(5), 235–250 (1971).
[12]M. Barron, "Subjective study of British symphony concert halls," Acta Acust. Acust. **66**(1), 1–14 (1988).
[13]T. Lokki, J. Pätynen, A. Kuusinen, and S. Tervo, "Disentangling preference ratings of concert hall acoustics using subjective sensory profiles," J. Acoust. Soc. Am. **132**(5), 3148–3161 (2012).
[14]ISO 3382-1:2009. *Acoustics—Measurement of room acoustic parameters—I: Performance spaces* (International Organization for Standardization, Geneva, Switzerland, 2009).
[15]R. Lacatis, A. Gimenez, A. B. Sevillano, S. Cerda, J. Romero, and R. Cibrian, "Historical and chronological evolution of the concert hall acoustics parameters," in *Acoustics'08*, Paris, France (June 29–July 4, 2008), pp. 2151–2156.
[16]A. H. Marshall, "A note on the importance of room cross-section in concert halls," J. Sound Vib. **5**(1), 100–112 (1967).
[17]A. H. Marshall, "Levels of reflection masking in concert halls," J. Sound Vib. **7**(1), 116–118 (1968).
[18]W. V. Keet, "The influence of early lateral reflections on the spatial impression," in *Proc. 6th International Congress on Acoustics*, Tokyo, Japan (1968), Vol. 3, pp. E53–E56.
[19]M. Barron, "The subjective effects of first reflections in concert halls—The need for lateral reflections," J. Sound Vib. **15**(4), 475–494 (1971).
[20]M. R. Schroeder, D. Gottlob, and K. F. Siebrasse, "Comparative study of European concert halls: Correlation of subjective preference with geometric and acoustic parameters," J. Acoust. Soc. Am. **56**(4), 1195–1201 (1974).
[21]T. Okano, L. L. Beranek, and T. Hidaka, "Relations among interaural cross-correlation coefficient (IACCe), lateral fraction (LFe), and apparent source width (ASW) in concert halls," J. Acoust. Soc. Am. **104**(1), 255–265 (1998).
[22]L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture* (Springer, New York, 2004), pp. 1–661.
[23]R. Wettschureck, "Über die Abhängigkeit raumakustischer Wahrnehmungen von der Lautstärke" ("On the dependence of room acoustic perception by sound level"), Ph.D. thesis, TU Berlin, Germany (1976), pp. 1–129.
[24]W. Kuhl, "Räumlichkeit als Komponente des Raumeindrucks" ("Spaciousness as a component of total room impression"), Acta Acust. Acust. **40**(3), 167–181 (1978).
[25]J. Bradley and G. A. Soulodre, "Objective measures of listener envelopment," J. Acoust. Soc. Am. **98**(5), 2590–2597 (1995).
[26]M. Barron and A. H. Marshall, "Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure," J. Sound Vib. **77**(2), 211–232 (1981).
[27]J. Blauert and W. Lindemann, "Auditory spaciousness—Some further psychoacoustic analyses," J. Acoust. Soc. Am. **80**(2), 533–542 (1986).
[28]D. Luce, "Dynamic spectrum changes of orchestral instruments," J. Audio Eng. Soc. **23**(7), 565–568 (1975).
[29]J. Meyer, *Acoustics and the Performance of Music* (Springer, New York, 2009).
[30]B. C. J. Moore and B. R. Glasberg, "Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns," Hear. Res. **28**, 209–225 (1987).
[31]In the individual vocabulary profiling method, the subjects elicit their own perceptual attributes before assessing the stimuli using these criteria. See Ref. 32 for examples.
[32]T. Lokki, J. Pätynen, A. Kuusinen, and S. Tervo, "Concert hall acoustics: Repertoire, listening position, and individual taste of the listeners influence the qualitative attributes and preferences," J. Acoust. Soc. Am. **140**(1), 551–562 (2016).
[33]M. Sams, R. Hari, J. Rif, and J. Kuutila, "The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence," J. Cognit. Neurosci. **5**(3), 363–370 (1993).
[34]J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," Acta Acust. Acust. **94**(6), 856–865 (2008).
[35]J. Pätynen, S. Tervo, and T. Lokki, "Simulation of the violin section sound based on the analysis of orchestra performance," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Piscataway, NJ (October 16–19, 2011), pp. 173–176.
[36]J. Pätynen, "A virtual symphony orchestra for studies on concert hall acoustics," Ph.D. thesis, Aalto University School of Science, 2011, pp. 1–111.
[37]T. Lokki, "Tasting music like wine," Phys. Today **67**(1), 27–32 (2014).
[38]S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," J. Audio. Eng. Soc. **61**(1/2), 16–27 (2013).
[39]J. Pätynen, S. Tervo, and T. Lokki, "Amplitude panning decreases spectral brightness with concert hall auralizations," in *Proc. 55th Audio Eng. Soc. Conference, Helsinki, Finland*, Audio Eng. Soc., New York (August 27–29, 2014), paper no. 49.
[40]P. Courcoux and M. Semenou, "Preference data analysis using a paired comparison model," Food Qual. Prefer. **8**(5), 353–358 (1997).
[41]R. A. Bradley and M. E. Terry, "Rank analysis of incomplete block designs: I. The method of paired comparisons," Biometrika **39**, 324–345 (1952).
[42]R. D. Luce, *Individual Choice Behaviour: A Theoretical Analysis* (Wiley, New York, 1959), pp. 1–176.
[43]S. Choisel and F. Wickelmaier, "Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preference," J. Acoust. Soc. Am. **121**(1), 388–400 (2007).
[44]Y. Benjamini and Y. Hockberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," J. R. Stat. Soc. **57**(1), 289–300 (1995).
[45]M. Barron and L.-J. Lee, "Energy relations in concert auditoriums. I," J. Acoust. Soc. Am. **84**(2), 618–628 (1988).
[46]M. Barron, "Late lateral energy fractions and the envelopment question in concert halls," Appl. Acoust. **62**(2), 185–202 (2001).
[47]M. Long, "What is so special about shoebox halls? Envelopment, Envelopment, Envelopment," Acoust. Today **5**(2), 21–25 (2009).
[48]A. H. Marshall and M. Barron, "Spatial responsiveness in concert halls and the origins of spatial impression," Appl. Acoust. **62**(2), 91–108 (2001).
[49]Y. Ando, *Concert Hall Acoustics* (Springer, Berlin, Germany, 1995), pp. 1–151.
[50]T. Hidaka, "Supplemental data of dependence of objective room acoustical parameters on source and receiver positions at field measurement," Acoust. Sci. Technol. **26**(2), 128–135 (2005).