
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Haapaniemi, Aki; Lokki, Tapio

Identifying concert halls from source presence vs room presence

Published in:
Journal of the Acoustical Society of America

DOI:
[10.1121/1.4879671](https://doi.org/10.1121/1.4879671)

Published: 01/01/2014

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

Please cite the original version:
Haapaniemi, A., & Lokki, T. (2014). Identifying concert halls from source presence vs room presence. *Journal of the Acoustical Society of America*, 135(6), EL311-EL317. <https://doi.org/10.1121/1.4879671>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Identifying concert halls from source presence vs room presence

Aki Haapaniemi, and Tapio Lokki

Citation: [The Journal of the Acoustical Society of America](#) **135**, EL311 (2014); doi: 10.1121/1.4879671

View online: <https://doi.org/10.1121/1.4879671>

View Table of Contents: <http://asa.scitation.org/toc/jas/135/6>

Published by the [Acoustical Society of America](#)

Articles you may be interested in

[Concert halls with strong and lateral sound increase the emotional impact of orchestra music](#)

[The Journal of the Acoustical Society of America](#) **139**, 1214 (2016); 10.1121/1.4944038

[Engaging concert hall acoustics is made up of temporal envelope preserving reflections](#)

[The Journal of the Acoustical Society of America](#) **129**, EL223 (2011); 10.1121/1.3579145

[Tasting music like wine: Sensory evaluation of concert halls](#)

[Physics Today](#) **67**, 27 (2014); 10.1063/PT.3.2242

[Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses](#)

[The Journal of the Acoustical Society of America](#) **133**, 842 (2013); 10.1121/1.4770260

[Disentangling preference ratings of concert hall acoustics using subjective sensory profiles](#)

[The Journal of the Acoustical Society of America](#) **132**, 3148 (2012); 10.1121/1.4756826

[Lateral reflections are favorable in concert halls due to binaural loudness](#)

[The Journal of the Acoustical Society of America](#) **130**, EL345 (2011); 10.1121/1.3647866

Identifying concert halls from source presence vs room presence

Aki Haapaniemi^{a)} and Tapio Lokki

Department of Media Technology, Aalto University School of Science, P.O. Box 15500,
FI-00076 Aalto, Finland
aki.haapaniemi@aalto.fi, tapio.lokki@aalto.fi

Abstract: Identification of concert halls was studied to uncover whether the early or late part of the acoustic response is more salient in a hall's fingerprint. A listening test was conducted with auralizations of measured halls using full, hybrid, and truncated impulse responses convolved with anechoic symphonic music. Subjects identified halls more reliably based on differences in early responses rather than late responses, although varying the late response had more effect on acoustic parameters. The results suggest that in a typical situation with running symphonic music, the early response determines the perceptual fingerprint of a hall more than the late response.

© 2014 Acoustical Society of America

PACS numbers: 43.55.Hy, 43.55.Gx [NX]

Date Received: February 7, 2014 Date Accepted: April 28, 2014

1. Introduction

When listening to music in a concert hall, the perception of acoustics can be divided into two basic constituents: *source presence* and *room presence*.¹ The direct sound and the early part of the hall's acoustical response create the source presence, whereas the late response creates the room presence. The early part is generally taken to be the hall's response up to until 80 ms after the direct sound, which is also reflected in the ISO 3382-1:2009 standard parameters.² However, the perceptual crossover time between the early and late response varies with the type of musical content, and is more or less governed by the *echo threshold*, i.e., the limit of the *fusion* aspect of the *precedence effect*.³

The precedence effect is a psychoacoustic phenomenon that aids the hearing system in localizing sound sources in reflective environments. With a sufficiently short time delay, a direct sound followed by a reflection will not be perceived as two separate sounds, but rather be perceptually fused into a single sound. However, the presence of the reflection affects the loudness, timbre, and spatial extent of the auditory event.⁴ When the time delay between the direct sound and the reflection exceeds the echo threshold, fusion no longer occurs and the reflection is perceived separately from the direct sound. In a concert hall the late response generally consists of myriad reflections that come from all directions and at an exponentially decaying level. The reflections are then not heard as separate echoes, which are generally detrimental to acoustics,⁵ but rather as a pleasant surrounding room presence that adds a sense of depth to the music.

With symphonic music, the situation is more complex with overlapping sound sources of different spectral and time-domain characteristics. Because the nature of music is continuous, the complete reflection decay process is only heard fully when the music stops. Yet the underlying phenomenon of fusion remains the same and affects the way the contributions of individual reflections are perceptually processed into the foreground stream (source presence) and the background stream (room presence).⁶

^{a)} Author to whom correspondence should be addressed.

The ISO 3382-1 parameters have been developed to address the need for physical data that could explain the perceptual differences between the acoustics of various halls. The parameters facilitate classification of concert halls by considering various aspects of the early and late response and their proportional balance. However, there are indications⁷⁻⁹ that the parameters fall short of explaining the more delicate aspects of subjective perception and preference, especially those that differentiate between “good” and “excellent” acoustics.

Therefore the question of what makes the acoustical “fingerprint” of a hall—i.e., what makes it recognizable from other halls—is still largely unanswered. In this paper, the question is explored by studying whether the most readily identifiable aspects of a concert hall’s acoustics are contained in the source presence or the room presence, i.e., the early or late response. For this purpose, a listening test was conducted with auralized acoustics of real concert halls.

2. Methods

The listening test consisted of identification tasks with auralized acoustics of eight unoccupied concert halls (listed in Table 1) using full, truncated, and hybrid spatial room impulse responses (SRIRs). The original SRIRs were derived from measurements of two listening positions in each hall with a calibrated loudspeaker orchestra.¹⁰ Figure 1(a) shows the positioning of the loudspeaker orchestra and the measurement positions P1 (11 m) and P2 (19 m) in one of the halls. P1 and P2 were measured at the same distances from the loudspeaker orchestra in each hall, thus enabling direct comparison.

Capturing the acoustics for one listening position consisted of recording a logarithmic sweep for each channel of the loudspeaker orchestra with a three-dimensional microphone array of six omnidirectional microphones. The impulse responses were analyzed with the spatial decomposition method¹¹ to estimate the direction of arrival of sound energy as a function of time and frequency. A 24-channel SRIR (i.e., a convolution reverb) was created for the multichannel system by using the vector base amplitude panning¹² method for spatial sound placement.

Three types of SRIR were used for the listening tests: full, truncated (up to 80 ms) and hybrid (two halls crossfaded at 80 ms). The truncated responses were created out of the first 80 ms of the response after the direct sound for each channel of the loudspeaker orchestra. A linear fade out was done from 75 to 80 ms to avoid discontinuities. The hybrid responses were created by taking the early response from one hall, the late response from another hall at the same position, and splicing them together channel-wise with a linear crossfade from 75 to 85 ms after the direct sound. An illustration of the splicing procedure is shown in Fig. 1(b). All combinations of hybrid responses were created for the two positions at each hall.

Table 1. A list of the halls and their parameter values [averaged over the 500 Hz and 1 kHz octave bands except for J_{LF} , and L_J (energy averaged) which are averaged over the 125 Hz, 250 Hz, 500 Hz, and 1 kHz octave bands]. The two values are for positions P1 (11 m) and P2 (19 m), respectively.

Hall	EDT (s)	C_{80} (dB)	J_{LF}	G (dB)	L_J (dB)	G_{80}
Amsterdam Concertgebouw (AC)	2.5/2.4	-2.0/-2.7	0.19/0.23	3.3/2.2	-3.1/-3.7	-0.9/-2.4
Berlin Konzerthaus (BK)	2.1/2.1	-1.7/-2.3	0.25/0.26	2.9/1.8	-2.6/-3.4	-1.1/-2.7
Brussels Palais des Beaux-arts (BB)	1.8/1.6	0.7/0.7	0.18/0.20	3.4/3.4	-4.1/-3.4	0.6/0.6
Berlin Philharmonie (BP)	2.1/1.9	2.4/0.2	0.09/0.14	2.5/1.2	-5.6/-5.7	0.4/-1.8
Helsinki Music Center (HM)	2.1/2.2	0.8/0.6	0.12/0.14	3.1/0.8	-3.9/-6.2	0.4/-2.0
Lahti Sibelius Hall (LS)	2.0/1.9	0.2/0.6	0.31/0.30	4.1/4.3	-2.2/-2.1	1.1/1.4
Stuttgart Beethovensaal (SB)	2.3/2.1	0.6/-1.2	0.11/0.11	2.1/1.1	-5.5/-5.8	-0.7/-2.6
Vienna Musikverein (VM)	2.8/2.9	-2.0/-4.0	0.23/0.19	4.6/3.3	-1.7/-2.5	0.5/-2.3

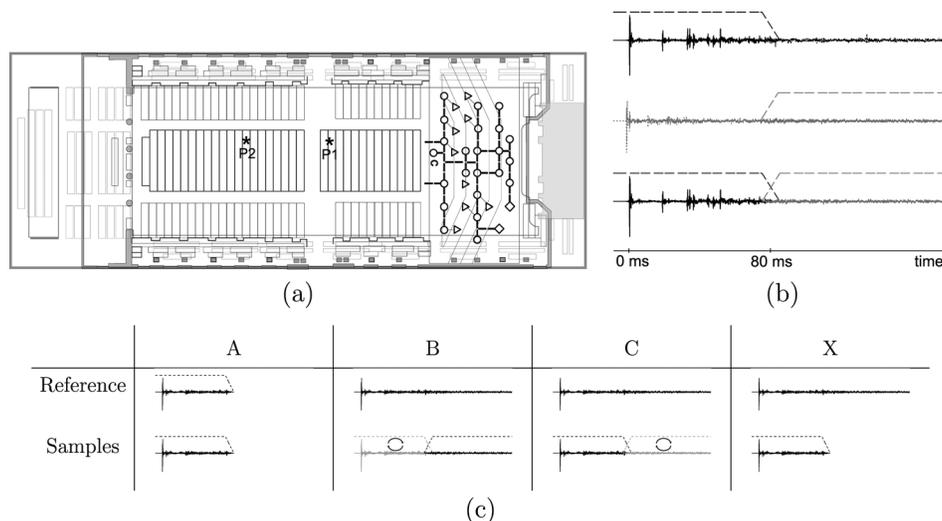


Fig. 1. (a) Positioning of the loudspeaker orchestra and the measurement positions P1 and P2 in Vienna Musikverein. (b) An illustration of the splicing procedure used to form the hybrid room impulse response of one channel. (c) Summary of the round types.

The audio samples were created by convolving the SRIRs with corresponding channels of anechoic recordings of a symphony orchestra.¹³ The recordings were of bars 33–36 (5.4 s) of the second movement of Anton Bruckner's symphony no. 8. The excerpt is in *fortissimo* (*ff*) and all the instrument groups are playing simultaneously. The resulting terminal reverberation tails were cut off in order for the samples to be representative of the general listening situation where only running reverberation is heard.⁵

The samples were reproduced in an acoustically treated listening room with a 24-channel setup of 20 Genelec 8020B and 4 Genelec 1029A loudspeakers (Genelec, Iisalmi, Finland). The loudspeakers were positioned on five levels of elevation: at 0° (ear level) [azimuth angles 0°, ±22.5°, ±45°, ±67.5°, ±90°, ±135°, 180°], 30° [azimuth angles 0°, ±45°, ±135°], 45° [azimuth angles ±90°], 90° (on top of the listening position), and at -35° [azimuth angles ±40°, ±150°]. The nominal loudspeaker distance was 1.5 m from the listening position. The loudspeakers directly on top and those elevated on either side of the listening position were positioned at 1.2 m for practical reasons. The distance differences were compensated with appropriate delays and attenuation in the loudspeaker signals. The output gains of the individual loudspeakers were calibrated with 0.1 dB accuracy.

The listening room has a reverberation time of 0.1 s at mid-frequencies and the background noise level is 31 dB (A-weighted). The peak-to-peak level difference between the direct sound and the strongest early reflection in the 1 to 8 kHz band was found to be 12.8 dB on average across all the loudspeakers, therefore complying with the ITU-R BS.1116-1 recommendation for subjective audio evaluation systems (early reflection peaks at least 10 dB below the direct sound peak). The sound pressure level (A-weighted, slow) was measured with a calibrated sound level meter (SINUS Messtechnik GmbH, Leipzig, Germany) at the listening position, and was found to vary between 70 and 77 dB for the samples. No attempt was made to regulate the differences in level because they were caused by the natural differences in the acoustics of the halls.

3. Standard parameters

In order to understand the differences between the halls and their hybrids, the standard parameter values were calculated for the five aspects of listener experience:²

reverberance (EDT), clarity (C_{80}), apparent source width (J_{LF}), sound level (G), and listener envelopment (L_J). An additional parameter, G_{80} , was included as a measure of the differences in the truncated responses. It is similar to G but takes only the first 80 ms of the response into account. The other parameters were omitted for the truncated response: There would be no means to calculate EDT, C_{80} would be infinite, J_{LF} would stay the same, and L_J would be minus infinity. The parameter values for receiver positions P1 and P2 at each hall are listed in Table 1 and Fig. 2 shows the same values with the values calculated for the hybrid responses.

It should be kept in mind that the parameter values presented are of a slightly different nature due to using the loudspeaker orchestra for the measurements instead of a single omnidirectional loudspeaker. They are more representative of the hall in performance with a real orchestra but not directly comparable to values measured using the typical method of one omnidirectional source.

4. Listening test

The listening test consisted of 4 separate rounds of 16 identification tasks each. Each round had identification tasks for positions P1 and P2 from each of the 8 halls ($2 \times 8 = 16$). In each task, the subject was presented with a reference sample and four samples out of which to identify the one corresponding to the reference. The four samples always included one that matched (fully in A, B, and C and partially in X) the reference sample and three other samples that were randomly drawn from the pool of other halls or hybrids, depending on the round. The round types are summarized in Fig. 1(c).

In round A all the samples were processed with truncated responses and one of the samples was identical to the reference. The listener's task was to distinguish the reference by the source presence. In round B the reference was processed with a full response and one of the samples was identical to the reference. The three other samples were processed with hybrid responses that had the same late response as the reference, but different early responses. The listener's task was to distinguish the reference by the source presence. Round C was contrary to B: The late response was varied while the

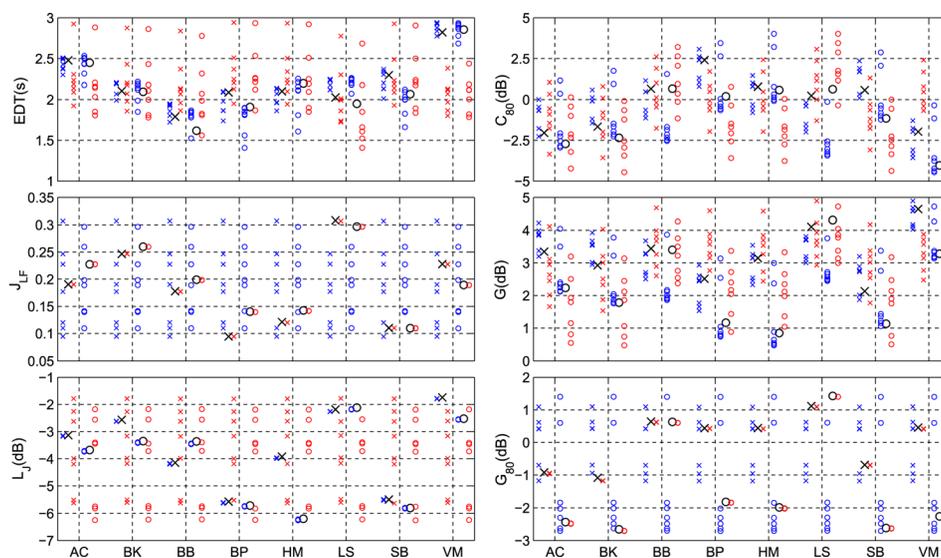


Fig. 2. (Color online) The parameter values for the two positions P1 (“×”) and P2 (“○”) at each hall (large symbols) and their hybrid versions (small symbols: left side = variable early response, right side = variable late response). The parameters are averaged over the 500 Hz and 1 kHz octave bands except for J_{LF} and L_J (energy averaged) that are averaged over the 125 Hz, 250 Hz, 500 Hz, and 1 kHz octave bands. The abbreviations for the halls are as listed in Table 1.

early response was kept the same. The listener had to identify the reference on the basis of the room presence. In round X none of the samples fully matched the reference. The reference was processed with a full response while the other samples were processed with truncated responses. One of the truncated responses was the same as the early response of the reference. The listener's task was therefore to identify the closest match on the basis of a similar source presence.

Before the listening test, the subjects were given brief written instructions that described the task and the proceeding of the test. The subjects were told that the test was about identification of acoustics, but the nature of the samples was not discussed. They were also told that in rounds A, B, and C their task was to identify an exact match while in round X the task was to find the closest match. The subject was free to listen to the samples for as many times as necessary and in any order before making the decision. The time spent on the tasks was monitored without the subjects' knowledge.

The test started with a short practice round of four identification tasks with truncated responses (identical to tasks of round A), after which came round A, followed by B and C in either order, and finally X. The order of rounds B and C were alternated between subjects. For one subject, the order of rounds A and B were accidentally switched, but the effect of this on the results is negligible. After each round was a short break when the subject was asked to answer one question on an answer form relating to the completed round: "How did you distinguish between the samples, i.e., what kind of differences did you use to identify the sample?" After the test the subject was also asked to rate the difficulty of the individual rounds on a scale of 1 to 5 (very easy to very difficult).

5. Results and discussion

Twelve male subjects of ages between 27 and 34 took part in the listening test. They were all acoustical engineers and/or musicians and had experience in critical listening. No one had any hearing impairments as far as they were aware. The results are shown in Fig. 3(a) in the form of percentage of correct identifications per subject and the median value for each round. A Kruskal-Wallis one-way analysis of variance shows that the differences between the rank means of the round results are significant ($\chi^2 = 23.41$, $p < 0.001$). Further analysis using Tukey's least significant difference criterion ($p = 0.05$) shows that the results of both rounds A and B differ significantly from those of rounds C and X. However, the difference between the results of rounds A and B is not significant, and the same applies to C and X. The corresponding times spent doing the rounds and the median values are shown in Fig. 3(b).

The subjects had no trouble in identifying the halls on the basis of the source presence in the case of the truncated halls in round A, as expected. However, they did equally well in round B in the presence of reverberation, but spent more time on the tasks. Thus the characteristic early response of a hall was easily identified by the subjects even in the presence of reverberation. In round C, the identifications were

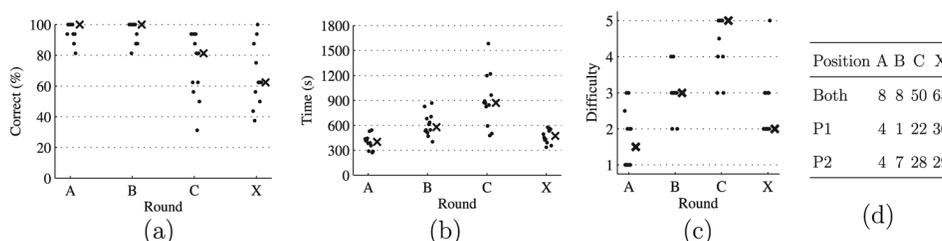


Fig. 3. (a) Percentage of correct answers per each round. (b) Spent time per each round. (c) The subjects' assessment of the difficulty of each round on a scale of 1 to 5 (very easy to very difficult). (d) Amount of incorrect identifications per position per round. In (a), (b), and (c) the "×" marks show the median values for each round.

significantly less reliable than in A and B, although the most time was spent. The subjects' assessment of the difficulty of the rounds [Fig. 3(c)] is in line with the results of rounds A, B, and C: A was felt to be easy, B moderate, and C difficult. In light of the subjective difficulty of round C the identification results were surprisingly good. It seems that the subjects were to some extent able to use repeated listening to focus on differences in the room presence characteristics even though the task was felt to be challenging.

For round X the median correct percentage was the lowest although overall the results resemble those of round C. The difference is that not much time was spent on the round and it was felt to be easy or moderate by most subjects. Two possible reasons for the discrepancy come to mind. Round X was done last by all subjects, so the results might carry some effects of fatigue. Furthermore, it was the only round where the samples did not include an exact match of the reference: The reference was processed with a full response and the samples were processed with truncated responses. It is likely that the difference in the context (room presence/no room presence) made it more difficult to connect the right source presence with the reference.

An interesting observation can be made about the parameter values (Fig. 2): The average deviation of the values is on the whole greater with variable late response than with variable early response. The exceptions are J_{LF} and G_{80} which are exclusive to the early response (likewise L_J is exclusive to the late response). Therefore the difference in the identification results of rounds B and C is not explained by the corresponding deviations in the parameter values.

Comparing the identification results between the two positions P1 and P2 reveals that the median values for the percentage of correct answers is identical for all rounds. However, a comparison of the amount of misidentifications [Fig. 3(d)] shows that for round B most of them were for position P2 (7 out of 96), which implies that at P1 (1 out of 96) the identification of the hall from the early response may be easier than at P2. This could be explained by P2 being further away from the orchestra, and thus more reverberant and quieter, with smaller values of C_{80} and G (and G_{80}) on average. These factors may contribute to make the assessment of differences in the source presence more difficult. However, the deviation of the parameters is greater for P2 compared to P1, which suggests that the contrary result would be more likely.

Finally, the subjects' answers to the question of what kind of differences they based their identifications on deserve some attention. For rounds A and B the two attributes that were by far the most often quoted were timbre and auditory width. Other often mentioned attributes were apparent distance, loudness, and the perceived amount and quality of bass. For round X the two top attributes were also timbre and width, although they were less dominant and more variety was seen in the other listed attributes. For round C, the top attributes were bass and timbre, followed by reverberance and width. It seems evident that timbre was on the whole among the most prominent variables in the listening test, and especially linked with the early response. Furthermore, none of the standard parameters specifically address timbre, which makes it more difficult to assess. While the subjects' comments do not yet warrant further conclusions, they serve as food for thought for planning future studies.

Acknowledgments

Thanks to Antti Kuusinen for the listening test interface. The research leading to these results has received funding from the Academy of Finland, Project No. 257099.

References and links

- ¹E. Kahle, "Room acoustical quality of concert halls: perceptual factors and acoustic criteria—return from experience," in *International Symposium on Room Acoustics (ISRA 2013)*, Toronto, Canada, June 9–11, 2013, Paper No. P074.
- ²ISO 3382-1:2009, *Acoustics—Measurement of room acoustic parameters—Part 1: Performance spaces* (International Organization for Standardization, Geneva, Switzerland, 2009), pp. 1–26.

- ³R. Y. Litovsky, H. S. Colburn, W. A. Yost, and S. J. Guzman, "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633–1654 (1999).
- ⁴J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization, revised edition* (MIT Press, Cambridge, MA, 1997), pp. 223–224.
- ⁵M. Barron, *Auditorium Acoustics and Architectural Design*, 2nd ed. (Spon Press, New York, 2009), p. 28.
- ⁶D. Griesinger, "The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces," *Acta Acust. Acust.* **83**, 721–731 (1997).
- ⁷T. Lokki, J. Pätynen, S. Tervo, S. Siltanen, and L. Savioja, "Engaging concert hall acoustics is made up of temporal envelope preserving reflections," *J. Acoust. Soc. Am.* **129**, EL223–EL228 (2011).
- ⁸J. S. Bradley, "Review of objective room acoustics and future needs," *Appl. Acoust.* **72**, 713–720 (2011).
- ⁹L. Kirkegaard and T. Gulsrud, "In search of a new paradigm: How do our parameters and measurement techniques constrain approaches to concert hall design?," *Acoust. Today* **7**, 7–14 (2011).
- ¹⁰T. Lokki, J. Pätynen, A. Kuusinen, H. Vertanen, and S. Tervo, "Concert hall acoustics assessment with individually elicited attributes," *J. Acoust. Soc. Am.* **130**, 835–849 (2011).
- ¹¹S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," *J. Audio Eng. Soc.* **61**, 17–28 (2013).
- ¹²V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.* **45**, 456–466 (1997).
- ¹³J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," *Acta. Acust. Acust.* **94**, 856–865 (2008).