Audio Engineering Society

Convention Express Paper 10270

Presented at the AES 160th Convention
2026 May 28–30, Copenhagen, Denmark

This Express Paper was selected on the basis of a submitted synopsis that has been peer-reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This Express Paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Sound Diffusion Properties of a Bending-Wave Loudspeaker Compared with a Conventional Speaker

Rina Mizukami[1] and Kazuhiko Kawahara[2]

[1]*Graduate School of Design, Kyushu University, Fukuoka, Japan*
[2]*Faculty of Design, Kyushu University, Fukuoka, Japan*

Correspondence should be addressed to Rina Mizukami (mizuna.mikan.0221@icloud.com)

ABSTRACT

The Panel-shaped Bending Wave Loudspeaker was proposed recently by Kawahara[1]. The authors conducted an objective evaluation of the diffusion characteristics of Bending Wave Loudspeakers (BWL) using the degree of interaural cross-correlation (DICC) in this paper. Conventional speakers exhibit strong directionality and rely on room reflections to create a spatial impression. In contrast, BWLs are considered less susceptible to room reflections due to complex mode vibrations across the entire diaphragm. To quantify this characteristic, the authors recorded sound in a real-world environment using a head-and-torso simulator (HATS). The authors compared the DICC of BWL with that of a conventional speaker. The results showed that the BWL exhibited significantly lower DICC values than the conventional loudspeaker at the front position (Center) under both broadband noise and music conditions, confirming its high diffusivity. Furthermore, this difference exceeded the Just Noticeable Difference (JND) for spatial perception, suggesting it is also significant to the human ear. In addition, analysis separating early reflections and late reflections suggested differences in diffusion characteristics between conventional speakers and BWL.

1 Introduction

1.1 Distributed Mode Loudspeaker and other Panel-shaped loudspeaker

Distributed Mode Loudspeakers (DMLs) utilize bending waves in the diaphragm to produce sound. Unlike conventional cone-diaphragm loudspeakers, DMLs radiate sound by generating bending waves in the diaphragm. Research into their characteristics and applications has been conducted since around 1990.

New Transducers Limited [2] announced the Distributed Mode Loudspeaker (DML) in the 1990s, introducing its acoustic characteristics and examples of its application in products. Kawahara [3] introduced the advantages of the DML to the Acoustical Society of Japan. The DML consists of a diaphragm, an actuator,

and a frame supporting the diaphragm. Conventional loudspeakers uniformly vibrate the cone paper back and forth. In contrast, the DML deliberately divides the diaphragm into sections for vibration. This article is about creating a diffuse sound field. Kawahara also stated that a quantitative evaluation of its properties and advantages is necessary for DML to be widely adopted.

Regarding the spatial representation of DML, Harris et al. [4] conducted a detailed investigation. They performed localization identification tests for stereo sources using conventional loudspeakers and DML in a listening room. The results demonstrated that DML maintains stable localization not only at the sweet spot but also at multiple listening positions, including off-axis seats, and is less susceptible to adverse effects from room reflections. Previous studies have investigated the relationship between the physical characteristics of DML and subjective spatial perception.

For example, Kleczkowski et al. [5] conducted a listening experiment comparing DML with cone loudspeakers and reported that DML outperformed cone loudspeakers in terms of listner envelopment (LEV). Importantly, they attributed this spatial advantage to the lower interaural correlation coefficient measured with DML. These results suggest that a low interaural correlation coefficient is an important physical indicator of a diffuse sound field and a rich spatial impression.

Regarding sound quality and design technology, Azima et al. [6] conducted research. They analyzed historical challenges, including irregularities in frequency response and unique directional control issues. They concluded that by combining mechanical approaches—such as actuator placement and diaphragm tuning—with electrical signal-processing approaches, DML can achieve sound quality competitive with that of conventional loudspeakers.

Heilemann et al. [7] suggested that flat-panel loudspeakers are highly compatible with display-equipped products due to their thinner profile compared to conventional cone loudspeakers. They also noted that sound quality issues, which were previously a challenge, are being improved through innovative design methods. Driven by these advantages, the development of novel flat-panel loudspeakers to replace conventional DMLs has actively progressed in recent years.

In this study, the authors focus on the Bending-Wave Loudspeaker (BWL) as one of these emerging approaches.

1.2 Bending Wave Loudspeaker

Based on the design proposed by Kawahara [1], the prototype BWL evaluated in this study was constructed using a corrugated plastic board as the diaphragm and a paulownia wood frame. An electrodynamic exciter (DAEX25, Dayton Audio) was attached to the panel to generate bending waves.(Figure 1) The BWL radiates sound by utilizing bending waves generated in the diaphragm rather than uniformly moving the entire diaphragm. This diffusive property was interpreted as "radiating a blurred sound." Therefore, the authors hypothesized that it contributes to an increase in auditory source width (ASW).

Inoue et al. [8] hypothesized that the ambiguous localization of the BWL enhances the immersive quality of applause sounds and conducted subjective evaluation experiments. Their results demonstrated that employing a BWL as a center speaker provides listeners with an enveloping sound spread.

The authors measured the frequency-dependent directivity characteristics using the same BWL prototype employed in this study. Figure 2, 3, and 4 compares the directivity patterns of the BWL and a conventional loudspeaker (YAMAHA HS5) at 500 Hz, 1 kHz, and 2 kHz. The figure-eight shape observed in the BWL's directivity at 500 Hz is characteristic of its enclosureless design. As the frequency increases to 2 kHz, the HS5 exhibits a decrease in rearward radiation compared to the front, whereas the BWL radiates sound over a wide area with an intensity comparable to the frontal direction. Maintaining this wide directivity at higher frequencies is a defining characteristic of the BWL, which is considered to be a key factor contributing to the spatial impressions discussed in section 2.1.

1.3 Structure of This Paper

In this study, as a preliminary step toward the subjective evaluation experiment for BWL, the authors measured the degree of Interaural cross-correlation (DICC) across frequency bands. First, the authors introduced the definition of DICC and its relationship with spatial impressions in section 2.1. Second, the authors report on the measurement methodology used in the laboratory where the subjective evaluation experiment was conducted and the resulting DICC values described in section 5. Subsequently, based on the definitions of ASW and LEV, the authors conducted an analysis

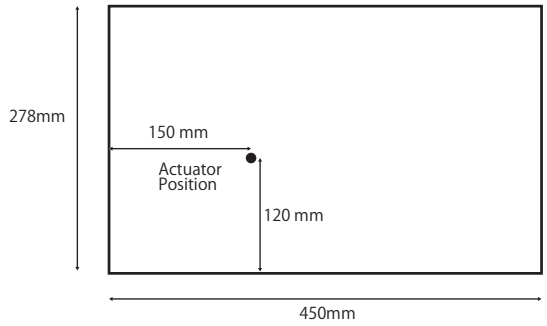


Fig. 1: Dimensional Drawing of a Bending Wave Loudspeaker (BWL)

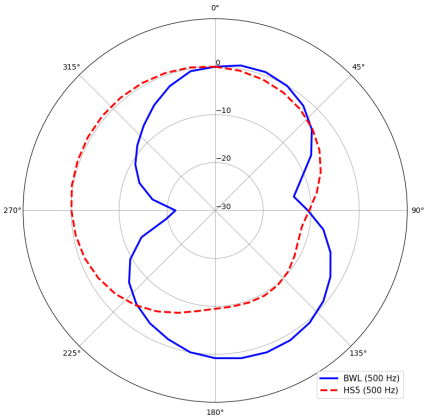


Fig. 2: The Anechoic Polar Pattern of the BWL(solid blue) and the conventional loudspeaker, HS5(dashed red), at 500 Hz.

using numerical values obtained from recorded data segmented into time windows described in section 5.2. Finally, the authors link differences in frequency and articulation mechanisms with differences in binaural correlation, concluding on the optimal environment and frequency bands for BWL evaluation experiments.

2 Theoretical Background on Measuring Spatial Impression

2.1 Evaluation Spatial Impression

Spatial impression is one of the subjective elements of sound quality in acoustic spaces such as concert halls. Hanyu [9] defines this spatial impression as consisting

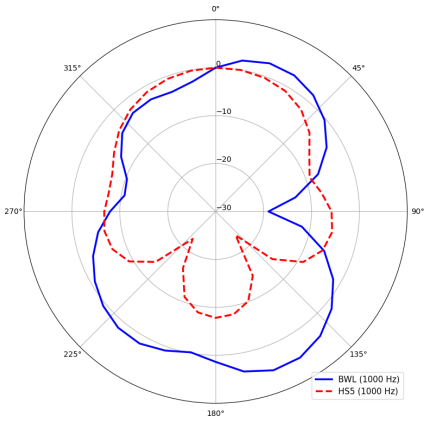


Fig. 3: The Anechoic Polar Pattern of the BWL(solid blue) and the conventional loudspeaker, HS5(dashed red), at 1 kHz.

of two aspects: auditory source width (ASW), which indicates the perceived spread of the sound source, and LEV, which suggests the sensation of being enveloped by sound. The interaural cross-correlation function (IACF), which indicates the dissimilarity of sounds received between the two ears, has been used as a physical metric for objectively evaluating spatial impression.

Marshall [10] demonstrated that varying the time window of the IACF processing allows access to different reflected sounds, stating that early reflections from 0 ms to 80 ms influence ASW, while late reflections beyond 80 ms influence LEV.

Furthermore, Blauert and Lindemann [11] conducted sound localization experiments using pink noise and band noise, confirming that as binaural correlation decreases, the sound image generated within the head becomes more separated and diffuse. This research provides theoretical support for the hypothesis that low binaural correlation directly correlates with increased ASW. However, previous studies reported that the conventional IACC showed low correlation with ASW.

Morimoto and Iida [12] experimented in an anechoic chamber, playing classical music in stereo to investigate the correlation between the binaural correlation function and ASW. They found no correlation between conventional ICC and ASW. However, they discovered that removing the ear simulator and A-weighting correction yielded a high correlation with ASW. They termed this metric the interaural correlation DICC. The

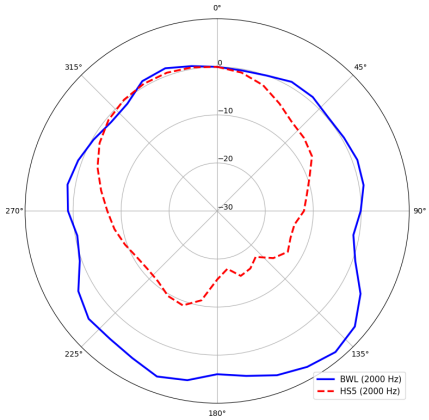


Fig. 4: The Anechoic Polar Pattern of the BWL(solid blue) and the conventional loudspeaker, HS5(dashed red), at 2 kHz.

authors adopt DICC as the ASW evaluation metric for BWL. Its definition is shown in Equation 1.

$$DICC = \max_{\tau} \left| \frac{\int_0^T p_L(t) \cdot p_R(t+\tau)\, dt}{\sqrt{\int_0^T p_L^2(t)\, dt \cdot \int_0^T p_R^2(t)\, dt}} \right| \quad (1)$$

where $p_L(t)$ and $p_R(t)$ are the sound pressure signals at the left and right ears respectively, τ is the interaural time delay, T is the integration time, and t represents time.

Furthermore, Morimoto et al. [12] hypothesized a negative correlation between DICC and ASW and conducted experiments using musical sources. Note that the maximum value of DICC is 1.00.

2.2 Selection of Sound Sources

To enhance the accuracy of spatial perception evaluation using DICC, it is necessary to focus on the frequency bands that contribute most to sound-image localization and perceived width. Previous research has examined the frequency bands used for sound-image localization.

Nakabayashi [13] demonstrated in his study of horizontal sound image localization that high-frequency components above 8 kHz are essential for the accuracy and clarity of sound image localization. When these high-frequency components are lacking, the sound image tends to spread out and sound blurred.

Morikawa and Hirahara [14] conducted horizontal sound-localization experiments using real sound sources and analyzed the frequency bands required for sound localization. Their results showed that localization becomes significantly more difficult with sounds containing only components below 2 kHz, and also more difficult with sounds containing only components above 12 kHz.

Furthermore, Morikawa and Shimakura [15] conducted sound localization experiments using white noise with HPF and LPF applied. Their results indicated that the 2-4 kHz band is necessary for localization. These findings suggest that the frequency components essential for sound localization primarily lie within the 2-12 kHz band.

Conversely, Kasbach et al. [16] evaluated an IACC-based ASW perception model and demonstrated that including high-frequency components degrades ASW estimation performance. This result suggests ASW prediction functions optimally primarily in the low-frequency band below 2 kHz.

This assertion necessitates a measurement-based investigation into whether DICC variations should be interpreted as high- or low-frequency changes when evaluating loudspeakers with high-frequency attenuation characteristics such as BWL. Considering the BWL frequency response and results from previous preliminary experiments, the authors employ two approaches. The first approach involves adopting band-noise-containing frequencies above 2 kHz. These results align with prior research using sound sources spanning 2-12 kHz. The second approach involves presenting musical tracks directly. The authors selected tracks featuring a variety of acoustic instruments. Comparing measurements of sources containing both low and high frequencies with band noise measurements enables an investigation into the importance of the low-frequency range. The authors considered that prolonged exposure could adversely affect the health of future human subjects and therefore rejected band noise centered above 4 kHz. Therefore, this experiment does not include an analysis of this frequency band.

2.3 Binaural Correlation JND

Considerations regarding the measurement of physical indicators require a criterion for whether changes in physical quantities are aurally distinguishable to

humans. This criterion is known as the Just Noticeable Difference (JND). ISO 3382-1 [17] estimates the JND for IACC when evaluating spatial impression to be 0.075, as stated in Annex B. The authors use this value as the threshold for when DICC changes manifest as perceptible auditory differences, and they evaluate measurement results accordingly.

3 Methods of Measuring DICC

3.1 Preliminary Review

The authors focused on conducting measurements and listening tests in environments closer to actual usage conditions when evaluating loudspeakers. Typically, loudspeaker performance is measured in an anechoic chamber to assess its characteristics. However, since listening ease is judged by subjects, authors considered that in real sound fields, the influence of room modes would be significant. In preliminary experiments, subjects listened to pop music played through HS5 and BWL for comparison. The results suggested differences in response tendencies between the two loudspeaker types. However, it became clear that the authors need further consideration regarding the selection of audio sources and the response methods.

3.2 Measurement of DICC in a Listening Room with Reverberation

The measurements were conducted in a standard listening room rather than an anechoic chamber. It is because evaluating the diffusion properties of the BWL using DICC requires capturing room reflections, which are essential for this metric. Furthermore, since typical music listening environments are not free fields, the authors considered it important to conduct measurements in a realistic, non-anechoic environment.

Measurements were conducted in the reverberation-controlled room at Ohashi Campus, Kyushu University (Figure 5). The reverberation time was 0.532[s], and the background noise level was 20 dB (A-weighted). As shown in Figure 6, HS5 and BWL were placed at three positions: one facing forward along the long axis of the room, and one each to the left and right. The HATS was mounted on a swivel chair, fixed in position, allowing only azimuth changes. Each loudspeaker was positioned 1.5 m from the top of the HATS head at the measurement point (listening point). The volume of each loudspeaker was adjusted so that the noise level at both ear positions was 70 dB(A-weighted) when playing a 1 kHz band noise signal. Subsequently, the sound pressure level of all sound sources remained unchanged. The HATS was positioned to face each loudspeaker directly, and the authors recorded for 40 seconds to calculate the DICC for each loudspeaker individually. The sound sources used were octave-band noise (center frequencies 1 kHz and 2 kHz) and a music track (Feelin' The Same Way/Norah Jones). For the music track, sections centered on the chorus were used, with vocals and instruments present.



Fig. 5: Variable Reverberation Chamber

4 Results of DICC Calculation

The comparison results of DICC obtained using the recorded sound sources are shown in Figure 7. The vertical axis represents DICC, with a maximum value of 1. The horizontal axis represents the sound source; when octave-band noise was used, the center frequency is listed. The Center condition indicates recording facing the room's long wall. Under this condition, BWL showed significantly lower values than HS5 for all sound sources (1 kHz, 2 kHz, Song). Specifically, at 1 kHz, the DICC value of BWL was 0.36, and the DICC value of HS5 was 0.82, confirming high diffusivity even at the front position. Under the Left condition, where recording was performed facing the short side of the room, the influence of reflected sound is more pronounced. However, even under this condition, BWL remained lower than HS5 for all sound sources.

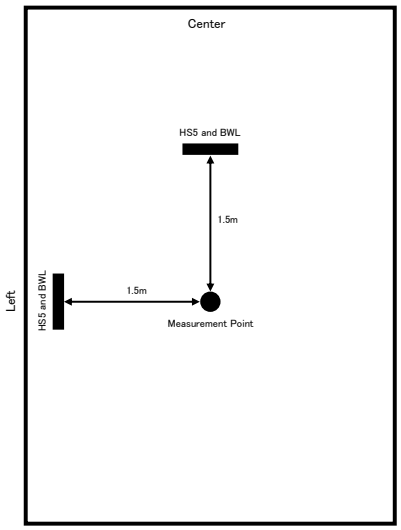


Fig. 6: Speaker Placement in the Experiment Room

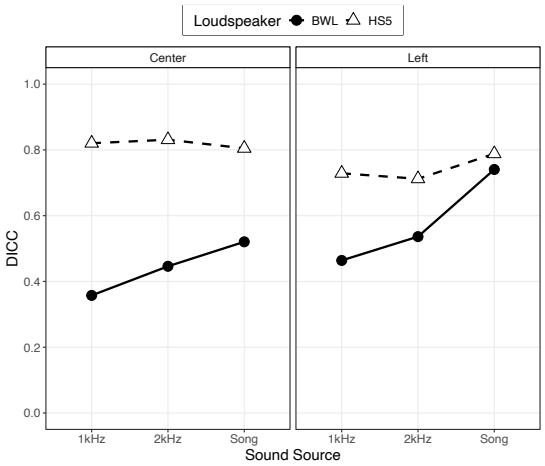


Fig. 7: Measured DICC

5 Discussion

5.1 Investigation of Auditory Perceptual Differences

The authors compared the minimum audible difference described in Background 1.5 with the measurement results. Figure 8 shows the difference between the DICC with BWL value and the DICC with HS5 value, compared to the minimum audible difference. Under the Center condition, all differences exceeded the JND, indicating a clear distinction. However, under the Left condition for the Song, the difference fell below the JND (0.075). This result suggests that when listening to complex music in a real reverberant environment, subjects may not perceive differences in loudspeaker diffusion characteristics.

5.2 Analysis Based on ASW and LEV Definitions

As defined in section 2.1, the authors analyzed the DICC by segmenting the impulse responses into early (0-80 ms) and late (>80 ms) reflection periods. This distinction, based on their respective contributions to ASW and LEV, allows us to evaluate whether the BWL diffusivity originates from initial reflections or the subsequent reverberant field. Figure 9 shows two types of DICC for loudspeaker placement, type, and sound source. Solid lines indicate DICC based on early reflections (Early_DICC), while dashed lines indicate DICC based on late reflections (Late_DICC). Examining the BWL values reveals that, for sound sources other than the 2 kHz octave-band noise (1 kHz, Song), the difference between Early_DICC and Late_DICC varies between the center and left conditions. The authors thought this occurred because the center condition has measurement points closer to the walls, allowing more reflections from the sides to reach them. For the HS5 song source (Song), a significant difference between Early_DICC and Late_DICC appeared under both the center and left conditions. The authors thought that the lower-frequency bands in the song source were related. Among the sound sources used this time, the song source was the only one containing frequencies below 800Hz.

There are two differences between BWL and HS5. The first is the difference between the Early_DICC and Late_DICC for all sound sources in the left condition, compared to the octave band noise (2 kHz) at the center frequency of 2 kHz in the center condition. Even when the difference correlated with HS5 exceeds the JND, the corresponding difference for BWL remains smaller. This result indicates that BWL does not produce significant differences in diffusion over time. The second difference is the progression of DICC due to source variations in the left condition. Even under conditions with strong lateral reflections, BWL exhibits similar progression in Early_DICC and Late_DICC, particularly for the song source, which includes low frequencies. The authors expected a difference between the

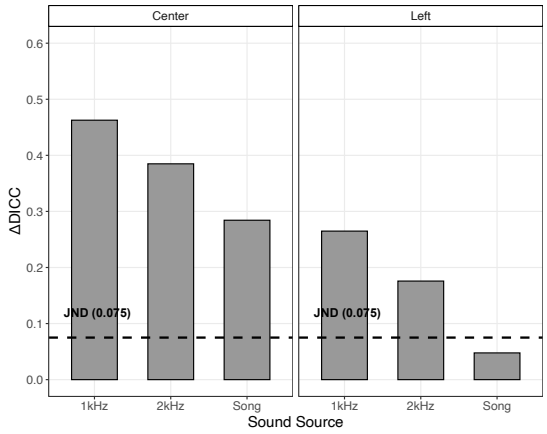


Fig. 8: Difference in DICC ($\Delta\text{DICC} = \text{DICC}_{\text{HS5}} - \text{DICC}_{\text{BWL}}$). The dashed line represents the JND of 0.075.

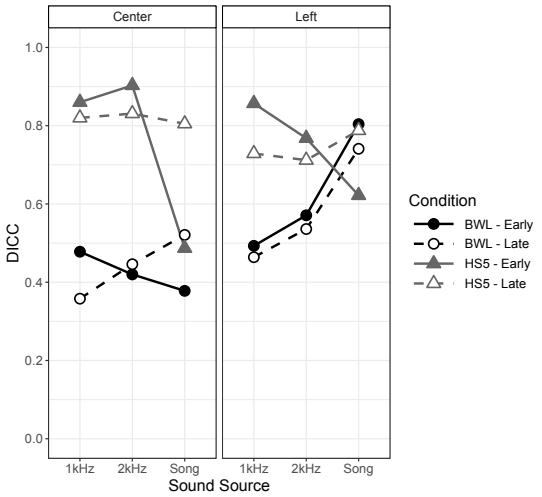


Fig. 9: Comparison of DICC based on early and late reflections.

two DICC values. However, the results did not match. This result suggests a relationship between BWL placement and its diffusion characteristics. The BWL lacks an enclosure and has an uncovered rear, allowing sound radiation toward the rear. In the left condition, where the loudspeaker is close to the wall, sound radiated toward the rear reflects and reaches the measurement point. This result is considered the reason Late_DICC did not decrease.

6 Conclusion

In this paper, the authors conducted a comparative evaluation using DICC to quantify the radiation characteristics of BWL physically. The experimental results yielded the following conclusions.

1. Compared to conventional loudspeakers, the BWL consistently demonstrated lower DICC values regardless of placement position or sound source type. The difference was particularly pronounced at the front position, confirming that the BWL exhibits high diffusivity from direct sound components.

2. The difference in DICC between the two loudspeakers exceeded the JND for spatial perception (0.075) under many conditions, suggesting that the sound field generated by BWL is clearly distinct from conventional systems in terms of auditory perception.

3. Analysis using time windows confirmed that BWL forms a diffuse sound field starting from the early reflection phase. In contrast, conventional loudspeakers exhibited behavior strongly dependent on the low-frequency components of the music and room reflections. Based on these results, the authors concluded that BWL enables stable spatial expression independent of room acoustics, as assessed from the perspectives of ASW and LEV.

In this study, differences in the diffusion characteristics between the HS5 and the BWL were observed across several musical stimuli. However, as a preliminary study, further investigations are required regarding both the physical properties of the BWL and the measurement conditions.

First, regarding the physical characteristics of BWL, a detailed analysis of its fundamental resonance frequency and panel mode shapes remains insufficient. Future research will investigate these vibration characteristics to clarify their direct relationship with the radiation characteristics. Additionally, the authors will examine how rear radiation affects the diffusive properties of BWL by testing the unit with a rear panel blocked.

Second, regarding measurement conditions, the authors plan to conduct measurements in other room environments with different acoustic properties to generalize

our findings. The authors also aim to analyze the impact of low-frequency components using high-pass-filtered music stimuli.

Finally, based on these objective findings, the authors plan to conduct subjective listening tests focusing on ASW using the sound sources employed in this study. The authors will analyze results alongside the JND and objective predictions, incorporating the Binaural Sound Pressure Level (BSPL).

References

[1] Kawahara, K., "Implementation examples of panel loudspeakers using bending waves of the diaphragm (in Japanese)," in *Proceedings of the Acoustical Society of Japan*, 2024.

[2] New Tranducers Limited, *nxt technologyreview*, 2002.

[3] Kawahara, K., "Distributed Mode Loudspeaker (DML Panel Speaker) (in Japanese)," *Journal. Acoust. Soc. Jpn.*, 62(11), pp. 798–801, 2006.

[4] Harris, N., Flanagan, S., and Hawksford, M. J., "Stereophonic localization in rooms, comparing conventional and distributed-mode loudspeakers," in *105th Audio Engineering Society Convention*, 4794, 1998.

[5] Kleczkowski, P., Makuch, T., Król-Nowak, A., and Czesak, K., "Equalized distributed mode loudspeakers rival conventional ones in listening tests," *The Journal of the Acoustical Society of America*, 157(4), pp. 2528–2541, 2025.

[6] Azima, H. and Mapp, P., "Diffuse field distributed-mode radiators and their associated early reflections," in *104th Audio Engineering Society Convention*, 4759, 1998.

[7] Heilemann, M. C., Anderson, D. A., Roessner, S., and Bocko, M. F., "The Evolution and Design of Flat-Panel Loudspeakers for Audio Reproduction," *Journal of the Audio Engineering Society*, 69(1/2), pp. 27–39, 2021.

[8] Inoue, S., Kawahara, K., Sato, T. G., Ishikawa, K., and Kamamoto, Y., "Immersive Applause Playback System Using Flat-Panel Loudspeakers," *Proceedings of 2024 IEEE 13th Global Conference on Consumer Electronics (GCCE)*, pp. 518–521, 2024.

[9] Hanyu, T., *Science of Concert Halls (Acoustic Science Series 6) (in Japanese)*, Corona Publishing Co., Ltd., 2012, section 2.1.

[10] Marshall, A. H., "A note on the importance of room cross-section in concert halls," *Journal of Sound and Vibration*, 5, pp. 100–112, 1967.

[11] Blauert, J. and Lindemann, W., "Spatial mapping of intracranial auditory events for various degrees of interaural coherence," *The Journal of the Acoustical Society of America*, 79(3), pp. 806–813, 1986.

[12] Morimoto, M. and Iida, K., "A practical evaluation method of auditory source width in concert halls," *Journal of the Acoustical Society of Japan (E)*, 16(2), pp. 59–69, 1995.

[13] Nakabayashi, K., "Sound localization in the horizontal plane (in Japanese)," *The Journal of the Acoustical Society of Japan*, 30(3), 1974.

[14] Morikawa, D. and Hirahara, T., "Sound localization in the horizontal plane with high-pass and low-pass noise (in Japanese)," *The Journal of the Acoustical Society of Japan*, 68(4), pp. 171–179, 2012.

[15] Morikawa, D. and Shimakura, N., "Frequency bands required for sound localization in the horizontal plane (in Japanese)," in *Proceedings of the Acoustical Society of Japan*, 2009.

[16] Kasbach, J., Hahmann, M., May, T., and Dau, T., "Assessing and modeling apparent source width perception," in *Proceedings of DAGA 2016*, pp. 1120–1123, 2016.

[17] International Organization for Standardization, "ISO 3382-1:2009, Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces," 2009.