forum acusticum 2026
8-12 September · Graz, Austria

PREDICTION OF LISTENING EFFORT IN MEETING ROOMS BASED ON OBJECTIVE METRICS

Leon Kaiser[1], Alois Sontacchi[1], Benjamin Pohler[1], Matthias Frank[1], Aron Karsai[1]

[1] *Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz, Austria*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

While traditional room acoustic metrics successfully evaluate speech intelligibility, recent research indicates that high intelligibility does not guarantee low listening effort. In modern meeting environments, signal-to-noise ratios are typically sufficient for high intelligibility due to short communication distances and adjustable playback levels. Consequently, this study reevaluates acoustic metrics for predicting listening effort under equal-loudness conditions to isolate perceptual effects beyond mere loudness. The mean opinion scores derived from a subjective listening test were regressed against established acoustic parameters using ordinary least squares. Leave-one-out cross-validation revealed that a linear combination of reverberation time and direct-to-reverberant energy ratio outperformed classic metrics such as clarity and the speech transmission index.

Keywords: *listening effort, speech intelligibility, room acoustics, predictive modeling*

1. Introduction

Room acoustic metrics are a widely researched topic usually in the form of evaluating speech intelligibility. First works date back to the 1930s, in which Aigner and Strutt proposed a ratio of useful and disturbing sound energy [1]. This was refined in 1964 by Lochner and Burger [2], who examined the beneficial aspects of early reflections for speech in auditoriums. In the 1970s, parameters were introduced that followed other approaches: Peutz proposed the Articulation Loss of Consonants (AL_{cons}) [3], while Houtgast and Steeneken developed the speech transmission index (STI) [4]. Returning to energy ratios, Bradley established the metrics Clarity (C_{50}) and Useful-to-Detrimental energy ratio (U_{50}) in his work from the 1980s to the 2000s as robust predictors of intelligibility [5], [6].

These metrics are typically designed for situations involving a single speaker facing a listening audience, where significant distances can lead to poor signal-to-noise ratios (SNRs). Energy ratios that include early and late reflections can also be interpreted as SNRs [7]. In modern meeting rooms, however, SNRs are often in reasonable ranges due to short communication distances or adjustable telecommunication levels. Under such conditions, it can be assumed that speech is intelligible. Nevertheless, current findings suggest that high intelligibility does not automatically guarantee low listening effort (LE) [8], particularly when influenced by inadequate acoustic treatment [9]. While recent studies have investigated LE in hybrid spaces through both qualitative user experiences [10] and objective acoustic metrics [11], the present work evaluates the accuracy of objective acoustic metrics as predictors of LE under equal-loudness conditions. To this end, a listening experiment was conducted using binaural renderings of twelve primary acoustic scenarios—varying in room volume, treatment status, and source–receiver distance. The acoustic scenarios were presented in four variants per scenario, incorporating both male and female speech at 0° and 90° source rotations. By normalizing stimulus loudness, we isolated perceptual effects independent of sound pressure level. The scenarios were captured using a first-order Ambisonics microphone and upmixed to a higher order via spatial decomposition [12]. We derived classical metrics (reverberation time (T_{60}), C_{50}, STI, direct-to-reverberant energy ratio (DRR)) supplemented by Short-Time Objective Intelligibility (STOI) and a novel spatial concentration (SC) parameter as experimental predictors of LE in room acoustics. These objective measures were regressed against mean opinion scores (MOS) using ordinary least squares (OLS). Finally, leave-one-out cross-validation was employed to assess the predictive accuracy of single metrics and their combinations.

***Corresponding author**: leon.kaiser@kug.ac.at.*

2. Methods

2.1. Acoustic scenarios

The acoustic scenarios utilized in the listening experiment encompass three rooms of varying volumes—small (S), medium (M), and large (L)—each evaluated in two distinct acoustic states: untreated and with supplemental acoustic treatment. Notably, the "untreated" condition for the medium room (a teaching studio) included existing permanent acoustic treatments that could not be removed for the study. For every room configuration, two source-receiver distances (near and far) were measured, resulting in a total of 12 distinct acoustic scenarios (3 rooms × 2 treatment states × 2 distances). The geometric layouts of the rooms, the tables, as well as the microphone and loudspeaker positions are schematically illustrated in Fig. 1. T_{60} values measured at the microphone positions are listed later in Tab. 1.

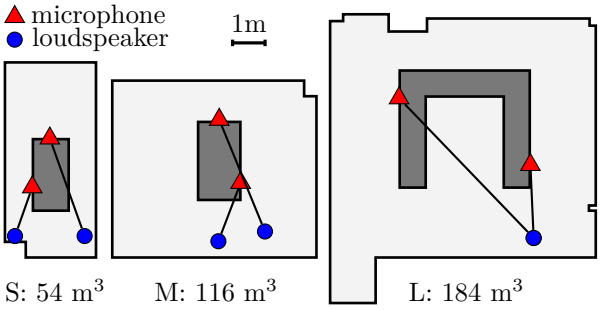

Figure 1: Floor plans of the investigated rooms showing tables and source–receiver configurations. Room volumes are specified below each respective plan.

Spatial impulse responses (IRs) were obtained using exponential sine sweeps. They were captured with a Soundfield ST450 first-order Ambisonics microphone. Sound reproduction was handled by a Genelec 8020B for the small and medium rooms, while a Yamaha DXR8 was employed for the large room, since it was already installed there. Inter-speaker differences are expected to be minimal; moreover, using directional loudspeakers yields more realistic DRRs than omnidirectional sources. Microphones and loudspeakers were installed at a typical sitting height of 1.22 m. To achieve high spatial resolution, the first-order Ambisonics IRs were spatially decomposed and upmixed to fifth-order Ambisonics using the Ṡ/D-ASDM algorithm proposed by Gölles and Frank [12]. The algorithm inherently identifies the direction of arrival of the direct sound component, used later for the orientation of the stimuli. The synthesized scenes were created by convolving fifth-order Ambisonics IRs with male and female English speech samples from the EBU SQAM library [13]. Samples no. 49 and 50 were cut to eight seconds. Each sample was presented in two spatial orientations: a frontal arrival of the direct sound and a lateral arrival from the left. Scene rotation and binaural rendering were performed using the IEM Plugin Suite [14], [15]. In summary, the 12 primary acoustic scenarios (3 rooms × 2 distances × 2 treatment states) were expanded into 48 stimuli by incorporating four variants: two speaker genders (male and female) and two spatial orientations (0° and 90° rotation). To ensure constant loudness across all stimuli, levels were normalized to the same integrated loudness units full scale (LUFS) [16] using REAPER digital audio workstation. Additional stimuli, such as artificial noise and other adaptations, were included in the listening test to gain deeper insights. However, these are not within the scope of this publication. The generated speech stimuli for all scenarios are available via open access on Zenodo: `https://doi.org/10.5281/zenodo.21510312`.

2.2. Listening Experiment Setup

Subjective quality ratings were obtained via a web-based listening test following the Absolute Category Rating (ACR) paradigm [17]. The experiment was implemented using the open-source webMUSHRA framework[1] and hosted online, allowing for remote participation via a distributed web link.

A total of 18 participants (17 male, 1 female) completed the test. The cohort consisted primarily of experienced listeners, with 14 identifying as experts. Regarding playback hardware, 16 participants used over-ear headphones, while the remaining two utilized on-ear models. At the start of the session, participants were instructed to imagine themselves attending a meeting and to rate the stimuli based on their suitability for such a scenario. The instructions explicitly noted that LE could be influenced by factors such as reverberation, speech clarity, timbre, echoes, and background noise. Before stimuli were presented, a 1 kHz sine with modulated frequency was played back and the users were asked to adjust the playback level, so that the test tone was at the threshold of perception. The LUFS-adjusted stimuli were between 54 and 55.5 $\mathrm{dB_{RMS}}$ above this test tone. After the level adjustment, a learning phase followed consisting of three stimuli that established the range of low and high LE. Finally, test stimuli were presented in a randomized sequence and evaluated on the five-point ACR scale, ranging from "Very Poor" (1) to "Excellent" (5). For each IR, a Mean Opinion Score (MOS) was calculated. Additionally, sub-scores for frontal ($\mathrm{MOS_f}$) and lateral ($\mathrm{MOS_l}$) arrivals were derived.

2.3. Room Acoustic Metrics

2.3.1. Classical Room Acoustic Metrics

All classical Room Acoustic Metrics were calculated from the omnidirectional component of the Ambisonic IRs. Single-value representations of T_{60} and C_{50} were derived from the arithmetic mean of the 500 Hz and 1 kHz octave bands according to [18]. For T_{60} estimation, the linear regression range of the energy decay curve was adjusted to mitigate direct-sound interference: a range of −5 to −25 dB was used for far source-receiver distances, while −10 to −30 dB was

[1] open-source webMUSHRA at audiolabs Erlangen

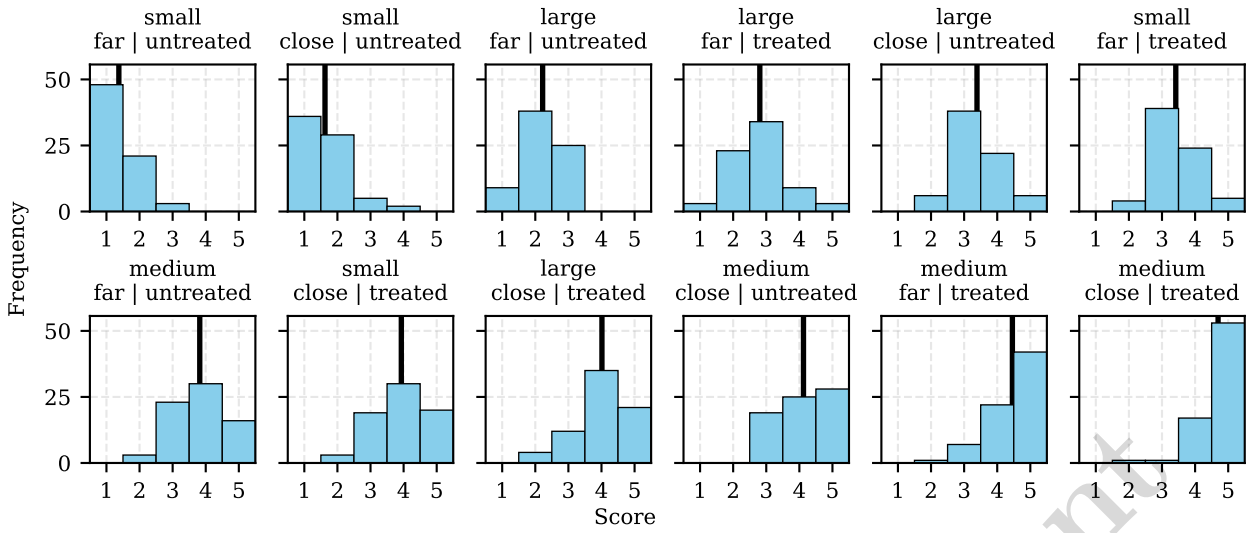

Figure 2: Distribution of listening test ratings with Mean Opinion Scores (MOS) underlaid in the background. Subplot titles denote room size, while subtitles indicate source–receiver distance and acoustic treatment status. The numbers 1-5 are the ACR-ratings: 1 - very poor, 2 - poor, 3 - fair, 4 - good, 5 - excellent.

applied for near-field positions. The STI value was calculated using the IR values in accordance with IEC 60268-16 [19] using the indirect method. The DRR was estimated via two methods: (i) by calculating the energy ratio using a temporal window 1 ms after the direct sound arrival (DRR_1), and (ii) via a diffuse-field approximation based on the source-receiver distance (DRR_d). Using the theoretical sound energy for a diffuse sound field [20] the DRR can be expressed for a point source:

$$DRR_d = 10 \log_{10} \left(\frac{Q/4\pi d^2}{cT_{60}/13.8V} \right) \quad (1)$$

where Q denotes the directivity factor, d the source–receiver distance, c the speed of sound, and V the room volume. For the calculation, a directivity factor of $Q = 2$ and a speed of sound of $c = 343\,\mathrm{m/s}$ were assumed. Both versions of the DRR utilized in this study are considered approximations. For DRR_1, it must be noted that early reflections arriving within the first millisecond are included in the direct sound measurement. On the other hand, since physical loudspeakers do not produce ideal impulses, the direct sound energy does not fully decay within the 1 ms window. Regarding DRR_d, the metric assumes a perfectly diffuse sound field. Moreover, because it was calculated using the single-value representation of T_{60}, it fails to account for frequency-dependent decay—specifically the faster attenuation of high frequencies—making it a simplified perceptual approximation.

2.3.2. Additional Metrics

Beyond standard acoustic parameters, the STOI measure was calculated [21]. While STOI is conventionally employed to evaluate noise reduction algorithms for non-reverberant signals, it was included here as an exploratory metric to observe its behavior in a room acoustic context. However, it must be noted that STOI evaluates the transformation between an input and output signal, making the results inherently dependent on the specific microphone and loudspeaker IRs. The reference signal for the STOI comprised a sequence of speech samples (39–54) from the EBU SQAM library [16], while the processed signal was derived via convolution with the omnidirectional component of the Ambisonic IR.

In addition, we calculated an experimental metric that we refer to as spatial concentration (SC). The experimental SC metric represents the ratio of energy within a 20° angular sector (defined by a raised-cosine taper) around the direct sound arrival within the first 50 ms, averaged across octave bands from 125 Hz to 1 kHz.

2.4. Prediction of Mean Opinion Score

To evaluate the predictive performance of the metrics and their combinations, a linear regression model was employed. Each of the K acoustic metrics M_k was standardized by its respective mean μ_k and standard deviation σ_k:

$$p_k = \frac{M_k - \mu_k}{\sigma_k} \quad (2)$$

Respective values for μ_k and σ_k are in Tab. 1. For a combination of N standardized acoustic metrics p_n, the predicted MOS score $\mathrm{MOS}_\mathrm{pred}$ is defined as:

$$\mathrm{MOS}_\mathrm{pred} = \beta_0 + \sum_{n=1}^{N} \beta_n \, p_n \quad (3)$$

where β_n represents the regression coefficients. These parameters were determined via OLS fitting against the observed MOS values from the listening experi-

Table 1: Acoustic characteristics of the acoustic scenario, sorted by MOS. Units are provided for distance d (m), T_{60} (s), C_{50} (dB), DRR_1 (dB) and DRR_d (dB); remaining values are dimensionless. The first three columns specify the acoustic scenarios (S is small, M is medium and L is large room). Column four to six show the MOS values from the listening test (MOS_f for frontal stimuli, MOS_l for lateral and MOS for the combined data). Column seven to thirteen show the acoustic metrics calculated from the IR, described in section 2.3.

Room	Treated	d	MOS	MOS_f	MOS_l	T_{60}	C_{50}	STI	DRR_1	DRR_d	$STOI$	SC
S	No	3.19	1.4	1.4	1.3	0.95	-1.4	0.59	-4.0	-14.5	0.59	0.14
S	No	1.59	1.6	1.6	1.6	0.96	1.3	0.65	2.8	-8.5	0.72	0.23
L	No	5.96	2.2	2.1	2.3	0.66	6.7	0.72	-6.1	-13.0	0.66	0.16
L	Yes	5.96	2.8	2.8	2.9	0.48	6.3	0.75	-4.6	-11.6	0.68	0.18
L	No	2.25	3.4	3.3	3.5	0.60	8.1	0.81	5.9	-4.1	0.85	0.4
S	Yes	3.19	3.4	3.3	3.6	0.45	5.5	0.79	1.5	-11.3	0.74	0.19
M	No	3.72	3.8	3.8	3.8	0.40	9.4	0.79	-1.0	-8.7	0.79	0.20
S	Yes	1.59	3.9	3.8	4.0	0.42	10.9	0.87	6.7	-4.9	0.86	0.30
L	Yes	2.25	4.0	4.1	3.9	0.50	10.7	0.85	7.1	-3.3	0.87	0.39
M	No	1.93	4.1	4.0	4.3	0.37	8.8	0.81	2.3	-2.7	0.86	0.32
M	Yes	3.72	4.5	4.5	4.4	0.28	14.0	0.88	1.9	-7.2	0.84	0.21
M	Yes	1.93	4.7	4.8	4.6	0.28	15.9	0.91	4.4	-1.4	0.90	0.34
metric mean (μ_k)						0.53	8.02	0.79	1.41	-7.60	0.78	0.26
metric standard deviation (σ_k)						0.22	4.66	0.09	4.28	4.17	0.10	0.09

ment. The unobserved error is defined as:

$$\epsilon = MOS_{pred} - MOS \tag{4}$$

The data pool is small with only 12 MOS values for the according $N_{IR} = 12$ IRs. To avoid the possibility of overfitting a leave-one-out cross-validation was performed. For various metric combinations a RMSE σ is calculated according to equation (5).

$$\sigma = \sqrt{\frac{1}{N_{IR}} \sum_{i=0}^{N_{IR}} \epsilon_i^2} \tag{5}$$

To identify significant differences between parameter configurations, independent Student's t-tests were performed on the distributions of the absolute errors $|\epsilon_i|$. Additionally, 95 % confidence intervals (CI) for the RMSE were derived via bootstrapping ($B = 10000$) to assess the reliability of the error metrics.

3. Results

The rating distributions from the listening test are illustrated in Fig. 2. Each IR received 72 total ratings, collected from 18 participants across four conditions: varying arrival direction (frontal vs. lateral) and speaker gender (male vs. female). The additional frontal and lateral sub-scores MOS_f and MOS_l each comprise 36 ratings per IR. The exact MOS, MOS_f and MOS_l values are presented alongside the respective objective metrics in Tab. 1. To analyze the relationships between the objective metrics, hierarchical cluster analysis was performed. The metrics were used in their z-score standardized version p_k. We utilized an absolute Pearson correlation distance metric ($1 - |r|$) to define dissimilarity, followed by the Average Linkage (UPGMA) method for cluster agglomeration. The resulting dendrogram (Fig. 3) illustrates the hierarchical organization of parameters, with a Cophenetic Correlation Coefficient of 0.72, indicating an acceptable preservation of the original dissimilarity matrix. The hierarchical clustering analysis reveals two distinct groups of metrics. Cluster A comprises the classical room acoustic parameters STI, C_{50}, and T_{60}, while Cluster B consists of STOI, DRR_d, SC, and DRR_1. Cluster B metrics are categorized as "DRR-related" metrics, as they exhibit a stronger dependency on source–receiver distance.

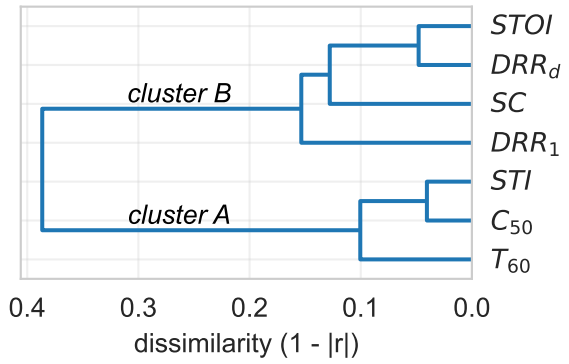

Figure 3: Dendrogram (UPGMA linkage) showing the relationships between standardized metrics based on absolute Pearson correlation distance ($1 - |r|$). The Cophenetic Correlation Coefficient of 0.72 validates the structural preservation of the original distances.

The regression analysis focuses on inter-cluster combinations. By selecting one metric from Cluster A and one from Cluster B, we maximize the inclusion of independent acoustic information. Preliminary testing indicated that intra-cluster pairings were functionally equivalent to single-metric models, yielding negligible performance gains. In contrast, cross-cluster com-

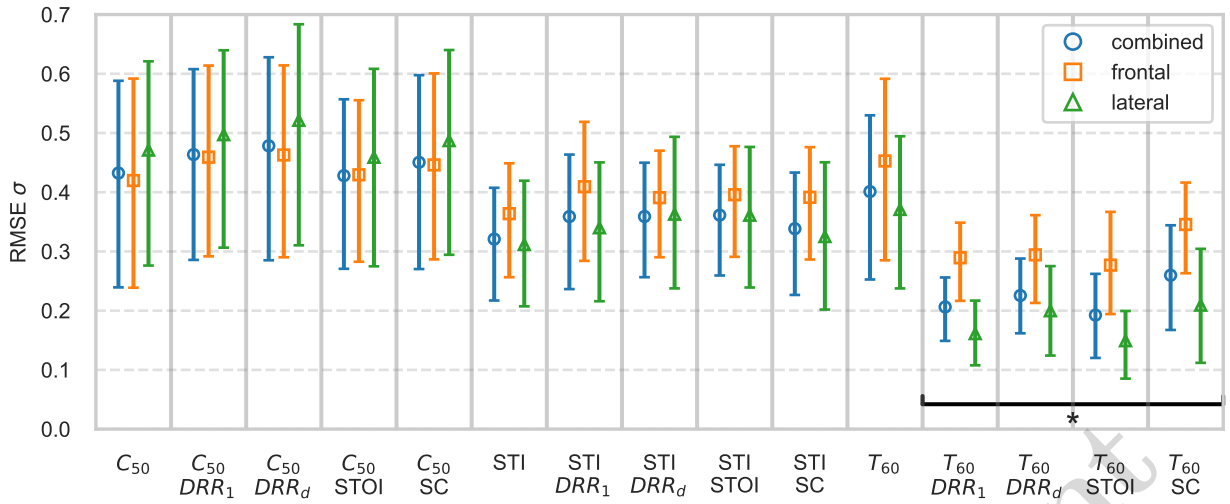

Figure 4: RMSE for single predictor metrics and cross-cluster combinations. Error bars represent 95% CI. The "combined" label refers to regression against MOS, while "frontal" and "lateral" denote regressions against MOS_f and MOS_l, respectively. The "*" symbol refers to the combinations of T_{60} and cluster B metrics

binations demonstrated significant improvements in specific configurations. Fig. 4 depicts the RMSE values with associated 95% CIs for regressions against the global MOS (combined), MOS_f (frontal), and MOS_l (lateral). The results indicate that the RMSE consistently decreases only when combining T_{60} with a metric from Cluster B (see * in Fig. 4). Shapiro-Wilk tests on the distributions of unobserverd errors showed no significant deviations from normality across the 45 distributions ($p > 0.05$), with a single marginal exception ($p = 0.0425$) expected as a false positive due to multiple testing, justifying the use of parametric t-tests. Given the exploratory nature of this analysis, Bonferroni-Holm corrections were omitted to avoid inflating Type II errors, meaning readers should consider the increased false-positive risk when evaluating the reported uncorrected p-values. Student's t-tests revealed that for the combined configurations ($p > 0.05$) and frontal arrivals ($p > 0.29$), the decrease in RMSE was not significant compared to the single-parameter predictors (C_{50}, STI, and T_{60}). In contrast, for lateral arrivals, the combination of T_{60} and Cluster B metrics yielded significant improvements ($p < 0.05$) compared to the single metric predictors in most configurations. Notable exceptions include the combination of $T_{60}+DRR_d$, which was not significant compared to the STI ($p = 0.12$), and $T_{60}+SC$, which showed marginal significance against T_{60} ($p = 0.051$) but was non-significant against the STI ($p = 0.12$). Ultimately, T_{60}+DRR_1 and T_{60}+$STOI$ demonstrated robust performance, consistently outperforming all single-parameter predictors across the lateral configuration with significance.

Furthermore, with the exception of T_{60} + DRR_d ($p = 0.15$), all T_{60} + Cluster B combinations showed significant differences for frontal and lateral arrival prediction ($p < 0.03$). The regression coefficients for the T_{60} + Cluster B models, obtained via OLS fitting across all 12 IRs, are shown in Tab. 2.

4. Conclusion

This study investigated the prediction of LE in meeting rooms using objective acoustic metrics. MOS were collected via a listening experiment encompassing three meeting rooms of varying volumes, each evaluated across two source-receiver distances, acoustic treatment states, and listener orientations. By applying OLS, we evaluated the performance of these metrics as objective predictors of LE. For lateral arrivals under equal-loudness conditions, combining T_{60} with some of the DRR-related metrics significantly improved the prediction of perceived LE compared to using STI, C_{50}, or T_{60}. Quantitatively, the regression weights suggest an equivalence where a 0.1 s change in reverberation time impacts listening effort similarly to a 4 dB shift in DRR. While these combinations also reduced the RMSE for frontal arrivals, the improvements did not reach statistical significance. Further research with a larger dataset is required to determine whether these synergistic effects of T_{60} and DRR can be statistically confirmed for frontal arrivals. In addition, the limits of assuming a linear relationship are of interest. Overall, these results suggest that multi-dimensional acoustic analysis may be more effective for designing communication spaces than relying on single parameters alone. Integrating metrics that target distinct physical drivers also enables a prediction of how varying individual attributes—such as the extent of acoustic treatment, source–receiver distance or directivity of the source—influence the perceived LE. Future work should also account for the SNR and other potential influencing factors.

5. Acknowledgments

Financial support by the Austrian Research Promotion Agency (FFG) (Grant No. 66369593 - Room

Table 2: Regression coefficients β (see equation 3) for different metric combinations, obtained via OLS fitting across all 12 IRs. Columns denote OLS fits to the corresponding combined, frontal, or lateral MOS values, while rows indicate the evaluated metric combinations (T_{60} + cluster B metric).

Metric Combination	combined			frontal			lateral		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
$M_1: T_{60}; M_2: DRR_1$	3.32	-0.88	0.33	3.29	-0.86	0.35	3.35	-0.90	0.32
$M_1: T_{60}; M_2: DRR_d$	3.32	-0.78	0.37	3.29	-0.75	0.40	3.35	-0.81	0.34
$M_1: T_{60}; M_2: STOI$	3.32	-0.65	0.47	3.29	-0.62	0.50	3.35	-0.68	0.44
$M_1: T_{60}; M_2: SC$	3.32	-0.88	0.31	3.29	-0.86	0.33	3.35	-0.90	0.30

Imagineer) and the cooperation with BETTERMEET-INGROOMS GmbH are gratefully acknowledged.

References

[1] F. Aigner and M. Strutt, "On a physiological effect of several sources of sound on the ear and its consequences in architectural acoustics," *The Journal of the Acoustical Society of America*, vol. 6, no. 3, pp. 155–159, 1935.

[2] J. Lochner and J. Burger, "The influence of reflections on auditorium acoustics," *Journal of Sound and Vibration*, vol. 1, no. 4, pp. 426–454, 1964.

[3] V. Peutz, "Articulation loss of consonants as a criterion for speech transmission in a room," *J. Audio Eng. Soc.*, vol. 19, pp. 915–919, 1971.

[4] T. Houtgast and H. Steeneken, "Evaluation of speech transmission channels by using artificial signals," *Acustica*, vol. 25, no. 6, pp. 355–+, 1971.

[5] J. S. Bradley, "Speech intelligibility studies in classrooms," *The Journal of the Acoustical Society of America*, vol. 80, no. 3, pp. 846–854, 1986.

[6] J. Bradley and S. R. Bistafa, "Relating speech intelligibility to useful-to-detrimental sound ratios (l)," *The Journal of the Acoustical Society of America*, vol. 112, no. 1, pp. 27–29, 2002.

[7] J. S. Bradley, H. Sato, and M. Picard, "On the importance of early reflections for speech in rooms," *The Journal of the Acoustical Society of America*, vol. 113, no. 6, pp. 3233–3244, 2003.

[8] M. B. Winn and K. H. Teece, "Listening effort is not the same as speech intelligibility score," *Trends in Hearing*, vol. 25, p. 23 312 165 211 027 688, 2021.

[9] C. Visentin, N. Prodi, F. Cappelletti, S. Torresin, and A. Gasparella, "Using listening effort assessment in the acoustical design of rooms for speech," *Building and Environment*, vol. 136, pp. 38–53, 2018.

[10] E. Marsja, C. Signoret, and V. Stenbäck, "The digital workplace and meeting accessibility: A qualitative study on listening effort in video meetings for employees with hearing loss," *WORK*, p. 10 519 815 251 398 498, 2025.

[11] R. Einig, S. Janscha, J. Schuster, J. Koch, M. Hagmueller, and B. Schuppler, "Room acoustics affect communicative success in hybrid meeting spaces: A pilot study," *arXiv preprint arXiv:2509.11709*, 2025.

[12] L. Gölles and M. Frank, "Ambisonic spatial decomposition method with salient/diffuse separation," in *Proceedings of the 158th AES Convention, Warsaw, Poland*, 2025.

[13] EBU. "Sqam cd: Sound quality assessment material. "[Online]. Available: `https://tech.ebu.ch/publications/sqamcd`.

[14] Institute of Electronic Music and Acoustics, *IEM Plug-in Suite*, Version 1.14.0, University of Music and Performing Arts Graz, 2024. [Online]. Available: `https://plugins.iem.at/`.

[15] M. Zaunschirm, C. Schörkhuber, and R. Höldrich, "Binaural rendering of ambisonic signals by head-related impulse response time alignment and a diffuseness constraint," *The Journal of the Acoustical Society of America*, vol. 143, no. 6, pp. 3616–3627, 2018.

[16] *Loudness normalisation and permitted maximum level of audio signals*, EBU, Tech. Rec. R128, 2014. [Online]. Available: `https://tech.ebu.ch/publications/r128`.

[17] *Subjective evaluation of speech quality with a crowdsourcing approach*, ITU-T, Recommendation P.808, 2021. [Online]. Available: `https://www.itu.int/rec/T-REC-P.808`.

[18] *Acoustics – measurement of room acoustic parameters*, ISO, 3382-1, 2009.

[19] *Objective rating of speech intelligibility by speech transmission index*, IEC Standard 60268-16:2020, 2020.

[20] S. Weinzierl, *Handbuch der Audiotechnik*. Springer Science & Business Media, 2008.

[21] C. H. Taal, R. C. Hendriks, R. Heusdens, and J. Jensen, "A short-time objective intelligibility measure for time-frequency weighted noisy speech," in *2010 IEEE international conference on acoustics, speech and signal processing*, IEEE, 2010, pp. 4214–4217.