The scores of speech intelligibility, obtained using objective and subjective methods for three university lecture rooms of the small, medium, and large sizes with different degrees of filling, were presented. The problem of achieving high speech intelligibility is relevant for both students and university administration, and for architects designing or reconstructing lecture rooms. Speech intelligibility was assessed using binaural room impulse responses which applied an artificial head and non-professional quality audio equipment for measuring. The Speech Transmission Index was an objective measure of speech intelligibility, while the subjective evaluation of speech intelligibility was carried out using the articulation method.

Comparative analysis of the effectiveness of parameters of impulse response as a measure of speech intelligibility showed that Early Decay Time exceeded the score of the T30 reverberation time but was ineffective in a small lecture room. The C50 clarity index for all the considered lecture rooms was the most informative. Several patterns determined by the influence of early sound reflections on speech intelligibility were detected. Specifically, it was shown that an increase in the ratio of the energy of early reflections to the energy of direct sound leads to a decrease in speech intelligibility. The exceptions are small, up to 30–40 cm, distances from the back wall of the room, where speech intelligibility is usually slightly higher than in the middle of the room. At a distance of 0.7–1.7 m from the side walls of the room, speech intelligibility is usually worse for the ear, which is closer to the wall. The usefulness of the obtained results lies in refining the quantitative characteristics of the influence of early reflections of sound on speech intelligibility at different points of lecture rooms

Keywords: binaural room impulse response, speech intelligibility, early reflections of sound

UDC 534.843.742
DOI: 10.15587/1729-4061.2021.228405

ASSESSMENT OF SPEECH INTELLIGIBILITY IN UNIVERSITY LECTURE ROOMS OF DIFFERENT SIZES USING OBJECTIVE AND SUBJECTIVE METHODS

Arkadiy Prodeus
Corresponding author
Doctor of Technical Sciences, Professor
Department of Acoustic and
Multimedia Electronic Systems*
E-mail: aprodeus@gmail.com
Maryna Didkovska
PhD, Associate Professor
Department of Mathematical
Methods of System Analysis*
*National Technical University of Ukraine "Igor
Sikorsky Kyiv Polytechnic Institute"
Peremohy ave., 37, Kyiv, Ukraine, 03056

Received date: 08.04.2021
Accepted date: 25.05.2021
Published date: 25 .06.2021

***How to Cite**: Prodeus, A., Didkovska, M. (2021). Assessment of speech intelligibility in university lecture rooms of different sizes using objective and subjective methods. Eastern-European Journal of Enterprise Technologies, 3 (5 (111)), 47–56. doi: https://doi.org/10.15587/1729-4061.2021.228405*

1. Introduction

One of the conditions for getting high-quality education is good intelligibility of the teacher's speech in the places where students are sitting. Meanwhile, the level of a speech signal decreases at an increase in distance between a listener and a teacher, which is especially noticeable in large lecture rooms. Since there is always noise in rooms (movement and conversations of nearby listeners, noise of a fan or an air conditioner, etc.), a signal level may not be high enough to ensure good speech intelligibility. Listeners with partial loss of hearing, the ones with hearing aids and cochlear implants are especially susceptible to harmful effects of noise. Even the category of listeners with normal hearing, such as pupils of elementary school, is more prone to harmful effects of noise on speech intelligibility than middle and high school students.

The masking effect of noise is partly compensated by the sound reflection from different surfaces (ceiling, walls, floor, furniture, etc.) so that the signal level in the rooms turns out to be higher than in the open space. However, late reflections of sound, as well as background noise, are harmful because they mask a speech signal. As for early reflections, their role is twofold. On the one hand, early sound reflections increase the level of a signal, improving speech intelligibility. On the other hand, early reflections impair the naturalness of sound, which leads to a slight decrease in speech intelligibility. Ill-conceived architectural solutions (unusual shape of a room, concave surfaces) can disrupt the diffuse nature of the sound field, leading to increased noise and even the appearance of an extremely undesirable echo.

The problem of low speech intelligibility can be solved by correcting the acoustics of the room by passive (noise-absorbing materials) or active (hardware-software systems for correction of the frequency characteristics of a room) methods. However, such correction should be preceded by an analysis of the acoustic condition of a lecture room, involving the assessment of noise characteristics and parameters of a room impulse responses. That is why the problem of establishing the comparative usefulness of the parameters of the room impulse response, containing information about speech intelligibility, is relevant and requires appropriate experimental and analytical study. Equally relevant are the issues of the influence of the lecture room volume on the informative parameters of the room impulse response, as well as the influence of reflecting surfaces of a lecture room on speech intelligibility at different points of a room.

2. Literature review and problem statement

The acoustics of school premises was quite well studied. Examples of poor architectural solutions that reduce speech intelligibility due to increased noise and reverb time, echo, or flutter echo are shown in paper [1]. Experimental studies have shown that noise in classrooms is more noticeable than late reflections of sound [2]. Acceptable speech intelligibility in lecture rooms is achieved due to early reflections of sound, even if the speaker's head is directed in the opposite direction from a listener [3]. The problem of ensuring good speech intelligibility is particularly relevant for schoolchildren with hearing impairments [4] and elementary school pupils [5], for whom noise levels should not exceed 30–35 dBA. Despite the considerable amount of research carried out in the field of the room acoustics, it is concluded in article [6] that there are a number of issues that require further development. Indeed, the results contained in [7] did not support the conclusion of paper [5] on the equivalence of the action of early reflections and direct sound. As the development of the results presented in papers [5] and [7], it was concluded in article [8] that the process of integrating early reflections with direct sound occurs regardless of the spatial processing of sound by the human hearing system.

Several national and international standards and recommendations prove a fairly good level of studying the acoustics of school premises. For example, in the standard DIN 18041:2004-05 (Germany), the premises are divided into categories A and B [9], the optimal values of reverb time are specified for category A premises ("music", "speech", "sport", "communication") [10]. Requirements and recommendations for acoustic characteristics and noise insulation of school premises are contained in the ANSI/ASA S12.60/ Part 1 (USA) standard [11]. Direct and indirect ways of assessing the STI speech intelligibility index and its modifications (STIPA, STITEL, RASTI) are described in the British standard BS EN 60268-16:2011 [12] and the international standard IEC 60268-16 [13]. Recommendations on taking into consideration the needs of ordinary schoolchildren and students with hearing impairments are outlined in the collection "Acoustic design of schools: performance standards. Building Bulletin 93" (UK) [14]. Similar recommendations, supported by experimental research, are outlined in the collection "The Essex Study. Optimized classroom acoustics for all" (UK) [15].

The acoustics of the concert halls is also well studied. Lists of the most important acoustic characteristics of concert halls, as well as the algorithms for their evaluation, are presented in the international standard ISO 3382-1 [16]. This demonstrates the relative completeness of scientific research in this area, despite the existence of a series of issues that need to be clarified [6].

As regards the acoustics of university lecture rooms, there are no explicitly relevant practical recommendations and standards, which indicates the lack of attention of scientists to the acoustics of the premises of this category. A possible reason for this situation may be the "intermediate" position of university lecture rooms.

Indeed, the average size of university lecture rooms is larger than that for school classes but less than the average size of concert halls. The variety of shapes and decoration of university lecture rooms is greater than that of school classes but less than that of concert halls. For example, in [17], the middle-sized lecture rooms include a room with a balcony, designed for 1,100 listeners and shaped like a "fan". The volume of twelve small and medium-sized university lecture rooms, designed for 30–70 students, considered in paper [18], ranges from 188 m^3 to 343 m^3. The acoustic characteristics of the rectangular lecture rooms having the volume of 530 m^3 and 1740 m^3 were explored in article [19]. An attempt to classify university lecture rooms by size was made in paper [20]. In this case, three classes of premises of small, medium-sized, and large size with the volume of less than 350 m^3, 350–650 m^3, and more than 650 m^3, respectively, were separated. A similar classification of premises was proposed in [21] where small, medium-sized and large lecture rooms included the premises of less than 230 m^3, 230–350 m^3, and more than 350 m^3, respectively. A typical feature of large lecture rooms, as well as concert halls, is the sloping floor [19, 20].

At the same time, university lecture rooms are smaller than concert halls. For example, the volume of ten halls in Boston (USA) intended for speech and music performances, considered in paper [22], varies from 3,000 to 60,000 m^3. Decorating materials in university lecture rooms [23] are more diverse than those in school classrooms [14] but are less diverse than in concert halls [22].

Given the lack of research in the acoustics of university lecture rooms, the increased interest of scientists in this problem seems understandable, which contributed to the development of useful recommendations for practical application. Thus, instead of the cumbersome and expensive hardware and software system of articulation tests described in paper [3] and developed in paper [7], the stimulus signals began to be synthesized more often by convolution of a clear signal and binaural room impulse response (BRIR) [17]. This approach has been greatly facilitated by the development of BRIR bank-building technology [24], as well as using special mathematical methods [25] when assessing the impulse responses of the premises [26]. The reliability of the evaluation of speech intelligibility masked by noise and reverb was significantly enhanced using an objective measure of speech intelligibility in the form of the STI index [27]. Automation of articulation tests [28] enabled a significant reduction in their resource intensity. Analysis of the features of the hearing system, presented in the form of a kind of spectral analysis system [29], made it possible to clarify the model of the middle ear [30]. In the development of measures aimed at correcting the acoustic characteristics of the room, the results of analytical [31] and experimental [32] studies of the acoustic field near reflecting surfaces are very useful.

To date, however, no scientifically grounded system of requirements for acoustic characteristics of university lecture rooms was created. Most of the relevant publications are fragmental in nature. For example, studies [19, 33] report the scores of the C50 clearness index and the STI speech intelligibility index for medium- and large-sized lecture rooms. At the same time, small-sized lecture rooms remain outside the field of interest of the authors, and instead of assessing the speech intelligibility using the objective articulation method, they conducted surveys of listeners, which significantly reduces the reliability of the results. Research [18] does not have these shortcomings, as speech intelligibility in university lecture rooms of small and medium size was assessed by both the articulation method and the instrumental method. However, large audiences were missing in the study, and EDT, T30, and C50 were evaluated instead of the STI index.

Insufficient attention was paid to quantifying the impact of reflecting surfaces on speech intelligibility. The known results indicate the lack of systemic search in this direction. For example, the results of measurements of speech intelligibility in the center and at the back wall of large rooms, cited in papers [22, 34], indicate a possible increase in speech intelligibility at the back wall. However, the authors of these works did not pay much attention to this fact. This drawback was corrected in research [35, 36] and it was pointed out that it is possible to increase the STI up to 7–14 percent at the back wall, compared to the STI values in the center of the room. However, no relation between the scores of speech intelligibility and the volume of a lecture room was detected, the links between the STI and BRIR parameters were not analyzed either. This shortcoming is partly compensated in [37] when considering two models of early sound reflections. The results of the analysis of these models are consistent with the conclusions in research [38] on the dual role of early reflections that increase the signal level but distort its spectrum. However, the models considered in paper [37] need to be further refined, as they do not take into account an increase in the density of early reflections over time.

Thus, the dependence of the effectiveness of objective measures of speech intelligibility on the volume of a lecture room has not been investigated to date. In addition, there are no quantitative estimates of the degree of influence of the reflecting surfaces of a lecture room (walls, windows, furniture, etc) on speech intelligibility at different points in a room. The lack of such data prevents the development of evidence-based practical recommendations and requirements for the acoustics of university premises. Such recommendations are needed both by the university administration and by architects designing or renovating university lecture rooms.

The above allows us to argue that it is appropriate to establish the informativeness of instrumental estimates of speech intelligibility in university lecture rooms of different sizes, as well as to identify the degree of influence of reflecting surfaces on speech intelligibility.

3. The aim and objectives of the study

The aim of this research was to establish the nature and the degree of influence of reflecting surfaces of a room on the evaluation of speech intelligibility by subjective and objective methods in university lecture rooms of different sizes. The obtained results will clarify the available information on the impact of early reflections of sound on speech intelligibility and better substantiate the choice of measures of speech intelligibility that are simple for calculation and at the same time reliable.

To achieve the set goal, the following tasks were set:

– to evaluate the BRIR parameters at different points of the premises;

– to assess speech intelligibility by objective and subjective methods at different points of the premises;

– to compare the estimates of the BRIR parameters with the results of subjective and objective assessment of speech intelligibility.

4. Methods for evaluating speech intelligibility in lecture rooms

4. 1. Characteristics of lecture rooms and measurement equipment

The arrangement of measurement equipment in small-size lecture room No. 1 and medium-size lecture room No. 2 are shown in Fig. 1, *a*, *b*, respectively. These lecture rooms are in building No. 12 of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Ukraine). A self-manufactured artificial head (Fig. 2, *a*) and audio equipment of different quality were used to measure the BRIR of these rooms. These are the household active speaker Genius SP-HF 2.0 500 (Taiwan), omnidirectional condenser measuring microphones Superlux ECM-999 (Taiwan), and outside sound card Superlux ECM-999 Steinberg UR242 (developed in Germany, manufactured in China).

The binaural pulse characteristics of large-size lecture room No. 3 (Fig. 1, *c*) were borrowed from the BRIR base presented and described in papers [23, 24].

According to ISO 3382-1 [16], the sound source was placed at the height of 1.5 m, and microphones were placed at the height of 1.2 m during measuring in all rooms.

The basis of the test signal was a mls-sequence having the length of 2^{16} samples, which corresponds to the duration of the signal of 1.49 s with the sampling rate of 44.1 kHz and 1.36 s at the sampling rate of 48 kHz. This mls-sequence was repeated 17 times during radiation. When calculating the BRIR of premises, the last 16 bursts of the BRIR score were averaged to increase the signal-noise ratio by 12 dB.

The test signals were recorded at the sampling rate of 44.1 kHz for premises No. 1, 2, and 48 kHz for room No. 3. The quantization depth was 24 bits in all cases. When recording the signals, the microphones were in the areas of the ears of the artificial head at a distance of 1 cm from the ear canal.

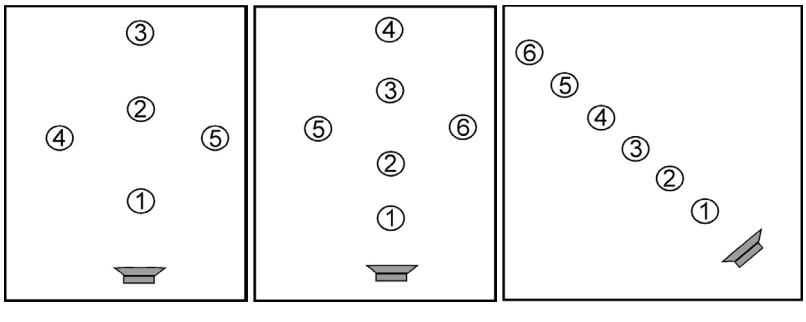

Fig. 1. Location of equipment in lecture rooms: *a* – No. 1; *b* – No. 2; *c* – No. 3

Fig. 2. Measuring audio equipment in the premises: *a* – No. 1; *b* – No. 3 [24]

The locations of the artificial head during recording the test signals are shown in Fig. 1 by numbered circles. The characteristics of the premises are given in Table 1.

Table 1

Characteristics of lecture rooms

Lecture room	No. 1	No. 2	No. 3
Volume	177 m^3	270 m^3	370 m^3
Dimensions	6.2×8.9×3.2 m	7.4×11.4×3.15 m	10.8×10.9×3.15 m
Places	37	59	25
Volume/person	4.8 m^3/person	4.6 m^3/person	14.8 m^3/person
Filling rate	12 people (32 %)	10 people (17 %)	Empty (0 %)
Distance from loudspeaker to microphone	2.4 m, 5 m, 6.9 m, 4.5 m, 4.5 m	3.5 m, 5.2 m, 8.6 m, 10.3 m, 6.4 m, 6.4 m	2.25 m, 4.0 m, 5.56 m, 7.1 m, 8.68 m, 10.2 m
Walls	3 windows, brick-plaster, back wall	4 windows, brick-plaster	3 windows, concrete-plaster
Ceiling	concrete-plaster	suspended, acoustic slabs	suspended, acoustic slabs
Floor	parquet	parquet	parquet
Furniture and its location	Wooden desks and chairs (3 rows by 6 chairs)	Wooden desks and chairs (1 row by 9 desks, 2 rows by 10 desks), 2 bookcases, and 1 wardrobe	Wooden desks and chairs (diagonal location of desks)

Comparing the data in Table 1, note a few distinguishing features of lecture room No. 3. This is a diagonal location of the desks in the room, a small number of seats, as well as the absence of students in the lecture room when taking measurements of BRIR.

4. 2. Assessing binaural pulse characteristics of lecture rooms

When measuring the BRIR, the test signal $x(t)$ was emitted by a loudspeaker located in the teacher's seat. The response $y(t)$ to stimulus $x(t)$ was recorded using a pair of microphones attached to the artificial head placed at certain points in the room (Fig. 1).

In the case of a broadband test signal with a uniform spectrum, the cross-correlation function of $K_{xy}(\tau)$ of signals $x(t)$ and $y(t)$ is proportional to BRIR $h_r(t)$ with proportionality factor k_0 [26]:

$$K_{xy}(t)\approx k_0\cdot h_r(t).$$

Given the existence of hardware distortions in the loudspeaker and in the microphone, the convolution will be estimated instead of BRIR h_r

$$h_\Sigma(t)=h_r(t)\otimes h_{lm}(t),$$

where $\otimes$ is the convolution symbol, $h_{lm}(t)=h_l(t)\otimes h_m(t)$ is the impulse response of a loudspeaker-microphone (LM) subsystem, $h_l(t)$ is the impulse response of the loudspeaker, and $h_m(t)$ is the impulse response of the microphone.

That is why at measurements, the BRIR $h_r(t)$ of lecture rooms No. 1 and No. 2 were estimated according to the expression

$$h_r(t)=\mathbb{F}^{-1}\left\{H_r(f)\right\}=\mathbb{F}^{-1}\left\{\frac{H_\Sigma(f)}{\left|H_{lm}(f)\right|}\cdot M_R(f)\right\},$$

where $H_\Sigma(f)=\mathbb{F}\left\{h_\Sigma(t)\right\}$, $H_{lm}(f)=\mathbb{F}\left\{h_{lm}(t)\right\}$, $\mathbb{F}$ and $\mathbb{F}^{-1}$ are the symbols of Fourier transforms, respectively, $|\cdot|$ is the symbol of a module, $M_R(f)$ is the regularizing multiplier [25], used to decrease the variance of estimate $h_r(t)$:

$$M_r(f)=\begin{cases}0{,}5\left[1+\cos\left(\pi f/\Delta F\right)\right], & |f|\le\Delta F;\\ 0, & |f|>\Delta F,\end{cases}$$

where ΔF (accepted that ΔF=18 kHz) is the regularization parameter [25].

4. 3. Assessment of parameters of binaural room impulse responses

In accordance with ISO 3382-1 recommendation, the BRIR parameters, such as EDT, T30, and C50 clarity index were evaluated. The ISO 3382-1 recommendations were designed primarily for concert halls, the size, shape, and decoration of which are significantly different from those for university lecture rooms. That is why verification of the suitability of these recommendations for lecture rooms is of practical interest.

In addition to these parameters, the ERB parameter was evaluated:

$$ERB=10\lg\left\{\left[\left(E_0^{50}/E_0^{10}\right)_{500}+\left(E_0^{50}/E_0^{10}\right)_{1{,}000}\right]/2\right\},$$

where $E_a^b=\int_a^b h_r^2(t)\mathrm{d}t$, a and b are the moments of time in milliseconds. The ERB parameter proposed in [3] characterizes the ratio of energies of early reflections and direct sound. Indices 500 and 1.000 mean that $h_r(t)$ was pre-filtered by octave filters with central frequencies of 500 Hz and 1.000 Hz, respectively.

Since the direct-sound energy interval of 0–10 ms specified in [3] is not in line with the parameters of the test signal used in this work, the ERBc parameter was additionally estimated. When calculating it, the action of direct sound was taken into account in the interval of 0–2 ms, i.e. E_0^2. was used instead of E_0^{10}.

4. 4. Objective assessment of speech intelligibility

Objective assessment of speech intelligibility was carried out by the modulation method [27]. In this case, the Speech Transmission Index (STI) was calculated:

$$STI=\sum_{k=1}^{7}\alpha_k\cdot MT_k-\sum_{k=1}^{6}\beta_k\cdot\sqrt{MT_k\cdot MT_{k+1}},$$

$$MT_k=\frac{1}{14}\sum_{i=1}^{14}T_{ki},$$

$$T_{ki}=\begin{cases}0, & E_{ki}<-15;\\ \left(E_{ki}+15\right)/30, & -15\le E_{ki}\le+15;\\ 1, & E_{ki}<-15,\end{cases}$$

$$E_{ki}=10\lg\frac{m_{ki}}{1-m_{ki}},$$

$$m_{ki}=\left|\int_0^\infty h_{rk}^2(t)\exp\left(-j2\pi F_i t\right)\mathrm{d}t\right|\Big/\int_0^\infty h_{rk}^2(t)\mathrm{d}t,$$

where α_k is the weight factors, β_k is the redundancy factors, $h_{rk}(t)$ is the result of filtration $h_r(t)$ with the k-th octave-band filter, F_i=0.63–12.5 Hz. Filtration $h_r(t)$ was implemented using 7 octave-band filters with central frequencies from 125 Hz to 8 kHz [27].

4. 5. Subjective assessment of speech intelligibility

Articulation measurements of speech intelligibility in all lecture rooms were performed by listening to test signals distorted by noise and reverberation. Listening was performed using special software, headphones, and computers [28].

The subjects randomly listened to 3 sets of single-syllable words for different combinations of the signal-to-noise ratio and BRIR. Each set contained 50 monosyllabic words such as consonant-vowel-consonant (CVC). 9 speakers, which included 7 men and 2 women took part in the formation of these sets. The anchor phrase was used in reading every word, which made it possible to take into consideration the impact of reverberation on speech intelligibility. For example, the word "niav"/niav/was read as "Write down niav now". After a listener recorded the perceived word with the keyboard, speech intelligibility was calculated automatically with the appropriate software.

Test stimuli were synthesized with the same software. At the same time, white noise $n(t)$ of the controlled level was added to the $s(t)$ signal, after which the convolution of the obtained mixture with the initial part of BRIR $h_r(t)$ at the length of 50ms was calculated. This method of synthesis made it possible to generate speech signals distorted only by early reflections and background noise [4].

Fourteen participants took part in the testing of rooms No. 1 and No. 2, and 20 listeners took part in testing in room No. 3. All listeners, aged 20–23, had normal hearing, Ukrainian was their native language.

5. Results of evaluation of parameters of binaural room impulse response and speech intelligibility

5. 1. Assessing the parameters of binaural room impulse response

The BRIR parameters were evaluated at two stages. At stage 1, according to the procedure described in p. 4. 2, the BRIR of each lecture room was assessed at different points of the rooms (Fig. 1). Parameters T30, EDT, C50, ERB, and ERBc were assessed at stage 2 in accordance with the procedure described in p. 4. 3.

The results of assessing parameters T30, EDT, C50, ERB, and ERBc are shown in Fig. 3–6.

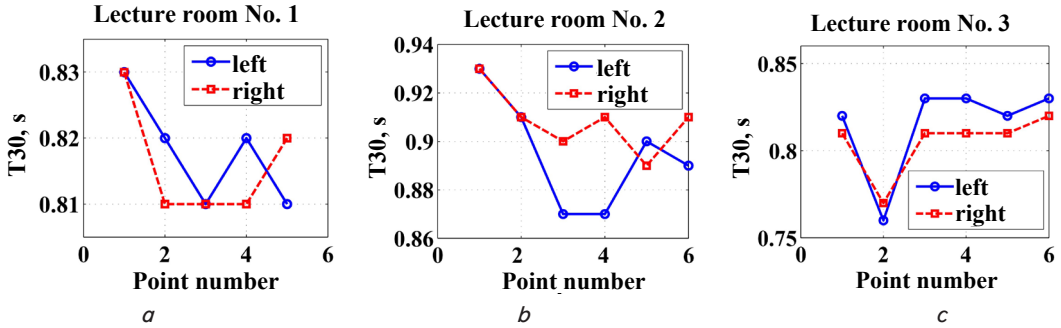

Fig. 3. Scores of T30 in lecture rooms: *a* – No. 1; *b* – No. 2; *c* – No. 3

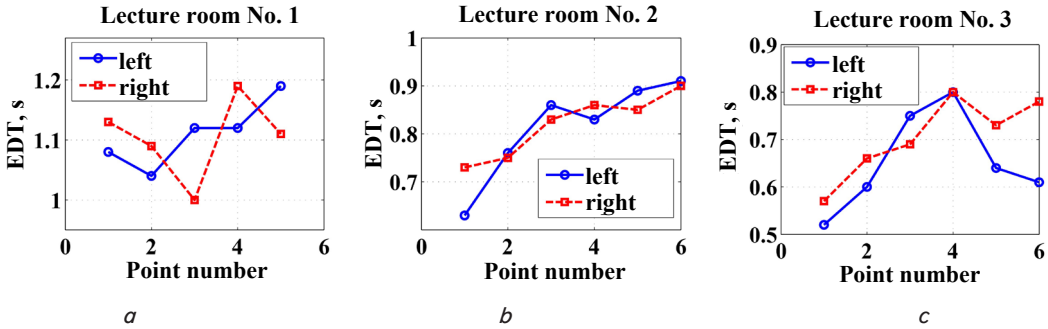

Fig. 4. Scores of EDT in lecture room: *a* – No. 1; *b* – No. 2; *c* – No. 3

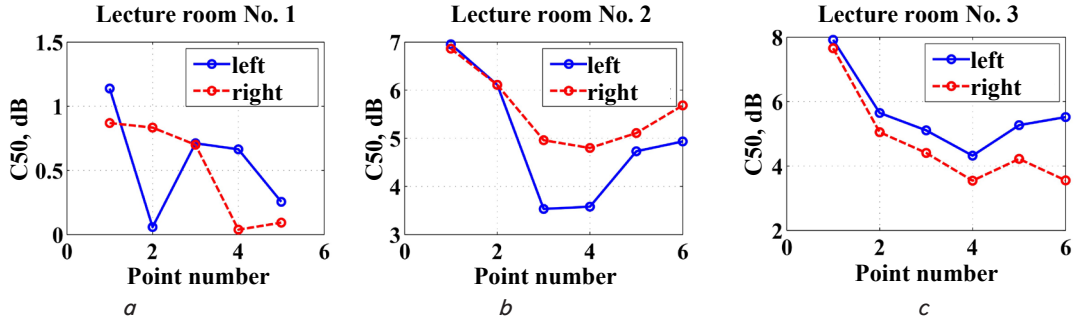

Fig. 5. Scores of C50 in lecture rooms: *a* – No. 1; *b* – No. 2; *c* – No. 3

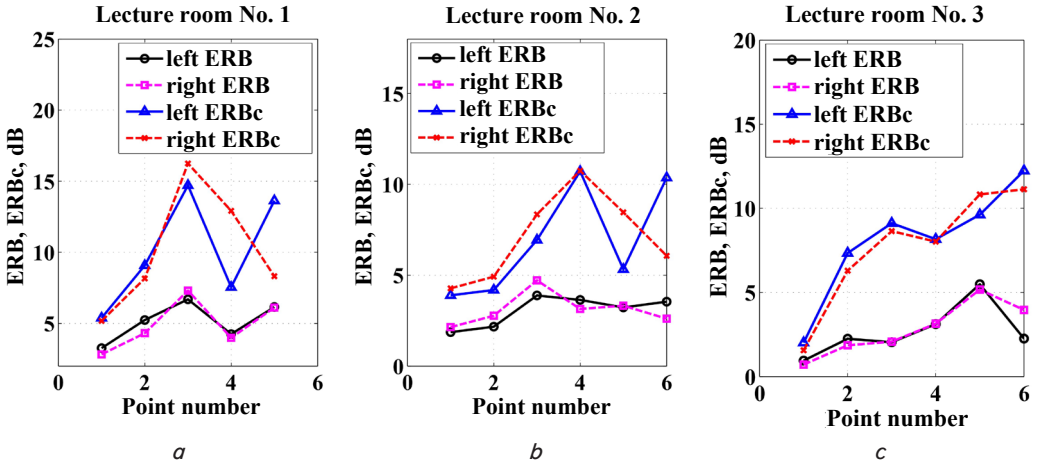

Fig. 6. Scores of ERB and ERBc in lecture rooms: *a* – No. 1; *b* – No. 2; *c* – No. 3

Owing to the existence in Fig. 3–6 of the scores corresponding to the left and right channels of the human auditory system, it is possible to take into account more fully the impact of reflecting surfaces on the BRIR parameters at different points of a room.

5. 2. Assessing speech intelligibility by objective and subjective methods

The scale of intelligibility rating and values of speech transmission index STI, which is an objective measure of speech intelligibility, is shown in Table 2 [27]. According to the German standard DIN 18041:2004-05 [9], STI>0.56 is an acceptable value for school classes. The British standard BS EN 60268-16 [12] recommends STI≥0.62 in school classrooms and lecture rooms in the "theater" form.

STI values obtained in accordance with the assessment procedure described in p. 4. 4 are shown in Fig. 7.

Table 2

Relations of STI values of speech intelligibility [27]

Intelligibility rating	STI
Excellent	>0.75
Good	0.60–0.75
Fair	0.45–0.60
Pour	0.30–0.45

Speech intelligibility was subjectively assessed in accordance with the procedure, described in p. 4. 5, only for points between the speaker and the back wall. The values of the subjective assessments of speech intelligibility, averaged by the number of listeners, are shown in Fig. 8 for two values of signal-to-noise ratio (SNR). The values of the standard deviation of these scores are marked upwards and downwards from average intelligibility values.

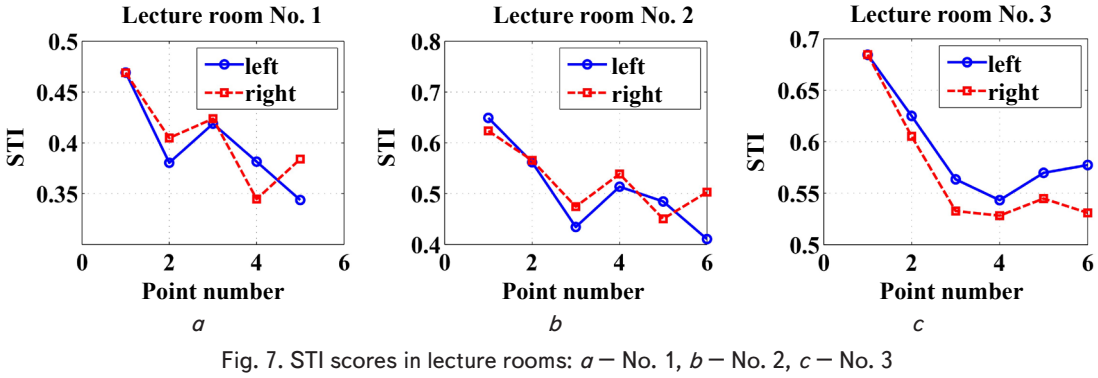

Fig. 7. STI scores in lecture rooms: *a* – No. 1, *b* – No. 2, *c* – No. 3

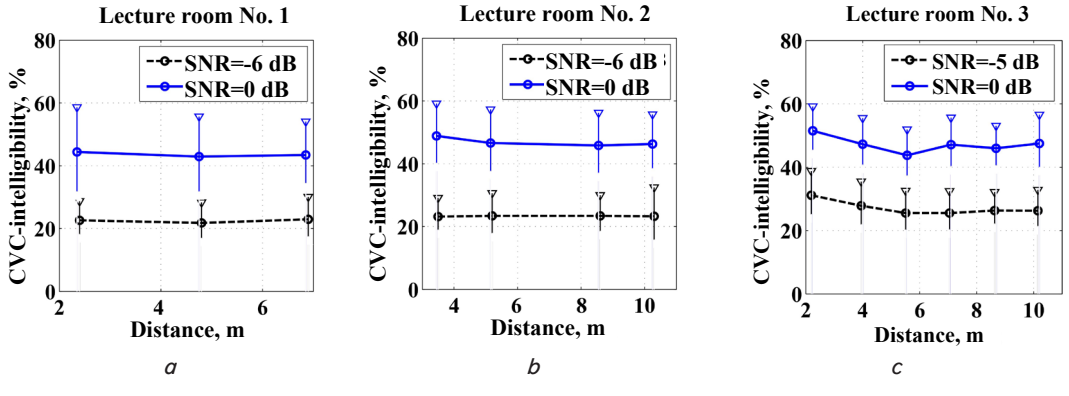

Fig. 8. Scores of CVC-intelligibility in lecture rooms: *a* – No. 1; *b* – No. 2; *c* – No. 3

The results of subjective and objective assessment of speech intelligibility serve as reference points in the analysis of comparative effectiveness of the BRIR parameters used as measures of speech intelligibility.

5. 3. Correlation of speech intelligibility scores

Table 3 shows the scores of Pearson correlation coefficient *r*, which characterize the statistical relationship between the values of the BRIR parameters discussed above and the values of the STI scores.

Table 3

Correlation factors between speech intelligibility scores

Parameter	Lecture room No. 1			Lecture room No. 2			Lecture room No. 3		
	Left	Right	Mean	Left	Right	Mean	Left	Right	Mean
T30	0.66	0.78	0.700	0.81	0.91	0.89	–0.43	–0.28	–0.36
EDT	–0.50	–0.43	–0.519	–0.95	–0.72	–0.94	–0.86	–0.90	–0.93
C50	0.85	0.85	0.978	0.80	0.92	0.86	0.95	0.97	0.97
ERB	–0.54	–0.17	–0.353	–0.88	–0.73	–0.85	–0.59	–0.73	–0.66
ERBc	–0.49	–0.41	–0.369	–0.64	–0.58	–0.60	–0.81	–0.92	–0.88
CVC-intelligibility	0.99	0.99	1.00	0.93	0.91	0.92	0.80	0.82	0.81

The last line of Table 3 contains scores *r* for STI and the results of subjective evaluation of speech intelligibility. The columns "Left", "Right", and "Mean" contain *r* scores for left and right channels, as well as the mean for channels, respectively.

6. Discussion of the results of the evaluation of speech intelligibility in university lecture rooms of different sizes

6. 1. Dependence of scores of parameters of binaural room impulse response on the location of a listener

As it follows from Fig. 3, the values of T30 in each of the premises differ slightly at different points of the room (no more than 5 %, 7 %, and 13 % in premises No. 1–3, respectively). T30 values in the left and right channels are also slightly different (no more than 1 %, 4 %, and 1 % in rooms 1–3, respectively).

On the contrary, the EDT values (Fig. 4) differ significantly in rooms No. 1–3 (13 %, 30 %, and 35 %, respectively). Differences in the left and right channels (11 %, 14 %, and 22 %, respectively) are also more pronounced.

Given the significant variability of the EDT scores, it can be assumed that they better characterize the variability in speech intelligibility at different points in the room than the T30 scores. For example, one can make a preliminary conclusion about mediocre intelligibility in room No. 1 at any point in the room due to the increased EDT values due to the reverberation. In room No. 2, the situation noticeably improves at points 1 and 2, but at the rest points of the room, the EDT values do not meet the requirements of the DIN 18041:2004-05 standard [17]. In room No. 3, one can expect some deterioration in speech intelligibility at point 4, compared to the rest of the points, because at point 4, the EDT values are higher than those for the rest of the points.

Judging by the EDT, on average, the best intelligibility should be provided by room No. 3, and the worst – by room No. 1. However, a comparison of T30 values in rooms No. 1 and No. 2 leads to a different conclusion, namely, that room No. 2 has the worst intelligibility. The comparison of the T30 scores (Fig. 3) and the EDT scores (Fig. 4) with the STI index (Fig. 7) indicates that the EDT parameter used as a speech intelligibility measure provides higher accuracy than the T30 parameter. The results are well in line with the recommendations of ISO 3382-1 standard [16] to use T30 as a measure of the physical properties of a room, and the EDT as a measure of the sound perception by listeners.

The points on the line between the speaker and the back wall of the room are of special interest, as studies [35, 36] underlined a tendency of increasing speech intelligibility when one approaches the back wall closer. Analysis of Fig. 4, *c*, testifies that the EDT values decrease with the movement from point 4 to point 5, however, the EDT behavior in the left and right channels at point 6 is contradictory. There is also a similar inconsistency of behavior of the EDT values in rooms No. 1 (Fig. 4, *a*) and No. 2 (Fig. 4, *b*). That is why in general, we can conclude that the EDT parameter does not adequately reflect the expected increase in speech intelligibility at the back wall.

The behavior of dependences of C50 scores on the numbers of room points shown in Fig. 5 with approaching to the back wall of the room is equally contradictory.

The reason for the noted contradictory behavior of the two-channel assessments of parameters EDT and C50 near the back wall may be a slight degree of increased speech intelligibility with approaching the wall. In this case, the measurement results can be noticeably affected by both the error of measurements and the difference in the reflectivity of the left and right walls. Note that the difference in the reflectivity of the left and right walls can indeed prove to be a very significant factor. The windows that reflect sound better than walls are usually to the left of listeners, so the signal level in the left channel will be higher at the same distance of a listener from the side walls. The behavior of the C50 values in Fig. 5 for the points near the back wall of a lecture room is quite consistent with these considerations. However, it is clear that in order to increase the reliability of findings, the volume of statistics should be significantly increased by obtaining double-channel EDT and C50 scores for a large number of other university lecture rooms.

Dependences of the scores of the ERB and ERBc parameters on the numbers of points in a room, left and right channels, are shown in Fig. 6. One can see that the ERB values for points located between the speaker and the back wall of a room do not increase monotonously in rooms No. 2 and No. 3. This may seem strange, as the role of early reflections of a sound should increase as the distance from the speaker increases.

However, ERBc values for the same points increase monotonously almost everywhere, except for point 4 of room No. 3. Thus, correct accounting of the size of the BRIR section corresponding to direct sound makes it possible to obtain the scores of the ratio of energies of direct sound and early reflections, which are mainly consistent with the known results [3]. Note, however, that in order to obtain this

consistency, it was necessary to adjust the assessment of the ERB parameters proposed in [3].

As for the noticed "subsidence" of the ERBc diagram at point 4 in Fig. 6, *c*, it may indicate a deterioration in speech intelligibility in the center of a lecture room due to a change in the temporal structure of early reflections. In order to clarify the causes of this deterioration, it is advisable to conduct subsequent modeling and increase the volume of statistics in the future.

6. 2. Dependence of speech intelligibility scores on the listener's location

Analyzing the behavior of STI scores (Fig. 7), we see that for the points located between the speaker and the back wall of the room, the intelligibility is minimal in the middle of the room and increases near the back wall. For room No. 3, in the right channel speech intelligibility increases at point 5, although it is decreased at point 6. The behavior of the ETD and C50 characteristics is similar at point 6. Since point 6 is in the corner between the back and right walls and closer to the right wall, we can assume that similar behavior of intelligibility scores will be observed at point 4 of room No. 1 and at point 5 of room No. 2. Analysis of Fig. 7, *a*, *b*, proves the validity of this assumption. As we can see, in all rooms there is an effect of the negative influence of early reflections from the side wall closest to a listener. The observed phenomenon can be explained by distortion of the form of a sound signal and, as a result, deterioration of the quality of a perceived signal, due to the discrete structure of the BRIR in the interval of 0–50 ms [37].

Comparing the STI scores near the back wall and in the middle of the room (Fig. 7), we see that the increase in STI was 5–10 % in room No. 1, 13–16 % in room No. 2 and 2–7 % in room No. 3. In the subjective evaluation of intelligibility (Fig. 8), this excess is less noticeable and amounted to 1–3 %, which can be explained by the large variance of measurement results (17–29 %) because of the difference in the psychophysical characteristics of the listeners.

The STI scores near the side walls obtained for rooms No. 1, 2, averaged for the channels, are close to similar scores in the middle of the room. This can be explained by the fact that in these rooms, measurements were made at distances of 0.7–1.7 m from the side walls, where an increase in sound level due to the interference of waves near reflective surfaces can be neglected.

During the subjective evaluation of speech intelligibility, listeners were offered test signals distorted not only by reverberation but also by noise. As it follows from the diagrams in Fig. 8, noise affects speech intelligibility much more noticeably than the reverberation. Indeed, the change in the signal-noise ratio by 5–6 dB led to a change in speech intelligibility by 20–25 %. At the same time, the change in speech intelligibility within the same room was 3 %, 6 %, and 15 % for rooms No. 1–3, respectively. The obtained results suggest that the variability of intelligibility scores inside a lecture room increases at an increase in the size of this lecture room. Of course, the correctness of this assumption should be verified in the future.

6. 3. Dependence between the parameters of binaural room impulse response and speech intelligibility

Analysis of the correlation between BRIR parameters and speech intelligibility scores (Table 3) revealed that the EDT scores are much more consistent with the STI scores compared to T30 scores. Indeed, the behavior of the T30 is contradictory: the correlation is positive for rooms No. 1, 2, but negative for room No. 3. At the same time, in the case of the EDT, the correlation is negative for all 3 rooms, in this case, the module of r value reaching high values of 0.9–0.95 for rooms No. 2, 3. The obtained results indicate that the known provisions on the possible use of the EDT as an objective measure of speech intelligibility in concert halls [16] are also true for medium- and large-size lecture rooms. However, a decrease in module r to 0.4–0.5 for room No. 1 shows that the EDT can be a very rough measure for small premises. The validity of this assumption should be tested in the future by analyzing the acoustic properties of a large number of lecture rooms.

The clarity index C50 proved to be most effective as an objective measure of speech intelligibility, as the values of r scores are consistently high (0.8–0.98) for lecture rooms of all sizes.

Since parameters ERB and ERBc are not designed to assess speech intelligibility, it should not be surprising that the corresponding values of the correlation factor are not always high. However, the stability of the sign of the correlation factor indicates that such relations exist. However, the negative values of r make it possible to conclude that the more energy of early reflection prevails over the energy of direct sound, the worse the speech intelligibility. This result is consistent with the results of research [7], which shows that the action of early reflections is not equivalent to the action of direct sound in the sense of ensuring the same speech intelligibility.

The degree of correlation of subjective scores of intelligibility (CVC-intelligibility) and STI scores for rooms No. 1, 2 may seem suspiciously high. However, this can be explained by a small number of points between the speaker and the back wall. That is why it seems appropriate to place at least 4 points between the reader and the back wall of a lecture room (provided that the shape of a lecture room allows doing it) in subsequent measurements.

7. Conclusions

1. The dependences of the scores of parameters T30, EDT, C50, and ERB of binaural room impulse response on the place of a listener in university lecture rooms of different sizes were obtained experimentally. It was shown that at an increase in the distance between a listener and a speaker, the increase in the ratio of the energy of early reflections to the energy of direct sound can be non-monotonous in large lecture rooms. A possible cause of this phenomenon may be a change in the temporal structure of early reflections. It is appropriate to test the validity of the obtained results in the future on an increased volume of statistical data.

2. The scores of speech intelligibility obtained using objective and subjective methods are of a reference nature, allowing to quantify the value of possible alternative measures of speech intelligibility T30, EDT, and C50. In addition, the use of these methods made it possible to detect a small increase in speech intelligibility at small distances from the back wall of up to 0.3–0.4 m. During the subjective assessment of speech intelligibility, this increase did not exceed 3 %, and during objective evaluation with the use of the STI measure, such increase is more noticeable, reaching 15 %. In addition, the two-channel nature of the STI assessment allowed revealing the effect of the negative influence of early

reflections on speech intelligibility in the auditory channel close to the side wall of the room at distances of 0.7–1.7 m from the wall. The most likely cause of this effect is a deterioration in sound quality due to the action of early reflections. The validity of this assumption should be verified in the future by assessing the degree of signal distortion in each of the measuring channels.

3. The calculation of correlation coefficients between the values of parameters T30, EDT, C50, and STI enabled quantifying the usefulness of these parameters as possible alternative measures of speech intelligibility. C50 was shown to be the most informative as a measure of speech intelligibility, while parameter T30 turned out to be the least informative. Parameter EDT as a measure of speech intelligibility takes an intermediate place, however, its use may be ineffective in small lecture rooms. It is subsequently advisable to clarify the obtained results by increasing the number and the variety of analyzed lecture rooms.

References

1. Eggenschwiler, K. (2005). Lecture Halls - Room Acoustics and Sound Reinforcement. Conference: 4th European congress on acoustics (Forum Acusticum 2005). Budapest. Available at: https://www.researchgate.net/publication/238114069_Lecture_Halls_-_Room_Acoustics_and_Sound_Reinforcement
2. Bradley, J. S., Reich, R. D., Norcross, S. G. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. The Journal of the Acoustical Society of America, 106 (4), 1820–1828. doi: https://doi.org/10.1121/1.427932
3. Bradley, J. S., Sato, H., Picard, M. (2003). On the importance of early reflections for speech in rooms. The Journal of the Acoustical Society of America, 113 (6), 3233. doi: https://doi.org/10.1121/1.1570439
4. Hu, Y., Kokkinakis, K. (2014). Effects of early and late reflections on intelligibility of reverberated speech by cochlear implant listeners. The Journal of the Acoustical Society of America, 135 (1), EL22–EL28. doi: https://doi.org/10.1121/1.4834455
5. Yang, W., Bradley, J. S. (2009). Effects of room acoustics on the intelligibility of speech in classrooms for young children. The Journal of the Acoustical Society of America, 125 (2), 922–933. doi: https://doi.org/10.1121/1.3058900
6. Bradley, J. S. (2011). Review of objective room acoustics measures and future needs. Applied Acoustics, 72 (10), 713–720. doi: https://doi.org/10.1016/j.apacoust.2011.04.004
7. Arweiler, I., Buchholz, J., Dau, T. (2009). Speech intelligibility enhancement by early re ections. ISAAR 2009: Binaural Processing and Spatial Hearing. 2nd International Symposium on Auditory and Audiological Research. Elsinore, Denmark. Available at: https://proceedings.isaar.eu/index.php/isaarproc/article/view/2009-29
8. Warzybok, A., Rennies, J., Doclo, S., Kollmeier, B. (2011). Influence of early reflections on speech intelligibility under different noise conditions. Forum Acusticum. Aalborg, Denmark. Available at: https://uol.de/f/6/dept/mediphysik/ag/sigproc/download/papers/SP2011_5.pdf
9. Nocke, C. (2016). New standards in Architectural Acoustics - a German view. Proceedings of ACOUSTICS 2016. Brisbane, Australia. Available at: https://www.acoustics.asn.au/conference_proceedings/AASNZ2016/papers/p114.pdf
10. DIN 18041:2004-05. Acoustic quality in small to medium-sized rooms (2004). Beuth Verlag GmbH, 39. Available at: https://www.beuth.de/de/norm/din-18041/69640391
11. ANSI/ASA S12.60-2010/Part 1 (R2015). Acoustical performance criteria, design requirements, and guidelines for schools, Part 1: permanent schools (2015). American National Standards Institute. Available at: https://webstore.ansi.org/standards/asa/ansiasas12602010partr2015
12. BS EN 60268-16:2011. Sound System Equipment - Part 16: Objective rating of speech intelligibility by speech transmission index (2011). British Standards Institution. Available at: https://infostore.saiglobal.com/en-us/Standards/BS-EN-60268-16-2011-217511_SAIG_BSI_BSI_512663/
13. IEC 60268-16:4.0. Sound system equipment - Part 16: objective rating of speech intelligibility by speech transmission index (2011). International Electrotechnical Committee. Available at: https://infostore.saiglobal.com/en-us/Standards/IEC-60268-16-4-0-569917_SAIG_IEC_IEC_1301856/
14. Acoustic design of schools: performance standards. Building bulletin 93 (2015). UK Department for Education, UK Education Funding Agency. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/400784/BB93_February_2015.pdf
15. Canning, D., James, A. (2012). The Essex Study. Optimised classroom acoustics for all. The Association of Noise Consultants. Available at: https://www.ndcs.org.uk/media/1808/the_essex_study_optimised_classroom_acoustics_for_all1.pdf
16. ISO 3382-1:2009(en). Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces. Available at: https://www.iso.org/obp/ui/#iso:std:iso:3382:-1:ed-1:v1:en
17. Sari, L. F., Utami, S. S., Sarwono, J. (2018). Objective and subjective acoustics measurement of audience seating areas in a medium size auditorium. Journal of Physics: Conference Series, 1075, 012015. doi: https://doi.org/10.1088/1742-6596/1075/1/012015
18. Choi, Y.-J. (2020). The intelligibility of speech in university classrooms during lectures. Applied Acoustics, 162, 107211. doi: https://doi.org/10.1016/j.apacoust.2020.107211
19. Eldakdoky, S. (2017). Optimizing acoustic conditions for two lecture rooms in Faculty of Agriculture, Cairo University. Ain Shams Engineering Journal, 8 (4), 481–490. doi: https://doi.org/10.1016/j.asej.2016.08.013

20. Leccese, F., Rocca, M., Salvadori, G. (2018). Fast estimation of Speech Transmission Index using the Reverberation Time: Comparison between predictive equations for educational rooms of different sizes. Applied Acoustics, 140, 143–149. doi: https://doi.org/10.1016/j.apacoust.2018.05.019
21. Nestoras, C., Dance, S. (2013). The Interrelationship between Room Acoustics Parameters as Measured in University Classrooms Using Four Source Configurations. Building Acoustics, 20 (1), 43–53. doi: https://doi.org/10.1260/1351-010x.20.1.43
22. Jacob, K. (1989). Correlation of Speech Intelligibility Tests in Reverberant Rooms with Three Predictive Algorithms. JAES, 37 (12), 1020–1030. Available at: https://www.aes.org/e-lib/browse.cfm?elib=6056
23. Jeub, M., Schafer, M., Vary, P. (2009). A binaural room impulse response database for the evaluation of dereverberation algorithms. 2009 16th International Conference on Digital Signal Processing. doi: https://doi.org/10.1109/icdsp.2009.5201259
24. Aachen Impulse Response Database. Available at: https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/
25. Tikhonov, A. N. (1965). On incorrect problems of linear algebra and a stable method for their solution. Doklady Akademii nauk SSSR, 163 (3), 591–594. Available at: http://www.mathnet.ru/links/c317fafa3fe690e2d1c52f383762a0e2/dan31374.pdf
26. Kuttruff, H. (2009). Room Acoustics. CRC Press, 392. doi: https://doi.org/10.1201/9781482266450
27. Steeneken, H. (2014). Forty years of speech intelligibility assessment (and some history). Proceedings of the Institute of Acoustics, 36 (3). Available at: http://www.steeneken.nl/wp-content/uploads/2014/04/IOA-Forty-years-keynote.pdf
28. Prodeus, A., Bukhta, K., Morozko, P., Serhiienko, O., Kotvytskyi, I., Shherbenko, I. (2018). Automated System for Subjective Evaluation of the Ukrainian Speech Intelligibility. 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO). doi: https://doi.org/10.1109/elnano.2018.8477568
29. Naida, S., Didkovskyi, V., Pavlenko, O., Naida, N. (2019). Spectral Analysis of Sounds by Acoustic Hearing Analyzer. 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO). doi: https://doi.org/10.1109/elnano.2019.8783915
30. Naida, S., Didkovskyi, V., Pavlenko, O., Naida, N. (2019). Objective Audiometry Based on the Formula of the Middle Ear Parameter: A New Technique for Researches and Differential Diagnosis of Hearing. 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO). doi: https://doi.org/10.1109/elnano.2019.8783502
31. Waterhouse, R. V. (1955). Interference Patterns in Reverberant Sound Fields. The Journal of the Acoustical Society of America, 27 (2), 247–258. doi: https://doi.org/10.1121/1.1907509
32. Jacobsen, F., Tiana-Roig, E. (2010). Measurement of the Sound Power Incident on the Walls of a Reverberation Room with Near Field Acoustic Holography. Acta Acustica United with Acustica, 96 (1), 76–81. doi: https://doi.org/10.3813/aaa.918258
33. Duran, S., Ausiello, L., Battaner-Moro, J. (2019). Acoustic Design Criteria for Higher-Education Learning Environments. Proceedings of the Institute of Acoustics, 41 (3), 1–12. Available at: https://pure.solent.ac.uk/ws/portalfiles/portal/10957857/Duran_Ausiello_Battaner_Moro_SOLENT_REPRODUCED_SOUND_2019_.pdf
34. Jedidi, M., Boulila, A. (2016). Acoustic study of an auditorium by the determination of reverberation time and speech transmission index. Int. J. Architect. Eng. Urban Plan, 26 (1), 25–32. Available at: http://ijaup.iust.ac.ir/article-1-312-en.pdf
35. Prodeus, A., Didkovska, M., Motorniuk, D., Dvornyk, O. (2020). The Effects of Noise, Early and Late Reflections on Speech Intelligibility. 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO). doi: https://doi.org/10.1109/elnano50318.2020.9088854
36. Prodeus, A., Didkovska, M. (2020). Objective assessment of speech intelligibility in small and medium-sized classrooms. Conference: 2020 IEEE International Conference "Problems of Infocommunications. Science and Technology" (PIC S&T'2020). Available at: https://www.researchgate.net/publication/347490796_Objective_Assessment_of_Speech_Intelligibility_in_Small_and_Medium-Sized_Classrooms
37. Prodeus, A., Didkovska, M., Kukharicheva, K., Motorniuk, D. (2020). Modeling the Influence of Early Sound Reflections on Speech Intelligibility. 2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC). doi: https://doi.org/10.1109/msnmc50359.2020.9255657
38. Lochner, J. P. A., Burger, J. F. (1964). The influence of reflections on auditorium acoustics. Journal of Sound and Vibration, 1 (4), 426–454. doi: https://doi.org/10.1016/0022-460x(64)90057-4