Date of publication JUN-30, 2024, date of current version JAN-25, 2024.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.23.2.3539

Comparison of Direct and Indirect Methods of Speech Transmission Index Assessment

ARKADIY PRODEUS[1], OLEKSANDR DVORNYK[1], ANTON NAIDA[1]
[1]Department of Acoustic and Multimedia Electronic Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine, (e-mail: aprodeus@gmail.com, alexanderdvornyk@gmail.com, naida.a.s.2001@gmail.com)

Corresponding author: Arkadiy Prodeus (e-mail: aprodeus@gmail.com).

ABSTRACT In this paper, direct and indirect methods of speech transmission index (STI) estimation are compared. Two versions of the indirect method of the STI estimating are considered. In the first version of the indirect method, a pair of signals is used as a test signal. It is a Maximum Length Sequence (MLS) signal with a uniform spectrum and a noise signal with a speech spectrum. In the second version of the indirect method, the test signal is an MLS signal with a speech spectrum. The comparison is carried out by means of computer modeling and by carrying out a field experiment in a medium sized university auditorium. Both versions of the indirect method use the same basic computer programs for STI calculating. It is shown that for the second version of the indirect method, the average values of the STI estimates differ from ones for the direct method by no more than 0.06 for signal-to-noise ratios from minus 20 dB to plus 20 dB. For the first version of the indirect method, this difference is significantly larger and reaches 0.24.

KEYWORDS speech transmission index; assessment; method; speech spectrum; simulation; field experiment

I. INTRODUCTION

The assessment of speech intelligibility in premises and communication lines is an important task, as it allows identifying deficiencies in the acoustic characteristics of the objects being tested and to develop reasonable recommendations for correcting these deficiencies.

The most common method for assessing speech intelligibility is the modulation method [1], the important advantage of which over the formant method [2] is the ability to take into account the distortion of the speech signal not only by noise, but also by reverberation.

In the full direct modulation method, the measure of speech intelligibility is the STI [1]. In addition, there are a number of so-called "rapid" full direct methods, such as STIPA (STI for Public Address systems), STITEL (STI by Telecommunication) [1]. The disadvantage of the full direct method is the need to use 98 test signals, each of which has a duration of at least 10 seconds, as a result of which the duration of STI measurements is close to 15 minutes with an estimation error of 0.02 [1]. Rapid methods, where only one test signal is used, can reduce the duration of STI measurements to 10-15 seconds.

STI can also be calculated using an indirect method. The indirect method of STI measurements, based on the preliminary assessment of the room impulse response and the use of the Schroeder equation, is also more advantageous than the direct method, since the duration of measurements is significantly reduced [1]. However, there is no information regarding the accuracy of STI measurements by the indirect method in [1], and there is no discussion of an alternative version of the indirect method, when two test signals emitted in series are used for measurements. The purpose of this article is to find answers to these questions.

II. STATE OF THE ART STUDIES

There are two possible ways of implementing the indirect method of STI measurement using test MLS signals:

1) the use of MLS signal with a uniform spectrum and a separate noisy test signal with speech spectrum (this variant of the indirect STI measurement method will be referred to as IN1);
2) the use of MLS signal with speech spectrum (this indirect method will be referred to as IN2).

The ability to measure STI by an indirect method is implemented in a number of commercial hardware and software applications.

In the DIRAC system [3], [4] it is recommended to use the IN2 method instead of the linear sweep or exponential sweep signal. The AURORA software application [5] can use the IN1 or IN2 method.

The CLIO software [6] implements only the IN2 method.

The measuring instrument XL2 Audio and Acoustic Analyzer [7], unlike the systems mentioned above, allows one to measure STIPA by a direct method.

It is possible to indicate a set of studies [8], [9], [10], [11], where the comparison of direct and indirect STI assessment methods was carried out using the applications mentioned above.

Note that the software scripts of all mentioned above applications are not available to the end user, so comparison of the accuracy of STI estimates obtained using these applications appears to be incorrect. A research method where the researcher does not completely control the algorithm and the calculation program is difficult to consider correct. Another shortcoming of the above studies is the absence of information on the difference in indirect STI estimates obtained using IN1 or IN2 methods. Moreover, in these studies, it is not clearly indicated which variant of the indirect method was used, IN1 or IN2.

To eliminate these shortcomings, a multicomponent test signal and algorithms for its processing are proposed in [12]. A trial comparison of IN1 and IN2 methods with the direct rapid formant-modulation (RFM) method was performed in [12].

II. PROBLEM STATEMENT

A general shortcoming of studies [8], [9], [10], [11] is the lack of information about the type of implemented indirect method, IN1 or IN2. As a result, there is no explicit comparison of STI assessment results according to IN1 and IN2 methods. In addition, the authors of these studies did not have full control over the software scripts of commercial hardware and software applications when evaluating STI.

A trial comparison of IN1 and IN2 methods performed by means of computer simulation is realized in [12] using author's computer programs. This comparison showed that both indirect methods lead to underestimated STI values compared to ones for the direct method. At the same time, it turned out that the use of IN2 method leads to a smaller error of the STI estimate, compared to the use of IN1 method. A disadvantage of the studies carried out in [12] is the lack of modeling of STI estimation for different values of the signal-to-noise ratio (SNR). Another drawback is the use of simplified model of speech distortion by noise and reverberation, which does not take into account possible nonlinear sound distortions in the room and in the sound equipment.

The objective of the paper is to eliminate these shortcomings by simulating the STI estimation procedure over a wide range of SNR values, and by conducting field studies estimating STI at different points in a real room.

III. EXPERIMENTAL SETUP

A. TEST SIGNAL

A multicomponent test signal [12]

$$x_{test}(t) = x_{sil}(t) \oplus x_{MLS}(t) \oplus x_{sil}(t) \oplus \\ \oplus x_{n\,spch}(t) \oplus x_{sil}(t) \oplus x_{AM\,spch}(t) \oplus , \\ \oplus x_{sil}(t) \oplus x_{MLS\,spch}(t) \oplus x_{sil}(t) \qquad (1)$$

was used for the studies, $x_{MLS}(t)$ is MLS signal with uniform spectrum lasting $T_{MLS} = 6$ s, $x_{n\,spch}(t)$ is a segment of Gaussian noise with speech spectrum lasting $T_{n\,spch} = 4$ s, $x_{AM\,spch}(t)$ is amplitude-modulated noise (1) with speech spectrum lasting $T_{AM\,spch} = 16$ s, $x_{MLS\,spch}(t)$ is MLS signal with speech spectrum lasting $T_{MLS\,spch} = 24$ s, $x_{sil}(t)$ is pause lasting $T_{sil} = 2$ s, $\oplus$ is a symbol of concatenation of individual signals and pauses into a single multi-component test signal.

The structure of the test signal (3) is shown in Fig. 1, and its form is shown in Fig. 2.

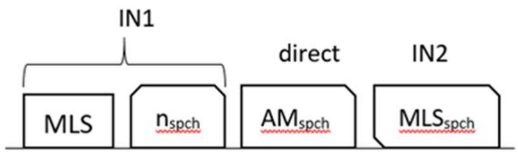


Figure 1. The structure of a multicomponent test signal [12]

The advantage of the test signal is the ability to perform STI measurements by different methods under the same interference conditions. Components of the multicomponent test signal are separated by 2-second pauses. This duration of pauses is sufficient for STI measurements in classrooms and offices, where the reverberation time usually does not exceed 1 s. Indeed, in this case, the pauses will be guaranteed to have intervals lasting at least 1 s, where the effect of reverberation is practically absent. The intervals are required for SNR estimating.

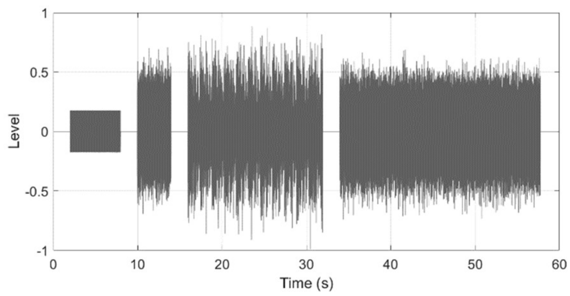


Figure 2. Form of multicomponent test signal $x_{test}(t)$ [12]

The variances of all components of the test signal are the same, which ensures further correct comparison of measurement results.

B. DIRECT METHOD OF STI ESTIMATION

The RFM method was used as a direct method of STI evaluating [13]. The test signal in this case has the form of an amplitude-modulated random process

$$x_{AM\,spch}(t) = \xi(t)\sqrt{f_5(t)}\,, \qquad (2)$$

$$f_5(t) = 1 + 0{,}32 \cdot \sum_{i=1}^{5} \sin 2\pi F_i t\,,$$

$$F_i = iF_m, \quad F_m = 0,7 \text{ Hz},$$

where $\xi(t)$ is a stationary random process with a speech spectrum, $f_5(t)$ is the law of $\xi(t)$ variance modulation with the basic modulation period $T_m = 1/F_m$ =1.43 s. Signal (2) is the third part of the multicomponent test signal in Fig. 1.

The processing of the signal $y(t)$ received by the microphone is aimed at calculating the STI:

$$STI = \sum_{k=1}^{7} \alpha_k \cdot MTI_k - \sum_{k=1}^{6} \beta_k \cdot \sqrt{MTI_k \cdot MTI_{k+1}}, \quad (3)$$

$$MTI_k = \frac{E_k + 15}{30}, \quad E_k = \frac{1}{5}\sum_{i=1}^{5} SNR_{eff\,k,i},$$

$$SNR_{eff\,k,i} = 10\lg\frac{m_{k,i}}{1 - m_{k,i}}, \quad m_{k,i} = \frac{6.25\,|\,A_k(F_i)\,|}{|\,A_k(0)\,|},$$

$$A_k(F_i) = \frac{1}{T}\int_0^T y_k^2(t) e^{-j2\pi F_i t} dt,$$

where MTI_k is modulation transfer index in the k th frequency band, α_k and β_k are weighting factors [1], $SNR_{eff\,k,i}$ is effective signal-to-noise ratio expressed in dB, i is modulation frequency number ($i = \overline{1,5}$), $y_k(t)$ is filtered signal $y(t)$ in the k th octave filter.

C. INDIRECT METHODS OF STI ESTIMATION

The basis of the indirect method is modulation transfer function [1]

$$m_{ki} = m_{k\,rev}(F_i) \cdot m_{k\,noise} = \frac{\int_0^\infty h_k^2(t)\exp(-j2\pi F_i t)dt}{\int_0^\infty h_k^2(t)dt} \cdot \left(1 + 10^{-0.1 \cdot SNR_k}\right)^{-1}, \quad (4)$$

$m_{k\,rev}(F_i)$ and $m_{k\,noise}$ $i = \overline{1,14}$, $k = \overline{1,7}$ are modulation transfer ratios of reverberation-distorted speech and modulation transfer coefficient of noise-distorted speech, respectively, $h_k(t)$ is the result of the filtering of RIR $h(t)$ with the k th octave filter, F_i is i th modulation frequency, SNR_k is SNR in the k th frequency band expressed in dB.

For IN1 method, a pair of signals emitted sequentially is used as a test signal. They are the first (MLS signal with a uniform spectrum) and the second (noise segment with a speech spectrum) parts of the multicomponent test signal shown in Fig. 1 and designed to estimate the modulation coefficients $m_{k\,rev}(F_i)$ and $m_{k\,noise}$, respectively.

Another variant of the indirect STI measurement method, recommended in [1], is IN2 method using a test signal in the form of a MLS signal with speech spectrum. The MLS signal is shown in Fig. 1 as the fourth part of the multicomponent test signal.

Each of the variants of the indirect method has its own characteristics. For example, an obvious advantage of IN1 is the ability to perform calculations in strict accordance with (4). The disadvantage is the need to use a test signal consisting of two components. This drawback is absent in the IN2 method, where the test signal is single-component. However, the disadvantage of IN2 is the distortion of the shape of the RIR $h(t)$ estimate, caused by the use of the MLS signal with the speech spectrum, which raises the question of the influence of the $m_{k\,rev}(F_i)$ estimation error on the STI assessment results.

Method IN1 is implemented according to (3)-(4), while the MLS signal $x_{MLS}(t)$ and stationary noise $x_{n\,spch}(t)$ with the speech spectrum are used for STI calculations. The maximum level of the side lobes of the autocorrelation function of the signal $x_{MLS}(t)$ lasting $L = 2^{18} - 1$ samples is minus 54 dB [12], which makes it possible to estimate the RIR $h(t)$ with high accuracy.

The IN2 method is also implemented according to (3)-(4), while the MLS signal $x_{MLS\,spch}(t)$ with the speech spectrum is used for STI calculations. The maximum side lobes level of the autocorrelation function of the signal $x_{MLS\,spch}(t)$ lasting $L = 2^{20} - 1$ samples is minus 40 dB [12]. Although the level of the side lobes is close to minus 45 dB at an interval of ±2 s in the vicinity of the maximum of the autocorrelation function, the main lobe of the autocorrelation function is extended and has a rather complex shape [12]. This means that the RIR estimate in the IN2 variant is inferior to the IN1 variant in terms of accuracy.

D. MODEL STUDIES

Signal recordings $y(t)$ distorted by reverberation and noise were simulated in model studies

$$y(t) = x(t) \otimes h(t) + n(t), \quad (5)$$

where $h(t)$ is the RIR, $n(t)$ is noise, $\otimes$ is the convolution symbol.

Two cases of interference were considered, namely, no reverberation (T60 = 0 s) and reverberation present (T60 = 0.8 s). The pink noise $n(t)$ model was used, as well as an estimate of the left channel RIR $h(t)$ of the fourth point of a university auditorium in Aachen (Germany) was used [14].

E. FIELD STUDIES

Field studies were carried out in auditorium 209 of building 12 of the National Technical University of Ukraine “Ihor Sikorsky Kyiv Polytechnic Institute” (Ukraine). The plan of the room and the location of the artificial head with two measuring microphones attached to it are shown in Fig. 3.

The measuring complex consisted of an active speaker Genius SP-HF 2.0 500 (Taiwan), a pair of measuring condenser microphones Superlux ECM-999 (Taiwan), an external sound card Steinberg UR242 (Germany-China) and a notebook.

The signal-to-noise ratio during field studies was 10-20 dB. The sampling frequency of the signals recorded from the microphone outputs was 44.1 kHz, with a bit depth of 24 bit.

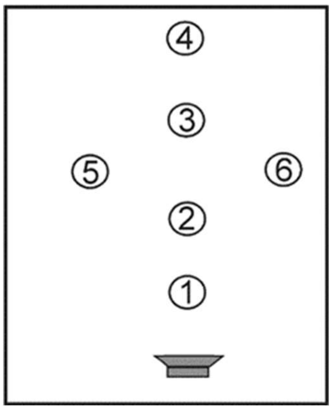


Figure 3. The room plan and locations of the artificial head

IV. RESEARCH REZULTS

The results of computer simulation based on model (5) are presented in Figs. 4-7. For each SNR value, 20 samples of the test signal (1) were generated, the average STI values and the corresponding 95% confidence intervals (vertical lines) were estimated (Fig. 4).

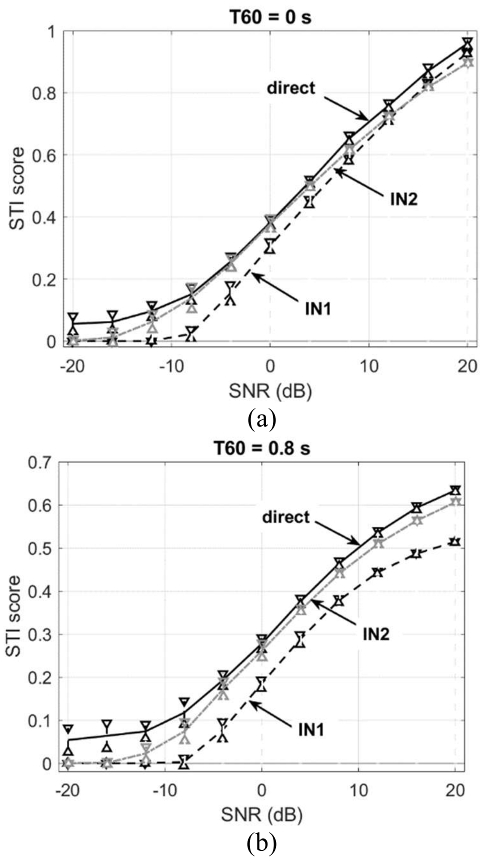


Figure 4. STI scores for noise (a) and noise plus reverberation (b)

Graphs of differences

$$\Delta_1 = STI_{IN1} - STI_{direct}\,, \quad \Delta_2 = STI_{IN2} - STI_{direct}\,, \quad (6)$$

for average values of STI estimates are presented in Fig. 5.

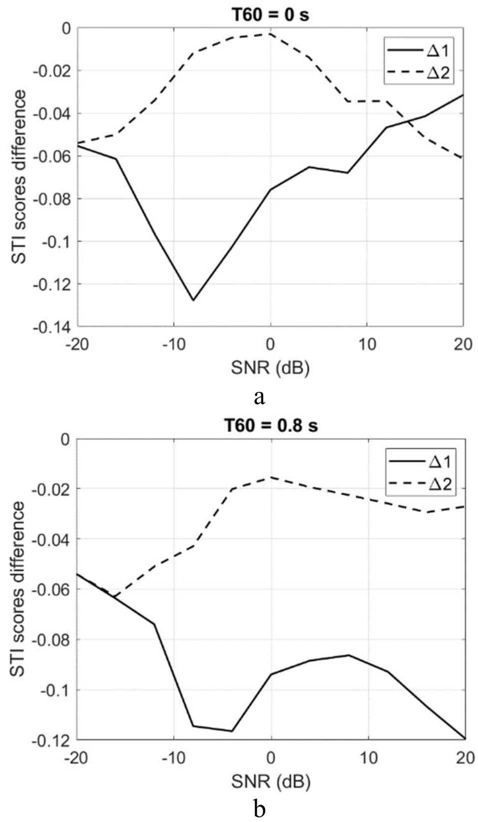


Figure 5. The difference in STI scores for noise (a) and noise plus reverberation (b)

It is accepted that the permissible error of STI estimation is the so-called just noticeable difference JND=0.03 [15]. Analyzing Fig. 4, it can be seen a somewhat strange behavior of the graphs for the direct method. These graphs for < -15 dB stabilize, not even reaching STI=0.05, although they should approach zero. This fact can be explained by the bias of the STI estimate, inherent in the direct RFM method at low SNR and at a finite duration of the test signal (1) [13]. Estimates of this bias dependence on the duration of the signal and the SNR for the case of the exclusive effect of noise are shown in Fig. 6.

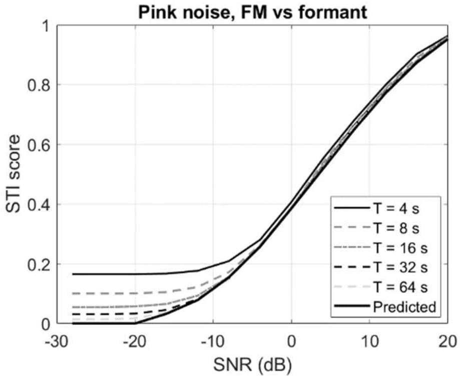
The bias was calculated relative to the predicted STI score obtained by the formant method [16], [17], [18]. The average STI scores for 100 samples, as well as the predicted values of STI are shown in Fig. 6a. The differences between averaged scores and predicted STI values are shown in Fig. 6b.

As can be seen in Fig. 6b, the bias does not exceed 0.03 at $T \geq 32$ s in a wide range of SNR values from minus 28 dB to plus 20 dB. At $T = 16$ s, the range of SNR values, at which the bias does not exceed 0.03, starts with minus 15 dB.

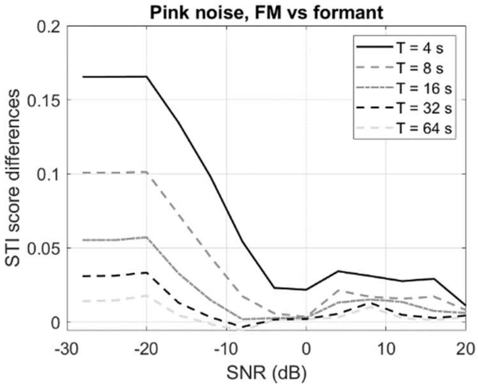
Graphs of differences

$$\delta_1 = STI_{IN1} - STI_{predict}\,, \quad \delta_2 = STI_{IN2} - STI_{predict}\,,$$

between averaged, over 20 samples, values of STI estimates obtained by indirect methods, and predictive estimates, is presented in Fig. 7.



(a)



(b)

Figure 6. Averaged and predicted STI scores (a) and corresponding differences (b)

Comparison of IN1 and IN2 estimates with predictive estimates (Fig. 6a and Fig. 7) allows obtaining not only more accurate, but also more optimistic results at very small (less than minus 10 dB) signal-to-noise ratios. For SNR>-10 dB, the comparison of IN1 and IN2 estimates with the scores of the direct RFM method leads to almost the same results as the comparison with predictive estimates.

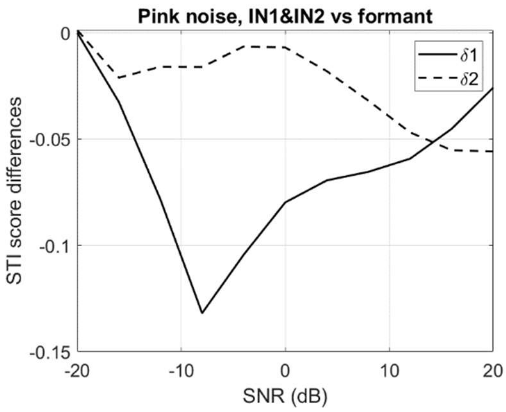
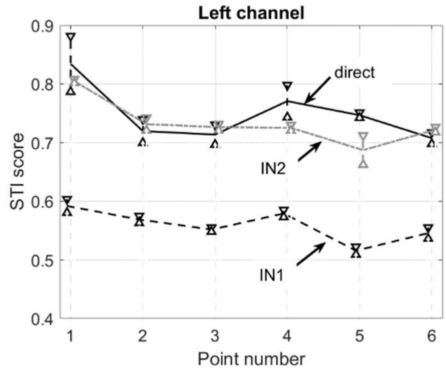


Figure 7. The difference between IN1, IN2 and predicted STI scores

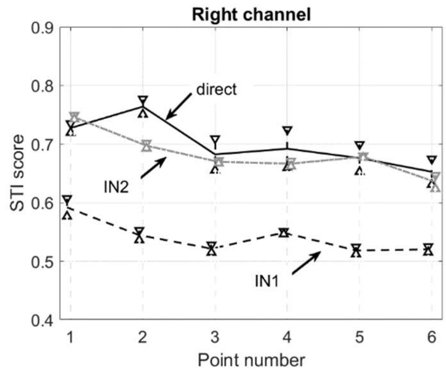
B. FIELD STUDIES

For each of the 6 points of the room (Fig. 3), five two-channel recordings were obtained using a multicomponent test signal (1). The results of the signals processing are presented in Fig. 8. The obtained results show that the STI estimates obtained by the IN2 method are little different from those for the direct method, while the estimates of the IN1 method are noticeably biased to a smaller value.

Graphs of differences (6) for the STI estimates averaged over 5 samples are presented in Fig. 9.

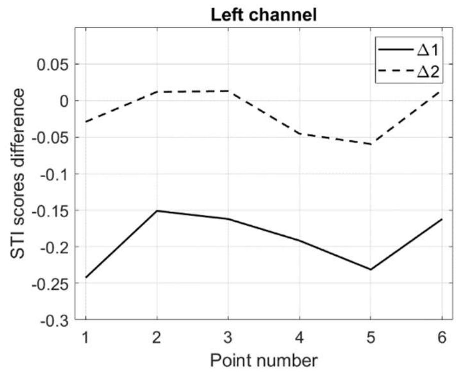


(a)

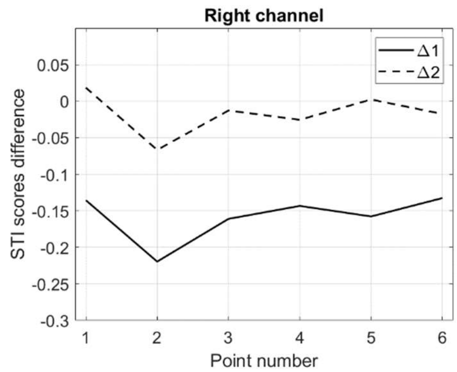


(b)

Figure 8. STI scores in room for left (a) and right (b) channels



(a)



(b)

Figure 9. The difference in STI scores in room for left (a) and right (b) channels

As can be seen from the graphs, the difference does not exceed 0.07 for the IN2 method, while for IN1 the difference can reach 0.24.

Comparing the results of field and model experiments, one can see that in the case of IN2 the results agree quite well and the difference Δ_2 did not exceed 0.06 in both cases. In the case of IN1, such agreement is absent. Indeed, if in the model experiment the difference Δ_1 did not exceed 0.13 (Fig. 5), in the field experiment this difference was in the range of 0.14-0.24 (Fig. 9). Such an increase in the difference can be explained by the inconsistency of the simple model (5) with the sound distortion in real room conditions.

V. DISCUSSION

In this paper, STI evaluation is performed by direct and indirect methods in the Matlab environment using the developed software, and a comparison of the obtained results is made. The advantage of this approach is the controllability of the software at the script level, which provides certain guarantees regarding the correctness of the obtained results.

The obtained study results show that, unlike IN2, the IN1 method leads to significantly underestimated results (by 0.13 in the model experiment and by 0.14-0.24 in the field experiment) compared to the direct method. The absence of an error in STI calculations by indirect methods IN1 and IN2 can be guaranteed by using the same software modules. This means that the resulting difference in STI scores is not caused by the difference in the software used, but by the difference in the estimation methods.

It should be noted that if in model studies the reverberation time was close to 0.8s, in the field experiment the reverberation time was somewhat lower. The values of T20 estimates of the reverberation time obtained for the room in Fig. 3 according to the results of processing the MLS signal and the MLS signal with the speech spectrum are shown in Fig. 10.

Averaged over all room points and across both channels, the value of T20 estimates for the MLS signal (used in IN1 method) is close to 0.6 s, while for the MLS signal with the speech spectrum (used in IN2 method) T20 value is significantly lower and close to 0.53 s. Since indoor speech intelligibility decreases with increasing reverberation time, these values of reverberation time estimates are in good agreement with the results of the model and field experiment, where lower intelligibility was obtained for IN1 method. This agreement can be considered as another proof of the correctness of the results obtained in this paper.

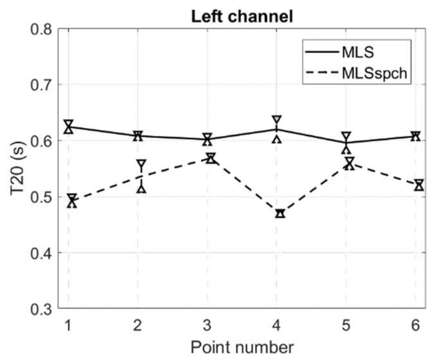
A similar agreement can also be seen when comparing the results of STI evaluation obtained from model and field studies. Indeed, during field studies, the integrated signal-to-noise ratio was 10-20 dB, and for estimates of T20 = 0.53-0.6 s, STI values were obtained in the range of 0.7-0.75 (Fig. 8). In model studies, the reverberation time was longer and close to 0.8 s, while for signal-to-noise ratios of 10-20 dB, STI estimates were smaller and were in the range of 0.5-0.6 (Fig. 4b).

For meeting rooms and offices, the reverberation time is usually shorter and close to 0.2-0.4 s, which is an intermediate case between the absence of reverberation and the reverberation time 0.8 s [14]. Therefore, the results presented in this paper can easily be extended to these cases as well. As for rooms with reverberation time T60>1 c, this case remains unexplored.

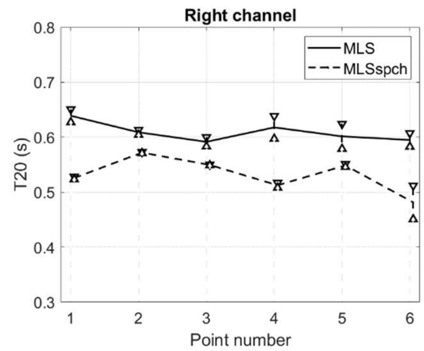
The influence of the long-term average speech spectrum (LTASS) shape on the STI measuring also remains unexplored. A number of papers aimed at evaluating the LTASS can be pointed [19], [20], [21]. It was noted in [19] that for the task of evaluating speech intelligibility it would be desirable to have a certain averaged speech spectrum, although at the same time it is obvious that there is a difference between the speech spectra of different languages and dialects, as well as between the speech spectra of men and women. It was proposed in [20] to use a new form of the English men speech spectrum instead of the one proposed in the standard [1]. Estimates of the spectrum of various languages, including the spectrum of Ukrainian speech, were obtained in [21]. At the same time, the influence of the shape of the speech spectrum on STI estimates was not investigated in the above-mentioned works. Thus, the specified direction of research also seems to be promising.

It should be noted that the issue of speech intelligibility assessment in classrooms is relevant and therefore has been considered quite often recently [22], [23].

At the same time, the effectiveness of STI prediction [24], as well as the effectiveness of STI assessment by analyzing the form of speech signals [25], [26] remains insufficiently studied.



(a)



(b)

Figure 10. Reverberation time T20 estimates for left (a) and right (b) channels

VI. CONCLUSIONS

In this paper, a comparison of STI estimates obtained by the direct and indirect method is made, while two variants of the indirect method are considered. A rapid formant-modulation method is used as a direct method of STI estimation. In the first version of the indirect method, called IN1, the test signal is a couple of MLS signal and a noise signal with a speech spectrum, while in the second version, called IN2, the test signal is single MLS signal with a speech spectrum. The comparison of STI estimates is made using computer simulation and by carrying out a field experiment. The same basic algorithms and computer programs are used for STI calculating in both variants of the indirect method. It is shown that for IN2 the average values of STI estimates are close to ones for the direct method and differ from them no more than 0.06 for SNR values from minus 20 dB to plus 20 dB. For IN1 method, this difference is significantly larger and can reach 0.24.

References

[1] British Standard BS EN 60268-16. *Sound system equipment. Part 16. Objective rating of speech intelligibility by speech transmission index*, 2011, 78 p. Available at https://resource.isvr.soton.ac.uk/staff/pubs/PubPDFs/BS%20EN%2060268-16.pdf

[2] ANSI S3.5-1969, *American National Standard, "Methods for Calculation of the Articulation Index,"* American National Standards Institute New York, 1969.

[3] Application Note, *Measuring Speech Intelligibility Using DIRAC — Type 7841*, Brüel & Kjær, 2013, 20 p. Available at: https://www.acoustics-engineering.com/html/dirac.html

[4] Acoustics Engineering, *Technical Note TN008 "DIRAC Stimuli"*, January 2008, 15 p. Available at https://www.acoustics-engineering.com/files/TN008.pdf

[5] A. Farina, *User Manual of Aurora 4.3*, Parma, Italy: University of Parma A/S; 2012, 69 p. Available at http://pcfarina.eng.unipr.it/aurora/download/Manual-HelpFile/Aurora43Manual.pdf

[6] D. Ponteggia AN-013, *Application note. Speech intelligibility assessment using CLIO 11*, 19 p. Available at: https://www.audiomatica.com/wp/wp-content/uploads/APPNOTE_013.pdf

[7] NTi Audio, *Application note. Speech Intelligibility. Measurement with the XL2 analyzer*, Dec., 2020, 28 p. Available at: https://www.nti-audio.com/en/

[8] P. Zhu, F. Mo, J. Kang, "Experimental comparison between direct and indirect measurement methods for the objective rating of speech intelligibility," *Proc. The 21st International Congress on Sound and Vibration (ICSV 21)*, Beijing, China, 13-17 July, 2014. Available at https://www.researchgate.net/publication/288781895_Experimental_comparison_between_direct_and_indirect_measurement_methods_for_the_objective_rating_of_speech_intelligibility

[9] P. Zhu, W. Tao, F. Mo, F. Guo, X. Lu, X. Liu, "Experimental comparison of speech transmission index measurement in natural sound rooms and auditoria," *Applied Acoustics*, vol. 165, pp. 1-21, 2020. https://doi.org/10.1016/j.apacoust.2020.107326.

[10] D. D'Orazio, E. Rossi, M. Garai, "Comparison of different in situ measurements techniques of intelligibility in an open-plan office," *Building Acoustics*, vol. 25, issue 2, pp. 111-122, 2018. https://doi.org/10.1177/1351010X18776431.

[11] J. Kotus, B. Kostek, A. Kurowski and P. Szczuko, "A comparison of STI measured by direct and indirect methods for interiors coupled with sound reinforcement systems," *Proceedings of the 2018 IEEE Joint Conference – Acoustics*, Ustka, Poland, 2018, pp. 1-6. https://doi.org/10.1109/ACOUSTICS.2018.8502277.

[12] A. Prodeus, O. Dvornyk, A. Naida, M. Didkovska, O. Grebin, "Multicomponent signal for comparing direct and indirect methods of speech transmittion index measurement," *Electronics and Control Systems*. vol. 75, no. 75, 2023, pp. 27-35. https://doi.org/10.18372/1990-5548.75.17546.

[13] A. Prodeus, "Rapid version of a formant-modulation method of speech intelligibility estimation," *Proceedings of the Perspective Technologies and Methods in MEMS Design*, Polyana, Ukraine, 2011, pp. 61-63. Available at: https://ieeexplore.ieee.org/document/5960269.

[14] M. Jeub, M. Schafer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," *Proceedings of the International Conference on Digital Signal Processing (DSP)*, Santorini, Greece, 2009. https://doi.org/10.1109/ICDSP.2009.5201259.

[15] J. Bradley, R. Reich, R.S. Norcross, "A just noticeable difference in C50 for speech," *Applied Acoustics*, vol. 58, pp. 99-108, 1999. https://doi.org/10.1016/S0003-682X(98)00075-9.

[16] A. Harasiuk, A. Prodeus, M. Myronov, V. Lozinsky, N. Vy, A. Darchuk, "Predictive estimation of speech intelligibility masked by noise interference using analytical modeling," *Microsystems, Electronics and Acoustics*, vol. 24, no. 5, pp. 48-55, 2019. https://doi.org/10.20535/2523-4455.2019.24.5.192893.

[17] V. Trushin, "The analysis of the formant method of speech intelligibility estimation as a method of performing indirect measurements," *Science Bulletin of the NSTU*, vol. 77, no. 4, 2019, pp. 135–146. https://doi.org/10.17212/1814-1196-2019-4-135-146.

[18] L. Galbrun, K. Kitapci, "Accuracy of speech transmission index predictions based on the reverberation time and signal-to-noise ratio," *Applied Acoustics*, vol. 81, pp. 1-14, 2014. https://doi.org/10.1016/j.apacoust.2014.02.001.

[19] D. Byrne, H. Dillon, K. Tran, "An international comparison of long-term average speech spectra," *J. Acoust. Soc. Am.*, vol. 96, issue 4, pp. 2108-2120, 1994. https://doi.org/10.1121/1.410152.

[20] L. Morales, G. Leembruggen, S. Dance, B. Shield, "A revised speech spectrum for STI calculations," *Applied Acoustics*, vol. 132, pp. 33-42, 2018. https://doi.org/10.1016/j.apacoust.2017.11.008.

[21] O. Pedchenko and S. Lunova, "Analysis of Ukrainian diagnostic articulation tables," *EUREKA Phys. Eng.*, vol. 1, no. 1, pp. 63–72, 2018. https://doi.org/10.21303/2461-4262.2018.00559.

[22] S. Di Loreto, F. Serpilli, V. Lori, and C. Di Perna, "Comparison between predictive and measurement methods of speech intelligibility for educational rooms of different sizes with and without HVAC systems," *Energies*, vol. 16, issue 6, 2719, 2023. https://doi.org/10.3390/en16062719.

[23] M. Engel, J. Herrmann, P. Zannin, "Assessment of the sound quality of classrooms through Speech Transmission Index (STI), Sound Definition (D50) and Reverberation Time (RT)," *Proceedings of the Forum Acusticum*, Lyon, France, December 7-11, 2020, pp. 2789-2792. Available at: https://hal.science/hal-03242464.

[24] Application Note, *ODEON Application Note – Calculation of Speech Transmission Index in rooms*, JHR, February 2014, 16 p. Available at: https://odeon.dk/pdf/Application_Note_SpeechTransmissionIndex.pdf.

[25] R. Coffeen, "Comparison of three methods for determining speech transmission index (STI)," *J. Acoust. Soc. Am.*, vol. 116, issue 4, p. 2638, 2004. https://doi.org/10.1121/1.4785524.

[26] K. Payton, L. Braida, "A method to determine the speech transmission index from speech waveforms," *J. Acoust. Soc. Am.*, vol. 106, issue 6, pp. 3637–3648, 1999. https://doi.org/10.1121/1.428216.

Arkadiy PRODEUS, received the BSc, MSc, PhD, and DSc degrees in electrical engineering from the Kyiv Polytechnic Institute (KPI), Kyiv, Ukraine, in 1970, 1972, 1982, and 2012, respectively. Currently is a professor at the Acoustic and Multimedia Electronic Systems Department, NTUU "Igor Sikorsky KPI". Current interests include digital signal processing, acoustic measure-

ments, pattern recognition, speech and music signals processing.

***Oleksandr DVORNYK** B.Sc., M.Sc., in Electronics and Acoustics Equipment (2017 and 2019 years respectively) in NTUU "Igor Sikorsky Kyiv Polytechnic Institute". Currently is a 4-year student of PhD program in electronics at the Department of Acoustic and Multimedia Electronic Systems of NTUU "Igor Sikorsky KPI". During university time, published scientific articles, organized elective about sound engineering and sound design.*

***Anton NAIDA** B.Sc., in Electronics and Acoustics Equipment (2022 year) in NTUU "Igor Sikorsky Kyiv Polytechnic Institute". Currently is a 2-year student of M.Sc. program in electronics at the Department of Acoustic and Multimedia Electronic Systems of NTUU "Igor Sikorsky KPI". During university time, published scientific articles, researched automatic security systems.*