UDC 621.391.83(045)
DOI:10.18372/1990-5548.75.16428

[1]**A. M. Prodeus,**
[2]**O. O. Dvornyk,**
[3]**A. S. Naida,**
[4]**M. V. Didkovska,**
[5]**O. P. Grebin**

MULTICOMPONENT SIGNAL FOR COMPARING DIRECT AND INDIRECT METHODS OF SPEECH TRANSMISSION INDEX MEASUREMENT

National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
E-mails: [1]aprodeus@gmail.com ORCID 0000-0001-7640-0850, [2]alexanderdvornyk@gmail.com, [3]naida.a.s.2001@gmail.com, [4]maryna.didkovska@gmail.com, [5]alexgstudio.2016@gmail.com

Abstract—*When evaluating the intelligibility of speech distorted by noise and reverberation, direct or indirect methods of measuring the speech transmission index are used. However, it remains insufficiently studied how significantly differ the results of measurements obtained by direct and indirect methods. To find an answer to this question, the use of a multicomponent test signal consisting of four "elementary" signals separated by pauses is proposed in this paper. As "elementary" signals, it is proposed to use a maximum-length sequence, a speech shaped maximum-length sequence, a speech shaped stationary noise, and a speech shaped amplitude-modulated noise. Use of amplitude-modulated noise allows estimating speech transmission index by a direct method. Other "elementary" signals make it possible to estimate speech transmission index by two variants of indirect method. The proposed algorithms and corresponding computer programs were tested on trial signal models, while the consistency of the obtained results with the results of previous studies was revealed. The results of the signal models studies show that both considered variants of the indirect speech transmission index measurement method lead to underestimated results compared to the direct method. For one of the variants of the indirect method, the value of the estimate bias is 0.03–0.04, regardless of the interfering conditions. For another variant of the indirect method, the estimate bias varies from 0.01 to 0.18, depending on the interference conditions.*

Index Terms—Test signal; speech intelligibility; direct method; indirect method; noise disturbance; reverberation.

I. INTRODUCTION

When using the indirect speech transmission index (STI) measurement method, BS EN 60268-16 standard [1] recommends using an exponential sweep signal or maximum-length sequence (MLS) as test signals. A certain preference is given to MLS signals with a uniform or speech shaped spectrum. This is explained by the fact that MLS signals are perceived by the ear as noise and therefore can be used in filled rooms. Another advantage of the MLS signal with the speech spectrum is the ability to simplify the assessment of the effects of both reverberation and background noise on speech intelligibility.

The specified recommendations of the standard [1] are implemented in a number of hardware and software applications, in particular, such as DIRAC [2], [3], AURORA [4], [5], CLIO [6], NTi Audio [7]. However, the following questions are relevant:

1) how significantly different are the results obtained by direct and indirect STI measurement methods?

2) how significantly different are the results of STI measurements obtained by the indirect method using MLS signals with a speech shaped spectrum from those for MLS signals with a uniform spectrum?

It is difficult to find a clear answer to these questions in the literature. In article [5], where the AURORA software application was used for acoustic measurements, a good agreement (the difference did not exceed 0.032) of the STI measurement results by the direct and indirect method in a wide range of signal-to-noise ratio (SNR) values from -5 dB to 20 dB was shown. However, although AURORA provides several ways to measure STI by an indirect method, it was not specified in [5] which method was chosen for experimental studies.

In article [8], the DIRAC software application was used to compare the results of STI measurements by direct and indirect methods. Exponential sweep signal was used in indirect STI measurements. According to the research results given in [8], STI estimates for the indirect method at high levels of background noise can exceed those for the direct method by 0.16. The shortcoming of the description of the research organization in [8] is that it is difficult to understand how the partial (in octave frequency bands) signal-to-noise ratios were estimated during STI measurements by the indirect method.

In article [9], the multicomponent signal consisting of three sweep signals (the instantaneous frequency varied linearly from 100 Hz to 10 kHz) of different durations and a signal in the form of amplitude-modulated noise was used. A test signal in the form of amplitude-modulated noise was used to estimate STI (version STIPA) by the direct method. All signal components were separated by pauses lasting 5 s. Measurements were made at different points in two rooms with a volume of 130,000 m^3 and 12,000 m^3. In the larger room, STI estimates obtained by the indirect method mostly exceeded those obtained by the direct method, with the largest excess being 0.17. In a smaller room, the difference in STI scores did not exceed 0.05. The shortcoming of [9] is the lack of information regarding the method of estimating partial signal-to-noise ratios during STI measurements by the indirect method.

In article [10], it is shown that the difference between experimental estimates of STI obtained by direct and indirect methods using the DIRAC, AURORA and NTi systems does not exceed 0.03. MLS with a pink spectrum, an exponential sweep signal, and a set of test signals for the implementation of the direct STI estimation method were used.

A common drawback of the above studies is the use of commercial hardware and software applications for STI assessment. The software in these applications is a trade secret product. Therefore, it cannot be excluded that the developers could programmatically compensate for the difference in the results of direct and indirect methods without notifying users. Another drawback is the lack of data on the difference in the results of STI measurements when using different variants of the indirect method. The existence of these variants is possible due to the fact that MLS signals with both uniform and speech shaped spectra can be used as test signals. Finally, given the somewhat contradictory results of the above studies, independent verification of these results is not superfluous. As for the review [11], its obvious drawback is excessive brevity, due to which the question of comparing the results of STI measurements by direct and indirect methods was not even considered. The object of this paper is the development of algorithms for the formation and processing of multicomponent test signals in the Matlab environment, which would allow to obtain "transparent" software and, based on it, to perform research aimed at eliminating the above mentioned shortcomings.

II. Problem Statement

According to the full version of the direct STI measurement method, a set of 14 non-stationary noises with a speech spectrum, the power of which varies according to a harmonic law, is used as test signals. Modulation frequencies of such noises vary from 0.63 Hz to 12.5 Hz [1]. A shortened version of the direct method is proposed in [12], where a segment of a random process with a duration of at least 16 s is used as a test signal

$$x_{AM\,spch}(t) = \xi(t)\sqrt{f(t)}\,, \tag{1}$$

$$f(t) = 1 + 0.32\left(\sin 2\pi\frac{t}{T_m} + \sin 2\pi\frac{2t}{T_m} + \sin 2\pi\frac{4t}{T_m} + \sin 2\pi\frac{8t}{T_m} + \sin 2\pi\frac{16t}{T_m}\right),$$

where $\xi(t)$ is speech shaped noise, $f(t)$ is the law of modulation of the noise variance with the "base" period T_m=1.43 s. When using such a modulation law, the modulation frequency range from 0.7 to 11.2 Hz is covered, which practically coincides with the modulation frequency range for the full version of the STI method [1].

According to the indirect STI measurement method [1], MLS signals are recommended to use as a test signals. They are, in particular, MLS signals with a uniform spectrum (variant 1 of the indirect method, which will be denoted as IN1) or filtered MLS signals, the spectrum of which is similar to the speech spectrum (variant 2 of the indirect method, hereafter IN2). Each of these versions, IN1 and IN2, has its advantages and disadvantages. Before pointing out these advantages and disadvantages, we note that the basis of the indirect method is an analytical expression [1]

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

where m_{ki}, $k=1,\ldots,7$, $i=1,\ldots,14$, are general modulation transfer coefficients, $m_{k\,rev}(F_i)$ are reverberation modulation transfer coefficients, $m_{k\,noise}$ are noise modulation transmission coefficients, $h_k(t)$ is the result of the room impulse response (RIR) $h(t)$ filtering by k th octave filter, F_i is i th modulation frequency, SNR_k is partial signal-to-noise ratio in dB for k th frequency band.

In the case of IN1, it is easy to calculate the first factor $m_{k\,rev}(F_i)$ of (2). Difficulties arise with the SNR_k evaluation, since the test signal does not contain information about the speech spectrum. One way to solve this problem is to use, in addition to the MLS signal, an additional noisy test signal with a speech spectrum, because the test signal does not contain information about the speech spectrum.

When using the IN2 variant, the signal-to-noise ratio $SNR_k, k = 1,\ldots,7$, in each of the seven octave frequency bands is easier for estimation. For this, it is sufficient to subject the recorded signal and background noise to spectral analysis using a comb of octave filters. However, in the case of IN2, the $h(t)$ estimate will be distorted due to the unevenness of the spectrum of the test signal. As a result, the $m_{k\,rev}(F_i)$ estimate will differ from the one obtained when using the MLS signal with a uniform spectrum. Thus, the question arises about the influence of the estimation error of the $m_{k\,rev}(F_i)$ value on the STI estimation results.

Although the authors of [5], [8], [9], [10] claim they have compared direct and indirect methods, it is not always clear which type of indirect method, IR1 or IR2, was implemented. In addition, attempts to compare the results of implementation of IR1 and IR2 variants are unknown. Finally, the authors of these papers used commercial software and hardware equipment for STI measurements, due to which the algorithms of the corresponding calculations are "hidden" from the researchers. Therefore, the purpose of the paper was developing of such test signals and algorithms for their processing, which would make it possible to make the comparison of STI estimates obtained by direct and indirect methods as transparent as possible.

III. SET UP OF THE STUDY

A. Variants of multicomponent test signals

To speed up and simplify the execution of model and experimental studies, it seems appropriate to use multi-component test signals consisting of a sequence of "elementary" test signals and pauses.

For example, three-component test signal “even spectrum MLS, speech spectrum noise, modulated speech spectrum noise” (Fig. 1a) is convenient for comparing the results of STI evaluation by the direct method and IN1. Indeed, after calculating the cross-correlation between the recorded room response and the MLS signal, a sharp high spike will be obtained at the beginning of the received array, which is convenient for determining the beginning of the first signal and the boundaries of other two signals against the intense background noise.

A two-component signal “speech shaped MLS, modulated speech spectrum noise” (Fig. 1b) is convenient for comparing the direct method with IR2.

However, proposed in this study the four-component test signal “even spectrum MLS, speech spectrum noise, modulated speech spectrum noise, speech shaped MLS” (Fig. 1c) is appears to be the most promising from the point of view of saving time and convenience of research.

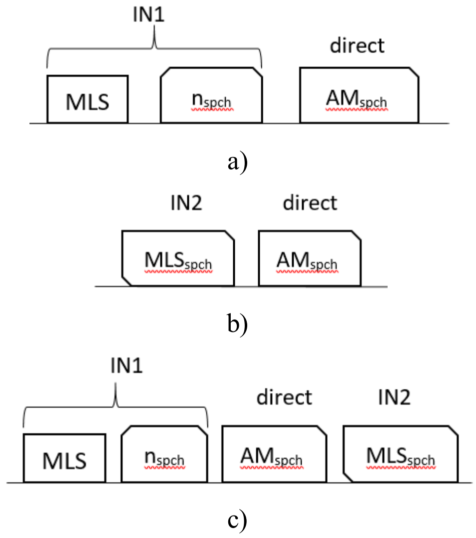


Fig. 1. Three-component (a), two-component (b) and four-component (c) test signals

This variant of the test signal significantly facilitates the end-to-end comparison of STI evaluation results by the direct method and both variants of the indirect method, IN1 and IN2. Moreover, when the individual components of the test signal are separated by relatively small pauses (lasting a few seconds), we can talk about conducting tests in practically unchanged conditions over time. Note that pauses between individual signals are required to estimate background noise parameters.

B. Choice of test signal parameters

The parameters of the test signal should be chosen, in particular, taking into account the expected reverberation time of the premises where speech intelligibility measurements are planned. If it is planned to carry out measurements in classrooms of educational universities and in offices, the reverberation time in the vast majority of cases does not exceed 1 s. Therefore, the duration of the pauses between the components of the test signal can be 2 s, which will ensure the presence of intervals lasting 1 s in the pauses, where the background noise interference prevails over the reverberation. Such intervals with background noise interference are needed both for estimating signal-to-noise ratios $SNR_k, k = 1,\ldots,7$ and for estimating the integral signal-to-noise ratio SNR.

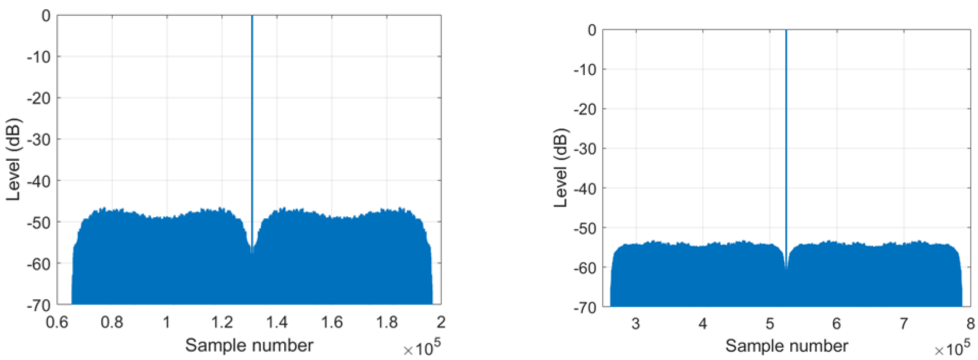
According to [1], the duration of modulated noise with the speech spectrum used when using the direct STIPA method should be 15–20 s. In our studies, an accelerated direct STI measurement method similar to the STIPA method [12] was used, according to which the duration of modulated noise with the speech spectrum should be at least 16 s.

The duration of an elementary test signal in the form of stationary noise with a speech spectrum in these studies is 4 s, which is sufficient to ensure the accuracy of STI measurements by IN1 method, close to the accuracy of the direct method under conditions of predominant noise interference [13].

The duration of MLS signals should be chosen taking into account the level of the side lobes of the autocorrelation function, which should preferably be no greater than minus 40 dB [1]. Graphs of the autocorrelation functions of MLS signals with a uniform spectrum are shown in Fig. 2.

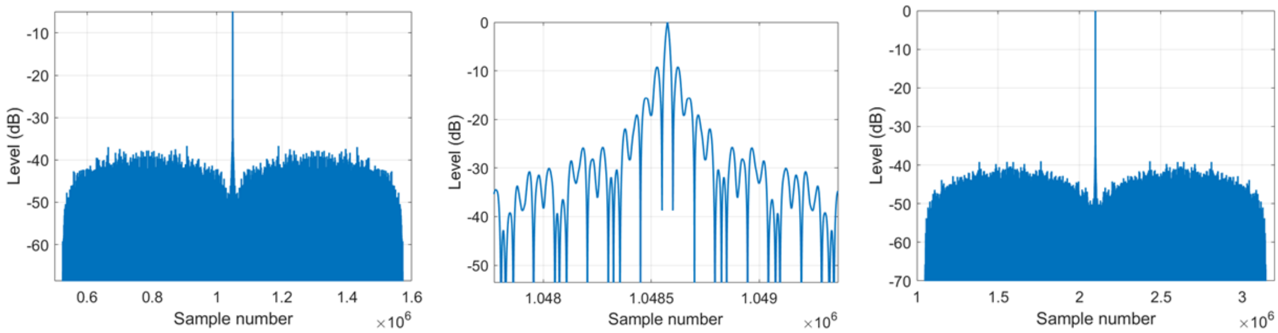
It can be seen that for the number of samples $L = 2^{16} - 1$ of the MLS signal, the maximum level of side lobes is close to -48 dB, while this level decreases to -54 dB for $L = 2^{18} - 1$, for the sampling frequency F_S = 44100 Hz. For research in this paper, $L = 2^{18} - 1$ was adopted, therefore the duration of the MLS signal with a uniform spectrum is close to 6 s.

For MLS with the speech shaped spectrum, the duration of the signals must be significantly increased due to the fact that the reduction of the frequency band leads to an increase in the level of the side lobes of the autocorrelation function. It is shown in Fig. 3 the graphs of the autocorrelation functions of MLS with the speech shaped spectrum.



b) b)

Fig. 2. Autocorrelation functions of the MLS signals for $L = 2^{16} - 1$ (a) and $L = 2^{18} - 1$ (b)



a) b) c)

Fig. 3. Autocorrelation functions of the speech shaped MLS signals for $L = 2^{19} - 1$ (a, b) and $L = 2^{20} - 1$ (c)

It can be seen that for the $L = 2^{19} - 1$, the maximum level of side lobes is close to -37 dB, although near the main peak (in the vicinity of ± 2 s) the level of side lobes is smaller and close to –40 dB. In the case $L = 2^{20} - 1$, the maximum level of side lobes decreases to –40 dB, and at an interval of ± 2 s in the vicinity of the maximum burst, the level of side lobes is close to -45 dB. In view of the specified levels of side lobes, $L = 2^{20} - 1$ for research in this work was chosen. In this case, the duration of the MLS signal with the speech shaped spectrum for the sampling frequency F_S = 44100 Hz is close to 24 s.

C. Test signal generation algorithm

The algorithm for generating a four-component test signal consists of the following stages:

1) generation of an MLS signal $x_{MLS}(t)$ with a uniform spectrum and T_{MLS} duration;

2) generation of the speech spectrum MLS signal $x_{MLS\,spch}(t)$ with $T_{MLS\,spch}$ duration;

3) generation of a segment $x_{n\,spch}(t)$ of Gaussian speech spectrum noise with $T_{n\,spch}$ duration ;

4) generation of amplitude-modulated speech spectrum noise $x_{AM\,spch}(t)$ with $T_{AM\,spch}$ duration;

5) equalization of variances of test signal components:

$$D[x_{MLS}(t)] = D[x_{MLS\,spch}(t)] = D[x_{n\,spch}(t)] = D[x_{AM\,spch}(t)];$$

6) formation of the final test signal by sequential addition (this procedure is indicated by the $\oplus$ symbol) of its individual components, with the insertion of pauses $x_{sil}(t)$ of the T_{sil} duration between these components:

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

7) normalization according to the maximum of the $x_{test}(t)$ signal and recording the result of normalization to computer disk.

Let's comment on some points of the given list. The formation of signals with a speech spectrum is proposed to be performed by filtering primary signals having a uniform spectrum with a comb of 7 octave filters.

The power equalization procedure (item 5 of the given list) is intended to facilitate the calibration of the measuring system. In the process of such calibration, the sound level emitted by the loudspeaker should be 60–70 dBA at a distance of 1 m from the loudspeaker. The power equalization allows for a correct comparison of the results obtained using different methods of measuring speech intelligibility.

The result of synthesis of a four-component test signal performed according to the above algorithm is shown in Fig. 4. As can be seen, the total duration of such a signal is about 60 seconds.

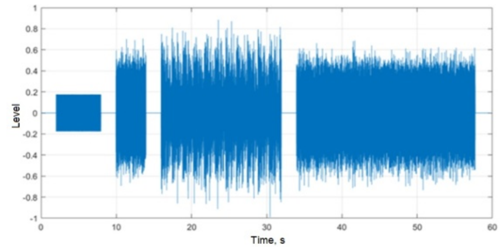


Fig. 4. Four-component test signal $x_{test}(t)$

IV. TRIAL COMPARISON OF STI MEASUREMENT METHODS

A. Processing the room reaction signal

The algorithm for processing the reaction $y(t)$ of the room to the test signal $x_{test}(t)$ consists of the following stages:

1) perform cross-correlation processing of the signal $y(t)$ and the MLS signal $x_{MLS}(t)$, as a result of which the signal $z(t)$ is obtained;

2) search for the maximum spike in the $z(t)$ signal, the position t_0 of which on the time axis indicates the beginning of the room's response to the MLS signal $x_{MLS}(t)$;

3) starting from the time value t_0, an estimate $\tilde{h}(t)$ of the room impulse response (RIR) of the duration of 1–1.5 s is cut out of the signal $z(t)$ for further calculation $m_{k\,rev}(F_i)$ according to (2);

4) focusing on the time value t_0, calculate the limits values of the three speech shaped signals (noise $y_{n\,spch}(t)$, modulated noise $y_{AM\,spch}(t)$ and MLS $y_{MLS\,spch}(t)$), after which arrays of relevant data are cut out of the signal $y(t)$ for further processing;

5) estimate STI by shortened direct method [12], using $y_{AM\,spch}(t)$;

6) estimate the STI by the indirect method of IN1 using expression (2), while the signal $\tilde{h}(t)$ is used for the $m_{k\,rev}(F_i)$ calculation, and the signal $y_{n\,spch}(t)$ is used for the $m_{k\,noise}$ calculation;

7) estimate the STI by the indirect method IN2 using expression (2), while the signal $\tilde{h}_{spch}(t)$ is used for the $m_{k\,rev}(F_i)$ calculation, where $\tilde{h}_{spch}(t)$ is the result of the cross-covariance of the signals $y_{MLS\,spch}(t)$ and $x_{MLS\,spch}(t)$, and the parameter $m_{k\,noise}$ is calculated using the signal $y_{MLS\,spch}(t)$.

B. Algorithms verification on signal models

To check the performance of the proposed algorithms for generating a test signal $x_{test}(t)$ and further processing the room's response $y(t)$ to this signal, four examples were considered, where the modeling of signals distorted by noise and reverberation took place according to the expression:

$$y(t) = x_{test} \otimes h(t) + n(t), \qquad (3)$$

where $n(t)$ is background noise disturbance.

These examples consider four situations:

1) reverberation and noise are absent;
2) reverberation is present, and noise is absent;
3) reverberation is absent, and noise is present;
4) noise and reverberation are present.

It was accepted $y(t) \approx x_{test}(t)$ in the first example. In the second and fourth examples, a record of the RIR $h(t)$ for a real room with a volume of 370 m^3 and a reverberation time of 0.8 s was used [14].

In the third and fourth examples, stationary white noise $n(t)$ with a normal distribution law was used.

The value of the integral signal-to-noise ratio SNR in the third example was calculated according to the expression

$$SNR = D_{x_{n\,spch}} / D_n, \qquad (4)$$

where $D_{x_{n\,spch}}$ is variance of the noise component with the speech spectrum, D_n is variance of background noise. In the fourth example, the SNR value was calculated similarly to (4) with the difference that convolution $x_{n\,spch}(t) \otimes h(t)$ was used instead of $x_{n\,spch}(t)$.

The results of the evaluation of STI and $E_k = SNR_k$ for the specified model examples are given in in Fig. 5.

As can be seen from the obtained results, both variants of the indirect STI measurement method lead to underestimated results compared to the direct method. The value of such a shift can reach 0.18 for IN1 and does not exceed 0.04 for IN2.

The smallest deviation of the IN1 estimate is observed for low noise levels and low reverberation time values (Fig. 5a). The largest deviation of the IN1 estimate can be expected in cases where the distorting effect of reverberation prevails over that of noise (Fig. 5b). As can be seen, the bias value of the IN1 estimate significantly depends on the interference situation.

As can be seen from the obtained results, both variants of the indirect STI measurement method lead to underestimated results compared to the direct method. The value of such a shift can reach 0.18 for IN1 and does not exceed 0.04 for IN2.

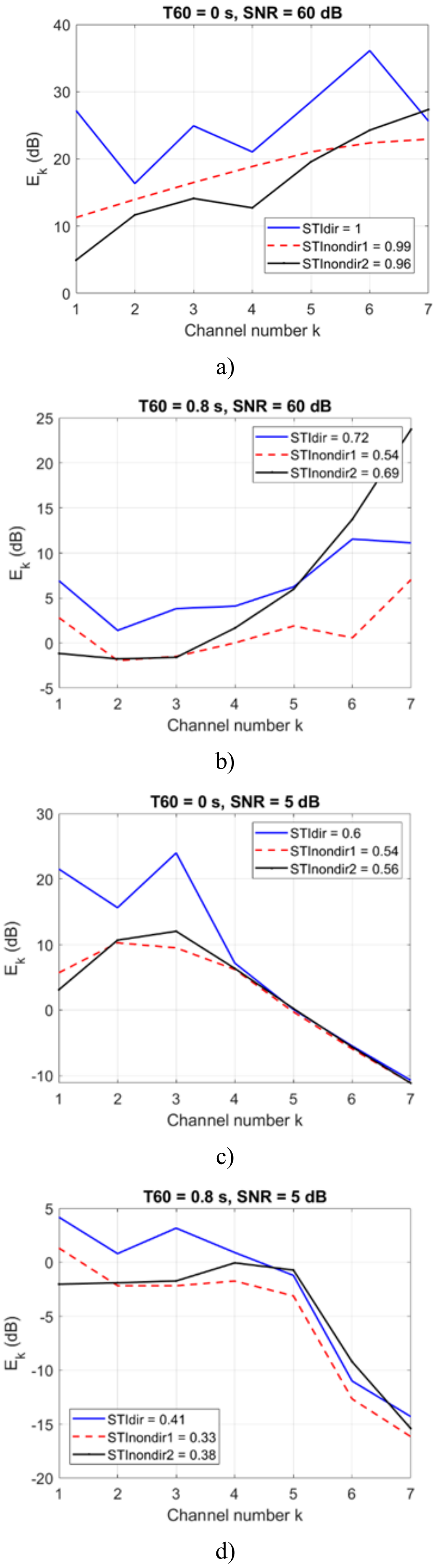


Fig. 5. STI and $E_k = SNR_k$ for examples 1 (a), 2 (b), 3 (c) and 4 (d)

The smallest deviation of the IN1 estimate is observed for low noise levels and low reverberation time values (Fig. 5a). The largest deviation of the IN1 estimate can be expected in cases where the distorting effect of reverberation prevails over that of noise (Fig. 5b). As can be seen, the bias value of the IN1 estimate significantly depends on the interference situation.

The priority of the IN2 estimate is, first, the bias insignificance. Second, the bias is practically independent of the interference situation.

The simulation results are in acceptable agreement with the research results [5], [8], [9], [10], which indicates the efficiency and correct functioning of the obtained software.

In the further study on similar model examples, it would be appropriate to try to optimize the parameters of the proposed test signal. In particular, since there are different opinions about the shape of the long-term speech spectrum [1], [15], [16], it would be useful to check the stability of the obtained results to a change in the shape of the long-term speech spectrum.

Finally, since model (3) is an approximation to the real effect of reverberation and noise on the speech signal, it is advisable to conduct experimental studies in real rooms.

V. CONCLUSIONS

It is proposed to use multicomponent test signals when comparing STI estimates by direct and indirect methods. It is shown that when considering two variants of the indirect method, IN1 and IN2, it is advisable to use a four-component test signal. Both the composition of such a signal and the parameters of its components are substantiated. Algorithms and computer programs for generating the test signal and its further processing have been developed. Performance testing of the developed algorithms and programs was performed on model examples.

As a result of such verification, it was found that both variants of the indirect STI measurement method lead to underestimated results compared to the direct method. The amount of underestimation can reach 0.18 for the IN1 variant and does not exceed 0.04 for IN2. The largest bias of the IN1 estimate should be expected in the case when the distorting effect of noise on the speech signal is weak compared to the effect of reverberation.

The difference between the STI estimates obtained by the direct method and the IN2 method is practically independent of noise and reverberation conditions. This makes it easy to adjust the STI estimate obtained by the IN2 method to match it with the estimate obtained by the direct method. For the estimation of STI obtained by the IN1 method, such reconciliation is more difficult to perform, as it is associated with the need to analyze the ratio of the contribution of noise and reverberation interference.

In the future, it is advisable to optimize the parameters of the test signal, as well as check the correctness of the obtained results by conducting experimental studies in real rooms.

REFERENCES

[1] British Standard BS EN 60268-16. Sound system equipment. Part 16. Objective rating of speech intelligibility by speech transmission index. 2011.

[2] Application Note, Measuring Speech Intelligibility Using DIRAC — Type 7841. Brüel & Kjær. 2013. Available at: https://www.bksv.com/media/doc/bo0521.pdf

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

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

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

[6] Ponteggia, AN-013, Application note. Speech intelligibility assessment using CLIO 11. 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] D'Orazio, E. Rossi, and M. Garai, "Comparison of different in situ measurements techniques of intelligibility in an open-plan office," *Building Acoustics.* 2018, 25(2), pp. 111–122. https://doi.org/10.1177/1351010X18776431

[9] 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," *Proc. 2018 Joint Conference–Acoustics, Ustka, Poland,* 2018, pp. 1–6. https://doi.org/10.1109/ACOUSTICS.2018.8502277

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

[11] H. Möller, "A review of STI measurements," *Forum Acusticum*, Dec 2020, Lyon, France. pp. 173–176.

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

[13] A. Prodeus, "Formant-Modulation Method of Speech Intelligibility Evaluation: Measuring and Exactness," *Proc. of the VII International Conference MEMSTECH 2011*. Lviv, Polyana, 2011, pp. 54–60. Available at: https://www.academia.edu/48296401/Formant_modulation_method_of_speech_intelligibility_evaluation_Measuring_and_exactness

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

[15] Birne, "An international comparison of long-term average speech spectra," *J. Acoust.Soc.Am.,* 96 (4), October 1994, pp. 2108–2120. https://doi.org/10.1121/1.410152

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

Received February 13, 2023

Prodeus Arkadiy. ORCID 0000-0001-7640-0850. Doctor of Engineering Sciences. Professor.
Department of Acoustic and Multimedia Electronic Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine (1972).
Research interests: digital signal processing.
Publications: 181.
E-mail: aprodeus@gmail.com

Dvornyk Oleksandr. Postgraduate student.
Acoustic and Multimedia Electronic Systems Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2019).
Research interests: digital signal processing.
Publications: 4.
E-mail: alexanderdvornyk@gmail.com

Naida Anton. Postgraduate student.
Acoustic and Multimedia Electronic Systems Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2022).
Research interests: digital signal processing.
Publications: 5.
E-mail: naida.a.s.2001@gmail.com

Didkovska Maryna. Candidate of Science (Engineering). Associate Professor.
Department of Mathematical Methods of System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2002).
Research interests: system analysis and management.
Publications: 78.
E-mail: maryna.didkovska@gmail.com

Grebin Oleksandr. Candidate of Science (Engineering). Associate Professor.
Acoustic and Multimedia Electronic Systems Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (1985).
Research interests: digital signal processing.
Publications: 44.
E-mail: alexgstudio.2016@gmail.com

А. М. Продеус, О. О. Дворник, А. С. Найда, М. В. Дідковська, О. П. Гребінь. Багатокомпонентний сигнал для порівняння прямого та непрямого методів вимірювання індексу передавання мовлення

Для оцінювання розбірливості мови, спотвореної шумом та реверберацією, використовують прямий або непрямий методи вимірювання індексу передавання мовлення. Проте залишається недостатньо вивченим, наскільки суттєво відрізняються результати вимірювань, отримані прямим та непрямим методами. Для пошуку відповіді на це питання в даній роботі пропонується використовувати тестовий сигнал, що складається з чотирьох «елементарних» сигналів, розділених паузами. В якості «елементарних» сигналів пропонується використовувати послідовність максимальної довжини, послідовність максимальної довжини із спектром мовлення, стаціонарний шум із спектром мовлення та амплітудно-модульований шум із спектром мовлення. Використання амплітудно-модульованого шуму дозволяє оцінити індекс передачі мовлення прямим методом. Інші «елементарні» сигнали дозволяють оцінити індекс передачі мовлення двома варіантами непрямого методу. Запропоновані алгоритми та відповідні комп'ютерні програми перевірено на модельних прикладах, при цьому виявлено узгодженість отриманих результатів з результатами попередніх досліджень. Результати пробних модельних досліджень свідчать, що обидва розглянуті варіанти непрямого методу вимірювання призводять до занижених результатів порівняно з прямим методом. Для одного з варіантів непрямого методу ця різниця становить 0,03–0,04 незалежно від завадових умов. Для іншого варіанту непрямого методу різниця варіюється від 0,01 до 0,18, в залежності від завадових умов.

Ключові слова: тестовий сигнал; розбірливість мовлення; прямий метод; непрямий метод; шумова завада; реверберація.

Продеус Аркадій Миколайович. ORCID 0000-0001-7640-0850. Доктор технічних наук. Професор.
Кафедра акустичних та мультимедійних електронних систем, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (1972).
Напрямок наукової діяльності: цифрова обробка сигналів.
Кількість публікацій: 181.
E-mail: aprodeus@gmail.com

Дворник Олександр Олександрович. Аспірант.
Кафедра акустичних та мультимедійних електронних систем, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (2019).
Напрямок наукової діяльності: цифрова обробка сигналів.
Publications: 4.
E-mail: alexanderdvornyk@gmail.com

Найда Антон Сергійович. Аспірант.
Кафедра акустичних та мультимедійних електронних систем, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна,, (2022).
Напрямок наукової діяльності: цифрова обробка сигналів.
Кількість публікацій: 5.
E-mail: naida.a.s.2001@gmail.com

Дідковська Марина Віталіївна. Кандидат технічних наук. Доцент.
Кафедра математичних методів системного аналізу, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (2002).
Напрямок наукової діяльності: системний аналіз та менеджмент.
Кількість публікацій: 78.
E-mail: maryna.didkovska@gmail.com

Гребінь Олександр Павлович. Кандидат технічних наук.
Кафедра акустичних та мультимедійних електронних систем, Національний технічний університет України «Київський політехнічний інститут ім. І.Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (1985).
Напрямок наукової діяльності: цифрова обробка сигналів.
Кількість публікацій: 44.
E-mail: alexgstudio.2016@gmail.com