Journal of Theoretical and Applied Information Technology
31st August 2020. Vol.98. No 16

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

SPEECH INFORMATION SECURITY ASSESSING IN CASE OF COMBINED MASKING SIGNALS

[1] YERZHAN N. SEITKULOV, [1] SEILKHAN N. BORANBAYEV, [1]NURLAN N. TASHATOV

[2]HENADZI V. DAVYDAU, [2] ALEKSANDR V. PATAPOVICH

[1] L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan

[2] Belarusian state university of informatics and radioelectronics, Minsk, Belarus

E-mail: [1] Seitkulov_y@enu.kz, [2] nil53@bsuir.edu.by

ABSTRACT

The paper presents the results of experimental studies directed to the speech intelligibility assessment while protecting it from leakage via acoustic channels by masking with combined acoustic signals, including “white” noise and speech-like signals. The speech information security assessment is closely related to environmental conditions, as a rule, the most common option is a certain room with various soundproofing properties of enclosing structural elements. It is possible to take into account the influence of soundproofing properties of building envelopes only on the basis of experimental measurements of the transfer characteristic of speech signals outside the premises and the additional protection space. It necessary to note that the results of the speech information security assessment also strongly depend on the methodology for conducting experimental studies. An important phenomenon in experimental studies is the resonance of the bending vibrations of the enclosing elements of the building structures. Difficulties in the speech information security assessment are caused by uncertainties associated with difficulties in the mathematical formulation of the protection problem on the one hand and a large number of factors affecting the speech information security, on the other hand. In the paper, it’s proposed to assess the speech information security by the determining of the speech intelligibility indicators in the limit states. For the correct assessment of the speech information security in terms of its intelligibility, a number of assumptions and limitations were adopted, which are based on practical experience and experimental studies of the speech information protection. The obtained results can be used to develop standards for the speech information protection from leakage via the acoustic channel.

Keywords: *Information Security, Speech Intelligibility; Combined Masking Signals.*

1. INTRODUCTION

Speech information is the primary information source, which is further reflected in the form of documents (projects, decisions and other forms). Therefore, the protection of voice information is the primary and most important element in the information security system. If at the initial stage of information appearance no measures have been taken to protect it, subsequent measures to protect it, up to cryptography, may not be effective.

Speech information security assessing is usually carried out according to the indicator of speech verbal intelligibility at the border of the protected area and is expressed in relation to the number of correctly recognized words to the total number of words contained in the information [1]. The protected area boundaries are usually enclosing structures of the premises. Each specific case of speech intelligibility assessing at the border of a protected area is a certain room with its own design, structural construction, form, materials and architectural construction. Therefore, it is not possible to use analytical expressions to find the speech signal transfer characteristic to the border of the protected area due to the wide variety of structural constructions and consideration of the acoustic properties of the used materials.

In practice, to assess the voice information security, the experimental and calculation method is widely used. The ratio of the sound pressure of the speech information signal to the sound pressure of speech-masking signals at the border of the security area is determined experimentally, since it is not possible to do this with sufficient accuracy (±3 dB)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

by calculation. In Russia, to assess the voice information security, the Sprut 7 and Sprut 11M software and hardware systems are used, the work of which is based on the theoretical principles [2–5]. Theoretical principles are based on the formant theory of speech intelligibility and suggest finding an integral index of articulation. In this case, the integral index of articulation is determined taking into account the weight coefficients for each of the 1/3 octave frequency bands as the sum of the products of the formant perception coefficient by the human hearing instrument and the weight coefficient for this frequency band. In a number of other countries, the geometric mean value of sound pressure ratios of the speech information signal to the sound pressure of speech masking signals for each of the 1/3 octave frequency band is used as an articulation index. Further, to evaluate the speech information security by the verbal intelligibility indicator, we use the experimentally obtained dependences of verbal speech intelligibility on the integral index of articulation or on the index of articulation SPI.

Thus, the voice information security is currently being evaluated. A significant difference between the proposed work is that verbal intelligibility is determined for combined speech-masking signals, when the required speech signal / masking noise ratio is significantly lower than for masking signals in the form of "white" or "pink" noise. Combined masking signals consist of "white" noise and speech-like signals formed taking into account the phonetic features of a given language [6–8]. The second significant difference is that the experimental dependence of verbal speech intelligibility on the speech signal / masking noise ratio was obtained for auditors with increased auditory sensitivity and specially trained speech recognition in the presence of noise.

2. METHODOLOGY FOR THE EXPERIMENTAL EVALUATION OF THE SPEECH SIGNALS TRANSFER FUNCTION

To assess the speech information security from leakage via the acoustic channel, it is necessary to know the speech signal / masking noise ratio at the boundary of the protected area. In the experimental evaluation of the transfer function, the speech-modeling signal in the Sprut 7 hardware-software complex can use "white" or "pink" noise with the possibility of controlling the spectral envelope or a harmonic signal in the frequency range from 1 to 20,000 Hz. However, the results of the experiment will differ from those that would have been obtained using a speech signal. This is due to the fact that the speech information transmission outside the premises through enclosing structural elements occurs due to the bending resonant vibrations of the latter, the excitation of which in the case of a speech signal takes 10–20 ms (this is about 10 periods of voiced speech sounds). In the case of "white" or "pink" noise using, the resonance oscillations are excited by spectral components equal to the resonant frequencies in terms of inferior amplitudes to the corresponding formants of the speech signal. Therefore, it is recommended to use speech-like signals formed on the basis of allophones and taking into account the phonetic features of a given language as a speech signal for experimental studies, and the amplitude spectrum of speech can be approximated for a frequency range from 100 to 8000 Hz and for an integral sound pressure level of $6.3\cdot10^{-3}$ Pa to 0.36 Pa (50 to 85 dB) by the expression

$$S(f)=\rho\cdot P\cdot K\left[\frac{1}{\rho+\left|(f_0-f)\right|}+\frac{1}{\rho+(f_0+f)}\right], \qquad (1)$$

where $\rho = 80$ Hz;

P – the sound pressure level of speech in the frequency band from 100 to 8000 Hz expressed in Pa;

$f_0 = 225$ Hz;

K – proportionality coefficient.

Coefficient ρ is close to the value of the frequency of the fundamental tone. The proportionality coefficient has a dimension of $s^{-1/2}$ and is equal to 0.0585 $s^{-1/2}$. The first term, taking into account the coefficient $\rho = 80$ Hz, characterizes the severity of the maximum in the speech spectrum. The second term characterizes the decrease in the spectral density of the sound pressure of speech with increasing frequency.

The spectral density of the speech signal at the places of assessing speech intelligibility outside the premises is determined from the expression

$$S_r(f)=\rho\cdot P\cdot K\left[\frac{1}{\rho+\left|(f_0-f)\right|}+\frac{1}{\rho+(f_0+f)}\right]\cdot K_r(f), \qquad (2)$$

where $K_r(f)$ – speech transmission coefficient as a function of frequency (transfer function for speech signals).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

If it is not difficult to implement a source modeling a speech signal in practice, there are uncertainties with the measurement of the speech signal at the boundary of the protected area. This is due to the fact that with a change in the position of the measuring sensor at the boundary of the protected area, the measurement results will also change. If the boundary of the protected area coincides with the boundaries of the building envelope, it is recommended to use an acoustic signal receiver with an antenna acoustic array to find areas with the maximum transfer characteristic value [9]. The differences in the amplitudes of vibrations on the surfaces of the building envelope are due to their bending vibrations during the transmission of a speech signal. These areas will be characterized as the most dangerous from the point of view of protecting voice information, and if the security assessment is determined with respect to these areas, then protection will be performed according to the limit states.

The block diagram of the experimental setup is shown in Figure 1.

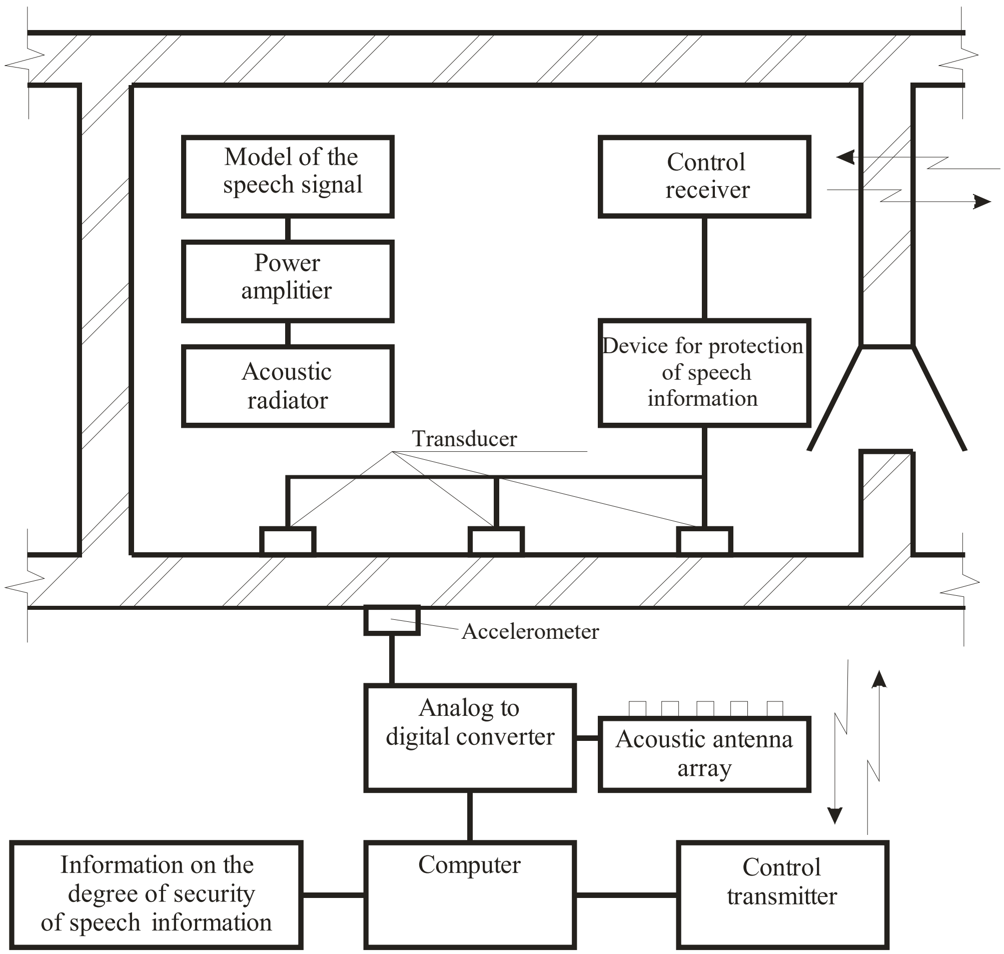


Figure 1. The block diagram of the experimental setup.

The "Model of the Speech Signal" block forms a speech-like signal on the basis of allophones with a spectrum whose envelope for speech is given by expression (1). This unit is controlled from a personal computer, to which a control transmitter is connected via a radio channel to a control receiver. The signal generated by the block "Model of the Speech Signal" is amplified by a power amplifier and fed to an acoustic emitter, which creates a given level of sound pressure in the room. The control of the sound pressure level in the room when modeling security for different speakers with different voice strengths is performed by changing the parameter *P* in expression (1). To ensure reliable reception of the speech-like signal, its level can be increased by 20 dB, and then take this into account when determining the ratio of the speech signal / masking noise. The sound pressure level outside the room at the border of the protected area is measured using an acoustic antenna array and transmitted through an analog-to-digital converter to a personal computer. At the same time, the sound pressure level outside the room at the border of the protected area consists of the sum of the signal simulating speech and acoustic industrial noise passed through the building envelope. In the same area of space, the level of production acoustic noise is measured separately in the absence of an acoustic speech-modeling signal.

Using a library of software tools for calculating fast Fourier transforms, a personal computer calculates the spectral density of incoming signals and finds the sound pressure levels of acoustic signals in 1/3 octave frequency bands with geometric mean frequencies 160, 200, 250, 315, 400, 500, 630, 800, 1000, 1250, 1600, 2000, 2500, 3160, 4000, 5000 Hz. The noise levels for each of the 1/3 octave frequency band are subtracted from the sum of the speech-simulating signal and production noise and the signal-simulating speech levels are obtained at the boundary of the protected area. Next, the integral sound pressure level of the signal simulating speech is calculated at the boundary of the protected area for the frequency range limited to 16 1/3 octave frequency bands. The permissible intelligibility of speech determines the integral level of masking effects.

3. SPEECH INFORMATION SECURITY ASSESSING

Speech information security assessing is proposed to be carried out according to the indicator of verbal intelligibility of speech in a number of papers. This is due to the fact that even one word carries a semantic meaning and thereby can reduce the security of speech information. Complete legibility of a phrase is the legibility of all words in a phrase. When this happens, the security of speech information can be significantly reduced, since the semantic content of speech information becomes more understandable.

To assess verbal speech intelligibility, it is proposed to use an articulation index defined for a frequency range with 16 1/3 octave frequency bands and geometric mean frequency bands from 160 to 5000 Hz from the expression

$$SPI = \sum_{f=160}^{5000} \left[L_{ts}(f) - L_n(f) \right] / 16, \quad (3)$$

where the sum is for each of 1/3 octave bands with an average frequency f;

$L_{ts}(f)$ – the transmitted level of the speech signal to the border of the protected area;

$L_n(f)$ – level of external noise and noise masking the speech signal at the boundary of the protected area.

If the difference in square brackets is less than –30 dB, these values are canceled and not taken into account. The articulation index is associated with the speech verbal intelligibility by addiction, obtained through experimental studies for a given language. The articulation index is the logarithm of the geometric mean signal-to-noise ratio in 1/3 octave bands. When this ratio is less than 0.03, it has been proposed to exclude this value.

The use of the formant theory of speech intelligibility and methods based on it [2–4], as indicated in a number of papers [10–13], can lead to rather large errors, therefore, the paper uses a simple method for evaluating verbal speech intelligibility based on experimental studies and integral signal-to-noise ratio.

The experimental studies of the speech spectral density made it possible to approximate its dependence on frequency and take into account the phonetic features of speech in assessing the speech information security using the formant method. This allowed us to create the speech signal model in the form of noise with the envelope of the speech spectrum.

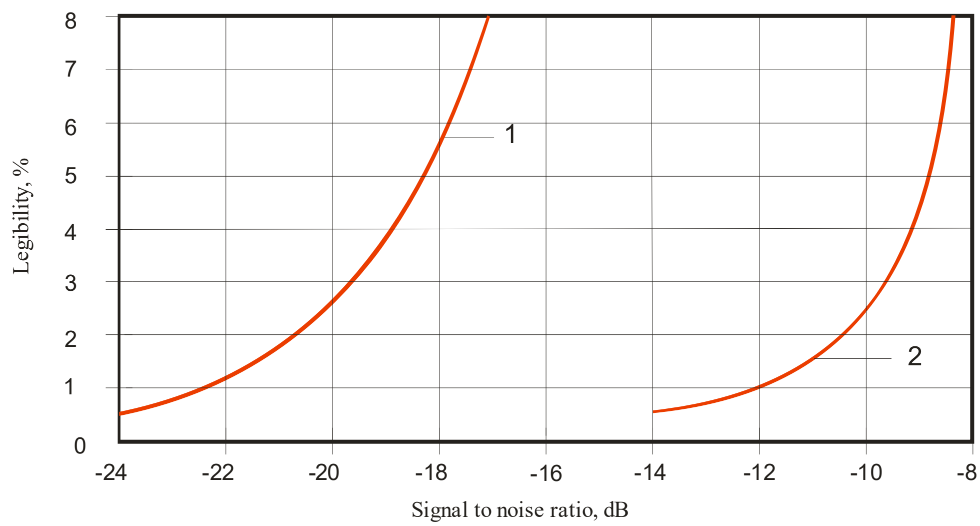
For experimental studies, combined masking signals were formed and they were superimposed on the phonograms of informational speech signals in the form of a coherent text lasting about 2 to 3 minutes and a volume of 198 to 202 words. Phonograms were voiced by selected and trained speakers. In this case, the signal-to-combined masking noise ratios were –14, –12, –10, –9 dB.

To assess the speech information security using combined masking signals, tests were conducted, for which 5 auditors were selected aged 20 to 30 years and with a differential sensitivity of hearing to a change in sound frequency of not more than 5 Hz at a frequency of 1000 Hz and high differential auditory sensitivity i.e. the ability to perceive changes in sound intensity from 0.5 to 0.9 dB according to Luscher. The measurements were carried out at an average sound intensity of 40

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

dB above the threshold of audibility and for each of the frequencies 500, 1000, 2000, 4000 Hz.

The results of experimental studies of verbal intelligibility by trained auditors by repeatedly listening to phonograms by various auditors are presented in Figure 2. At the same time, speech was masked by combined signals and white noise.



1 – for a masking signal in the form of “white” noise
2 – for combined masking signals

Figure 2. The dependence of intelligibility of speech on the signal-to-noise ratio for the entire frequency range.

To automate the calculation of the speech verbal intelligibility of from the integral signal-to-noise ratio for the masking signal in the form of “white” noise, the dependence 1 was approximated (Figure 2) by the expression

$$R = 6{,}9 \cdot SN^2 - 0{,}2 \cdot SN \, , \qquad (4)$$

where R – verbal intelligibility in relative units (not percentage);

SN – the integral signal-to-noise ratio in relative units (not in dB), which is valid for the SN range from 0 to 0.15.

For combined masking signals, verbal speech intelligibility can be determined from the expression

$$R = e^{-8{,}9 + 17{,}4 \cdot SN} \, , \qquad (5)$$

where R – verbal intelligibility in relative units (not in percent);

SN – the integral signal-to-noise ratio in relative units (not in dB), which is valid for the SN range from 0 to 0.35.

The parameter SN, used in expressions (4) and (5), characterizes the integral signal-to-noise ratio in relative units, and not in dB. The SPI articulation index is the logarithm of the geometric mean signal-to-noise ratio and is not used in the methodology presented, although it is close to the integral signal-to-noise ratio by its values. Table 1 presents the results of the work on the formation of combined speech masking signals and shows the distribution of their levels in 1/3 octave frequency bands.

Table 1 – Distribution of levels of masking signals with speech-like signals in the Kazakh language in 1/3 octave frequency bands

Geometrical average frequency 1/3 octave frequency band, Hz	Sound pressure levels in 1/3 octave frequency bands, dB			
	For white noise, dB	For speech-like signals, dB	For white noise and speech-like signals, dB	For a speech signal (at signal to noise ratio –10 dB), dB
160	48.14	52.2	53.6	49.2
200	49.14	55.9	56.7	52.9
250	50.14	57.5	58.2	54.5
315	51.14	57.1	58.1	54.1
400	52.14	54.7	56.6	51.7
500	53.14	52.6	55.9	49.6
630	54.14	50.6	55.7	47.6
800	55.14	48.9	56.1	45.9
1000	56.14	47.5	56.7	44.5
1250	57.14	46.2	57.5	43.2
1600	58.14	45.0	58.3	42.0
2000	59.14	43.9	59.3	40.9
2500	60.14	42.9	60.2	39.9
3150	61.14	41.9	61.2	38.9
4000	62.14	40.9	62.2	37.9
5000	63.14	39.9	63.2	36.9
141–5612	70.00	64.0	71.0	61.0

The data presented in table 1 are shown in Figure 3 in the form of dependences of sound pressure on frequency for clarity.

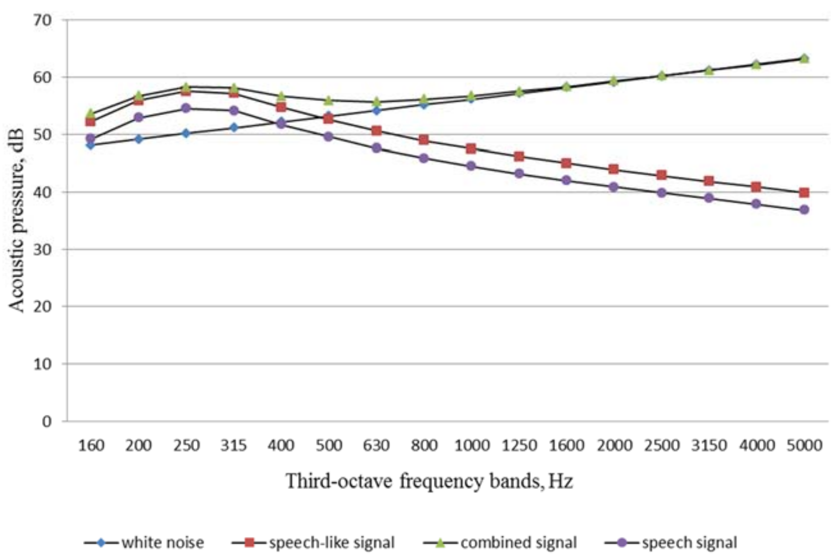


Figure 3. Distribution of masking signal levels in 1/3 octave frequency bands.

Journal of Theoretical and Applied Information Technology
31st August 2020. Vol.98. No 16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

As an information signal simulating a speech signal, a noise signal with an envelope corresponding to the spectrum of speech is used. When the acoustic vibrational information signal is excited by the bending vibrations of the building envelope and this information is transmitted outside the premises, the speech information security assessment in this case is determined by the ratio of the level of vibration accelerations caused by the acoustic speech signal and the level of vibration accelerations caused by industrial noise. As well as in acoustic calculations, the level of acoustic, speech-modeling, influence can be increased by 20 dB, and the level of acceleration of vibrations caused by this effect should be reduced by 20 dB when converted to the sound pressure level of the modeling signal of 70 dB.

The technique is based on experimental measurements of the information signal-masking noise ratios in the places where the intruder is possible. Using the obtained signal-to-noise ratio, speech intelligibility is determined using a graphical, experimentally obtained dependence of verbal intelligibility on the integral signal-to-noise ratio for the frequency range of a speech signal.

4. SOFTWARE TOOL

To automate the process of speech intelligibility assessing as an indicator of protection degree of the speech information masked by combined signals, the software product “Automatization001” has been developed. The work of the software product is carried out in accordance with the structural diagram shown in Figure 1, and the main part of the code for controlling the automation process is given below:

```
namespace Automatization001
{
   public partial class Form1 : Form
   {
      public String WavFileName1, WavFileName2, WavFileName3;
      public String XlsFileName1, XlsFileName2, XlsFileName3;
      public Form1()
      {
         InitializeComponent();
      }
      private void Form1_Load(object sender, EventArgs e)
            {
private void comboBox1_SelectedIndexChanged(object sender, EventArgs e)
      {
         switch (comboBox1.SelectedIndex)
         {
            case 0:
               {
                  OpenFileDialog openFileDialog = new OpenFileDialog();
                  openFileDialog.InitialDirectory = "d:\\";
                  openFileDialog.Filter = "Speech files (*.wav)|*.wav | All files (*.*)|*.*";
                  if (openFileDialog.ShowDialog() == DialogResult.OK)
                     WavFileName1 = openFileDialog.FileName;
               }
               break;
            case 1:
               XlsFileName1 = ProcessExcelFile(chart1, 2);
               break;
            case 2:
                     break;
}
      }
      private static string ProcessExcelFile(System.Windows.Forms.DataVisualization.Charting.Chart chart,
int columnIndex)
      {
```

```
    OpenFileDialog openFileDialog = new OpenFileDialog();
    openFileDialog.InitialDirectory = "d:\\";
    openFileDialog.Filter = "Excel files (*.xls)|*.xls | All files (*.*)|*.*";
    if (openFileDialog.ShowDialog() != DialogResult.OK)
      return null;
    string fileName = openFileDialog.FileName;
    var xlApp = new Excel.Application();
    var xlWorkbook = xlApp.Workbooks.Open(fileName);
    var xlWorksheet = xlWorkbook.Sheets[1];
    var xlRange = xlWorksheet.UsedRange;
    int rowCount = xlRange.Rows.Count;
    chart.Series[0].Points.Clear();
    for (int rowIndex = 1; rowIndex <= rowCount; ++rowIndex)
    {
      try
      {
       string xVal = xlRange.Cells[rowIndex, 1].Value2.ToString();
        double yVal = (double)xlRange.Cells[rowIndex, columnIndex].Value2;
              chart.Series[0].Points.AddXY(xVal, yVal);
}
      catch (Exception exx)
      {
        MessageBox.Show("Error in line" + rowIndex + ": " + exx.Message, "Error");
        fileName = null;
      }
    }
    xlWorkbook.Close();
    return fileName;
  }
  private void comboBox2_SelectedIndexChanged(object sender, EventArgs e)
  {
    switch (comboBox2.SelectedIndex)
    {
      case 0:penFileDialog();
          openFileDialog.InitialDirectory = "d:\\
        {
          OpenFileDialog openFileDialog = new O";
          openFileDialog.Filter = "Звуковые файлы (*.wav)|*.wav|Все файлы (*.*)|*.*";
          if (openFileDialog.ShowDialog() == DialogResult.OK)
            WavFileName2 = openFileDialog.FileName;
}
        break;
      case 1:
        XlsFileName2 = ProcessExcelFile(chart2, 5);
        break;
      case 2:
        break;
    }
        }
```

Figure 4 shows the appearance of the software product main window through which the choice of the noise data source is made: from wav. file; from xls. file from sound level meter or manual data entry into a given table.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

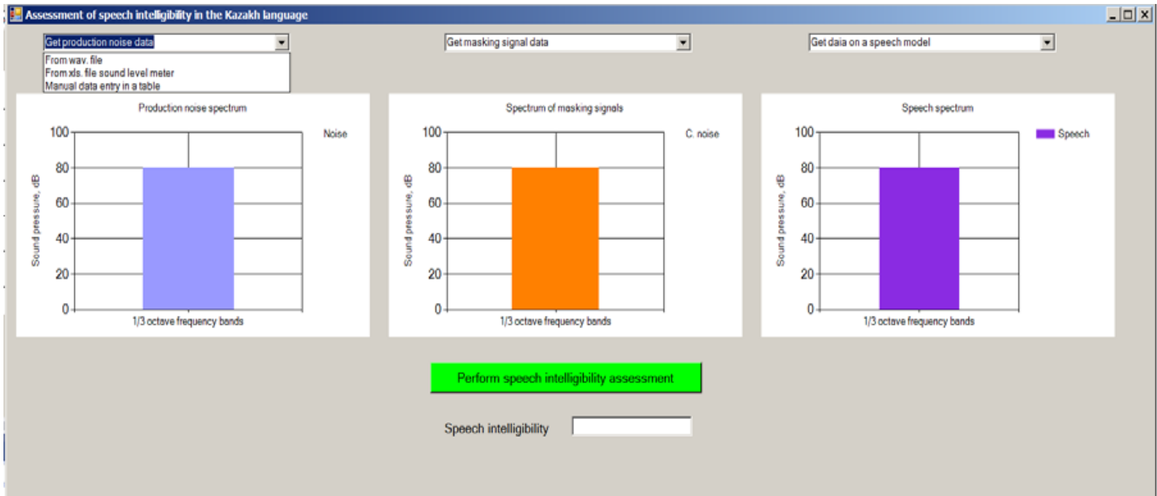


Figure 4. The appearance of the software product main window.

The data for combined masking signals and an informational speech signal at the boundary of the security zone is obtained in the same way. The received data is displayed in the main window in the form of histograms for 16th 1/3 octave frequency bands in the range from 140 to 5600 Hz with geometric mean frequencies of 160, 200, 250, 315, 400, 500, 630, 800, 1000, 1250, 1600, 2000, 2500, 3160, 4000, 5000 Hz.

The speech intelligibility calculation is performed in accordance with expression (5) for the integral ratio of the speech signal / combined masking noise with values at the boundary of the security zone. The calculation program code is given below:

```
//Variable declaration;
        double IntelligibilytiSpeech;
        //double SignalNoise = 0.2;
        double SignalNoise;
        double SignalSpeech;
        double MaskingNoise;
        double SummaSpeech;
        double SummaNoise;
        SummaSpeech = 0;
        SummaNoise = 0;
         //Variable Array;
         // In the future, the “SpeechArray” array must be passed data with yVal when forming chart2;
           double[] SpeechArray = chart2.Series[0].Points.SelectMany(t => t.YValues).ToArray();
         // In the future, the “NoiseArray” array must be passed data with yVal when forming chart3;
           double[] NoiseArray = chart3.Series[0].Points.SelectMany(t => t.YValues).ToArray();
           for (int i = 0; i < NoiseArray.Length; i++)
        {
           double deslog = 20;
           double opornzbuk = 50000;
           NoiseArray[i] = NoiseArray[i] / deslog;
           NoiseArray[i] = Math.Pow((double)10, (double)NoiseArray[i]);
           NoiseArray[i] = NoiseArray[i] / opornzbuk;
           NoiseArray[i] = NoiseArray[i] * NoiseArray[i];
           SummaNoise = SummaNoise + NoiseArray[i];
        }
         //Work with a linear array SpeechArray;
```

```
            for (int i = 0; i < SpeechArray.Length; i++)
          {
            double deslog1 = 20;
            double opornzbuk1 = 50000;
            SpeechArray[i] = SpeechArray[i] / deslog1;
            SpeechArray[i] = Math.Pow((double)10, (double)SpeechArray[i]);
            SpeechArray[i] = SpeechArray[i] / opornzbuk1;
            SpeechArray[i] = SpeechArray[i] * SpeechArray[i];
            SummaSpeech = SummaSpeech + SpeechArray[i];
          }
  }
          MaskingNoise = Math.Sqrt(SummaNoise);
          SignalSpeech = Math.Sqrt(SummaSpeech);
          SignalNoise = SignalSpeech / MaskingNoise;
          IntelligibilytiSpeech = Math.Exp(17.4 * SignalN
          textBox2.Text = IntelligibilytiSpeech.ToString();
               }
```

Figure 5 shows the appearance of the software product main window after the calculations performing. The data for speech intelligibility assessing in relative units (and not in percent) as an indicator of the speech information protection degree is given in the box “Speech intelligibility”.

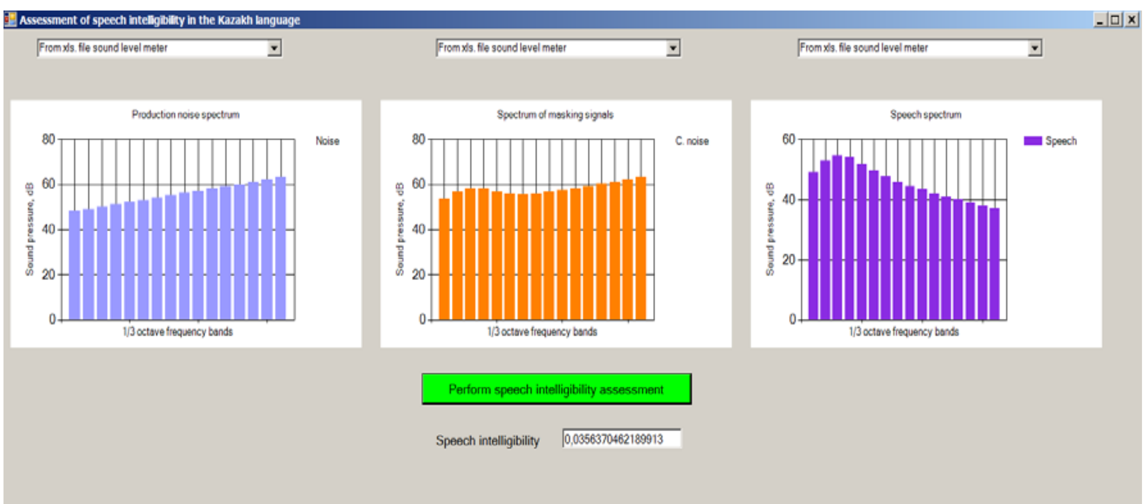


Figure 5. The appearance of the software product main window after the calculations performing.

The of voice information security degree in accordance with the values of its permissible intelligibility at the border of the security zone is established by the consumer in accordance with the the protected voice information importance degree.

5. CONCLUSION

So, the results of experimental studies confirmed the hypothesis about the effectiveness of combined speech masking signals for speech information protection, which was put forward in previous works [6 - 8]. This work is the final in a series of works about speech information protection against leakage via acoustic channels using combined speech masking signals and has a practical solution with a software product.

Experimental studies of intelligibility of the speech masked by only “white” noise by selected and trained auditors shown the following results: with a signal-to-noise ratio of –24 dB, speech intelligibility was less than 1 %. In general, the dependence coincides with the results presented in [16].

Intelligibility of the speech masked by the combined signals, including “white” noise and speech-like signals, which were 6 dB less than "white" noise, for the same selected and trained auditors was less than 1 % for a signal-to-combined masked noise ratio of –14 dB. Thus, it turned out that the usage of combined speech masking signals made possible to reduce the level of masking signals by 10 dB, and thereby to provide more effective protection and more comfortable conditions for a personnel in the protected room.

On the dependence of verbal speech intelligibility on the ratio of the speech signal / combined masking noise, there is a sharp increase in verbal speech intelligibility after the signal-to-noise ratio of more than –11 dB. This is due to the fact that the differences in levels between the protected speech signal and the speech-like signals included in the combined masking signals become smaller. In the study of speech intelligibility masked by combined signals, auditors first of all want to recognize more powerful signals – these are speech-like signals against a background of "white" noise. Information speech signals are lower in level of both speech-like signals and "white" noise, and the auditor has to recognize them against the background of speech-like signals and "white" noise.

The developed software product "Automatization001" allowed us to automate the process of speech intelligibility assessment, as well as to find the integrated speech signal / combined masking noise ratio for a given speech intelligibility value at the border of the protected area, which is important from the point of view of the practical use of the results of this series of works.

6. ACKNOWLEDGEMENT

The work was supported by the grant funding from the Ministry of Education and Science of the Republic of Kazakhstan, № AP05130293 and partially supported by the Belarus Foundation for Basie Research, grant № F20KITG-002

REFERENCES

[1] STB GOST R 50840-2000 (State Standard System). Transmission of speech through communication paths. Methods for assessing quality, intelligibility and recognizability. Minsk, 2000, 366 p. (in Russian)

[2] Pokrovskij N.B. Calculation and measurement of speech intelligibility. Moscow, Svyazizdat, 1962, 392p. (in Russian).

[3] Sapozhkov M. A. Speech signal in cybernetics and communication. Moscow, Svyazizdat, 1963, 452 p. (in Russian).

[4] Zheleznjak V.K., Makarov Ju.K., Horev A.A. Some methodological approaches to assessing the effectiveness of voice information protection. Special equipment, 2000, no. 4, pp. 39–45. (in Russian).

[5] Bradley S.J., Cover B.N. Designing and Assessing the Architectural Speech Security of Meeting Rooms and Offices: IRC Research Report, RR – 187, August, 2006, 45 p.

[6] Seitkulov Yerzhan N. Method for speech intelligibility assessment with combined masking signals. Yerzhan N. Seitkulov, Seilkhan N. Boranbayev, Banu B. Yergaliyeva, Henadzi V. Davydau, Aleksandr V. Patapovich, Journal of Theoretical and Applied Information Technology, 30th April 2020, – vol. 98, no 08, pp.1173–1186.

[7] Seitkulov Y.N., Davydau H.V., Patapovich A.V. The base of speech structural units of Kasakh language for the synthesis of speech-like signals. Proceeding of the IEEE 12th International Conference on Application of Information and Communication Technologies, Almaty, 17 – 19 October 2018.

[8] Seitkulov Y.N., Boranbayev S.N., Davydau H.V., Patapovich A.V. Algoritym of forming speech base units using the method of dynamic programming, Journal of Theoretical and Applied Information Technology, 15th December 2018, vol. 96, no. 23, pp.7928–7941.

[9] Kawan J. Antenna array simulation for location of sonic signal sources, Jamal Kawan, A. Davydau, H. Davydau, 17th International Congress on Sound and Vibration 2010 (ICSV 17) Cairo, Egypt 18–22 July 2010, vol. 5, pp. 3547–3554.

[10] Gavrilenko O.V, Didkovskij .V. S., Prodeus A.N. Calculation and measurement of speech intelligibility at small signal-to-noise ratios. Part 1. Correct measurement of the speech distribution function. Electronics and communications. Thematic issue "Problems of Electronics", part 1. 2007, pp. 137–141. (in Russian).

[11] Gavrilenko O.V, Didkovskij V. S., Prodeus A.N. Calculation and measurement of speech intelligibility at small signal-to-noise ratios. Part 2. Correction of perception

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

coefficients Correct measurement of the speech distribution function. Electronics and communications. Thematic issue "Problems of Electronics", part 1. 2007, pp. 142–147. (in Russian).

[12] Trushin V. A., Reva I. L., Ivanov A. V. On methodological errors in evaluating verbal intelligibility of speech in information protection problems. TUSUR reports, № 1(25), 2012, pp. 180–185. (in Russian).

[13] Bucula, A.P., Ivanov A. V., Reva I. L., Trushin V. A. On the reliability of the assessment of the security of speech information from leakage through technical channels. TUSUR reports, no. 1 (21), 2010, pp. 89–92. (in Russian).

[14] Seitkulov Y. N. Spekers and auditors selection technique in assessing speech information security, Y.N. Seitkulov, S.N. Boranbayev, H.V. Davydau, A.V. Patapovich, Journal of Theoretical and Applied Information Technology, 30th June 2019, vol. 97, no 12, pp. 3306–3316.

[15] Gotovko M.A., Assessment of the security of speech information, M.A. Gotovko, G.V. Davydov, E.N. Seitkulov, Information Technologies and Systems 2013 (ITS 2013): Proceedings of The International Conference, BSUIR, Minsk, 24th October 2013, pp. 268–269. (in Russian)

[16] Trushin V. A., Reva I. L., Ivanov A. V. Improving the methodology for assessing speech intelligibility in information security tasks. Polzunovsky Bulletin, no 3/2, 2012, pp. 238–241.(in Russian)