Ministry of Education of the Republic of Belarus
Educational Institution
Belarusian State University of Informatics and Radioelectronics

UDK 004.934

Abdulhadi
Haitham Dheyaa Abdulhadi

The speech intelligibility determining algorithm for information security systems

ABSTRACT
for the Degree of Master of Engineering
Specialization 1-98 80 01 «Methods and Systems of Information Protection, Information Security»

Dissertation Advisor
Oleg Borisovich Zelmansky
Candidate of Engineering Sciences,
Assistant Professor

CONTENTS

GENERAL DESCRIPTION OF THE WORK

Objectives and tasks of the research

The objective of this dissertation is the algorithm of determination of intelligibility of the speech that is affected by different types of noise signal generated by information security means and creation on its basis a hardware-software complex for assessing the speech data protection under the effect of speech-like noise.

To achieve the said objective it is necessary to solve the following tasks:

1. To analyze existing methods of speech intelligibility determination and noise synthesis.
2. To study the influence of the type of the noise signal on the speech intelligibility.
3. To design the bench-scale plant for determination of speech intelligibility that is affected by noise signal.
4. To develop the hardware-software complex of evaluation of speech data security when it’s influenced by speech-like noise.

Speech and noise signals were selected as the object of the study. The subject of the study is presented by algorithms of speech intelligibility analysis.

The topic of Dissertation work corresponds to the priority scientific areas of fundamental and applied research in the sphere of information and engineering and technical safety, creation of modern information protection systems.

Personal contribution of Master's degree student

Contents of Dissertation work demonstrate personal contribution of the author. Main scientific and practical results were obtained personally by the author.

In works published in cooperation, the author focused on speech intelligibility analysis applied to mixture of the speech signals and different types of noise signals.

Dissertation advisor O.B. Zelmansky, candidate of engineering sciences, assistant professor, is a co-author of main publications. He formulated objectives and tasks of the research, decided on research methods, participated in work planning and results discussion, interpreted and summarized obtained results, completed scientific editing of Dissertation materials.

Practical approval and publication of Dissertation results

Main provisions and results stated in Dissertation were presented and discussed at XIV Belarusian-Russian scientific and technical conference “Technical Means of Information Protection” (Minsk, 2016), 52nd Scientific conference of graduate students, undergraduates and students of BSUIR (Minsk, 2016).

INTRODUCTION

Speech data protection is one of the most important tasks in general set of activities on data security of a facility or an establishment. One of the most efficient ways of speech data protection from unauthorized audio interception is the active one which provides creation of masking noise in critical points of the premises. Active protection is implemented by different types of jamming devices.

Speech is a noise-like process with complicated amplitude and frequency modulation and consequently the best form of masking interfering signal is also noise process with normal law of distribution of density of instantaneous values probability (i.e. white or pink noise). Speech-like interference is especially effective in terms of features similarity.

The goal of this work is to study the efficiency of different types of interfering signals.

The following tasks deserve particular attention:

1. Speech-like interference.

2. The ways of speech-like noise generation.

3. Optimization of vibration sensors depending on the type of the used noise signal.

Noise efficiency is evaluated through speech intelligibility. Relevance of the present work is in solving tasks mentioned above.

CHAPTER 1
DETERMINATION OF THE INTELLIGIBILITY OF THE SPEECH BEING AFFECTED BY THE NOISE SIGNAL

1.1 Speech intelligibility determination by means of the most common methods. Comparative analysis

Speech intelligibility represents integral estimation of speech signal and is defined as ‘the extent to which speech can be understood (deciphered) by the listeners’ according to international standard ISO/TR4870. Thus, it is the extent to which the listeners can understand the meaning of a phrase, identify words, syllables and phonemes. Accordingly, intelligibility is divided into different types: phonemic, syllabic, word and phrasal, which are interconnected and can be converted one into another. Due to the fact that the degree of predictability during phrase listening is higher than during listening to separate words or syllables, phrasal intelligibility is higher than word intelligibility, word – higher than syllabic, syllabic – higher than phonemic.

The following factors can be singled out from those multiple influencing speech intelligibility first of all: masking with other sounds including noise, reverberation, sound reinforcement path.

Different expert methods and standards such as GOST 25902-83, GOST 51061-97, International standards ISO/TR4870, IEC 268-16, American standard ANSI S3.2-1989, etc. are practically applied to determine speech intelligibility, in particular, during estimation of acoustic properties of lecture halls, theaters, concert halls, studios and other rooms. The following methods of speech intelligibility determination can be referred to as expert: attenuation equivalent method, selection tables method, articulation tables method . Attenuation equivalent method consists in measurement of sound intelligibility dependence on attenuation for tested and standard paths. The disadvantage of such a method is the necessity of standard path. Moreover, presence of amplifier in the path limits usage of this method. Intelligibility measurement according to selection tables consists in measurement of mistakes number during transmission of separate words from the group of phonetically similar words through the tested path. Low training requirements to operators and small measurement duration can be referred to method advantages. The method using articulation tables represents measurement of a relative number of correctly transmitted words, syllables and sounds through the tested path. Usage of syllable tables has a disadvantage connected to the fact that during syllables reading naturalness is lost to a significant extent as well as intonation overtone of oral speech. Limited number of word tables and their high memorability make their frequent usage quite undesirable. Phrase tables are almost not used due to their great excessiveness and consequently low sensitivity to distortions. Digital tables are used in cases of extremely low intelligibility.

Given standardized rules first of all concern selection of test material: specially composed phrase, word or syllable tables which are recorded and broadcast by a speaker to estimate premises, public address or other communication systems. However, they suggest quite complicated, long and expensive procedures.

Besides, three groups of objective methods of speech intelligibility determination can be singled out: formant, modulation, empirical . The following formant methods can be pointed out among the foreign ones: Articulation Index (AI), Speech Intelligibility Index (SII). It is considered within the framework of AI version that speech intelligibility is proportional to average difference between peak level of speech and effective level of masking noise. AI improvement resulted in appearance of SII standardized into ANSI S3.5-1997. Domestic formant methods to which the methods of Pokrovsky , Bykov , Sapozhkov , Kalintsev can be referred are significantly based on foreign works. The methods that can be referred to modulation ones are STI (Speech Transmission Index), STIr (revised Speech Transmission Index), RASTI (Rapid Speech Transmission Index), STITEL (STI for telecommunication systems), STIPA (STI for Public Address). Usage of STI method makes it possible to take account of noise and reverberation interference simultaneously which is provided by a special choice of test signal in the form of noise with the spectrum identical with the specter of long lasting speech. This noise is modulated by periodic signal in each octave frequency band in such a manner that the envelope of momentary signal power would be sinusoid-shaped . According to “full” version of STI method, also called STIr or STI-14, there are 98 values of STI index that are obtained for 14 values of modulation frequencies which are averaged thereafter by means of special methods.

RASTI method represents cut version of STI method. Both STI and RASTI method allow taking account of reverberation interference. However, the account of background noise with irregular spectrum and of nonlinearity distortions is not taken correctly. Nevertheless, RASTI method can be used for approximate premises diagnostics.

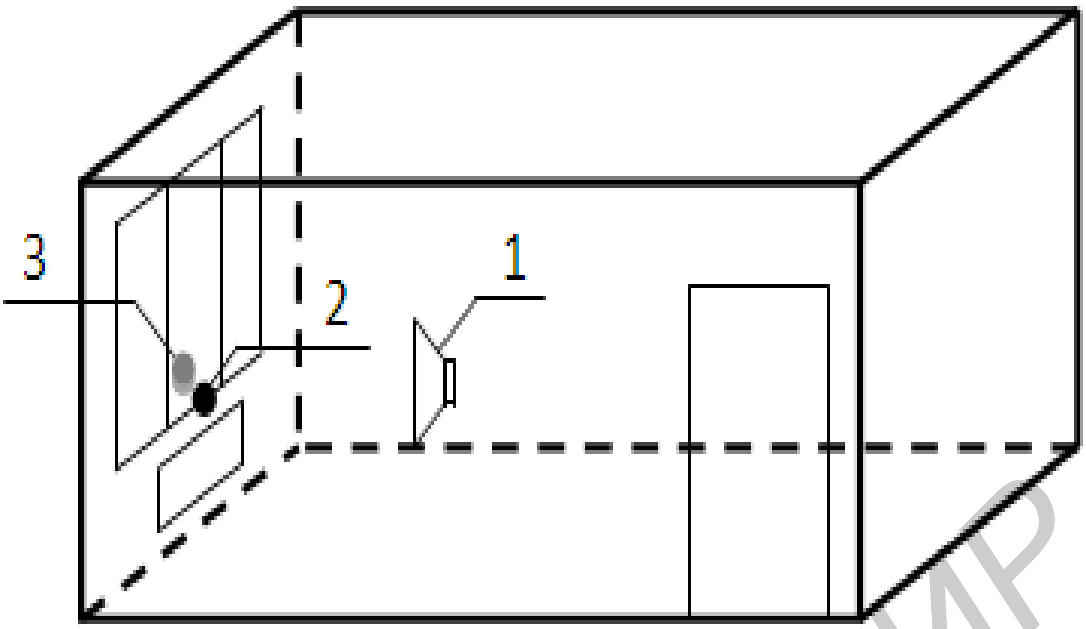
STIPA method is a modification of STI method for public address systems which allows taking account of reverberation as well as nonlinearity sound distortions in premises. Test signal is simplified in the sense that only two modulations are used in each of seven octave bands. In all other respects STIPA is identical with STIr method.

In STITEL method only one modulation frequency is used in each of seven 1/1 octave bands. Bearing noise for each 1/1 octave band has spectrum width of 1/2 octave in order to avoid influencing adjoining bands. STITEL method doesn't allow taking account of reverberation interference and nonlinearity distortions.

The most popular among empirical methods is %Alcons – the method of measurement of consonants articulation loss value expressed as a percentage. %Alcons method is widely used, especially in the USA, for approximate evaluation of speech intelligibility and reflects the loss of voiced consonants caused by indoor reverberation and sound absorption.

One more empirical method is C50 coefficient usage which allows to estimate legibility (clearness) of music sounding. This coefficient is calculated as a ratio of early energy (50 ms) of reflected sounds to late energy. At the same time, noise interference influence is not taken into account as in the case of %Alcons coefficient.

Thus, speech intelligibility for optically transparent components of building envelope was calculated using the example of objective formant method of N.B. Pokrovsky and expert method in compliance with the standard STB GOST P 50840-2000. All the measurements were taken in the room schematically shown in the Figure 1.1.



1 – sound source; 2 – indoor measurement point;
3 – outdoor measurement point (check point)

Figure 1.1 – Layout of the room where the measurements of speech intelligibility were taken

1.2 Study of the influence of interfering signal type on speech intelligibility

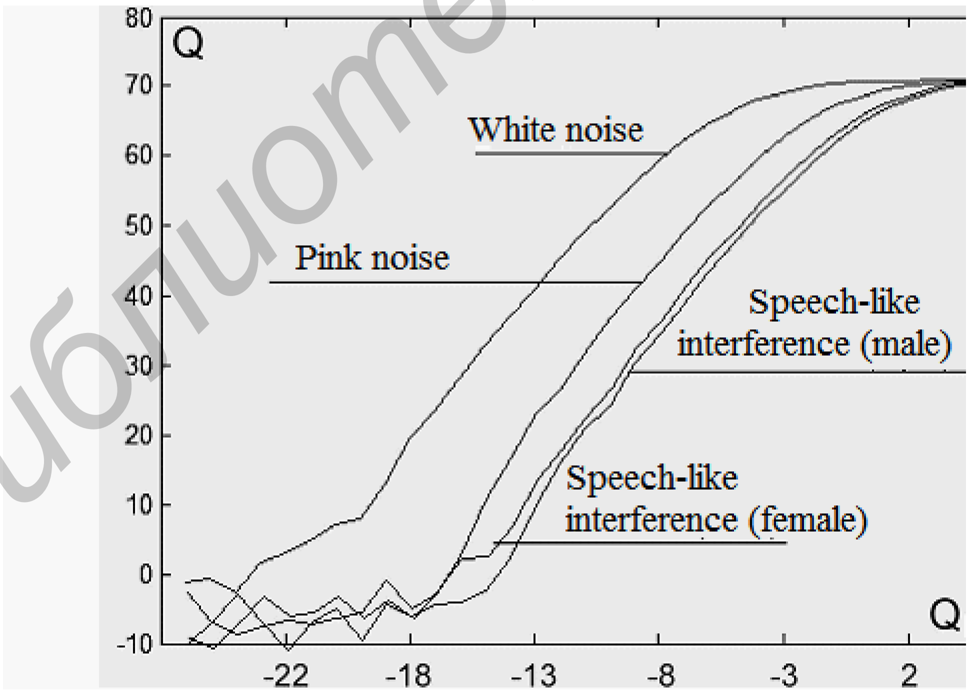
Noise optimization is a crucial practical task. The criterion according to which the optimization will be implemented is speech signal intelligibility with given signal-to-noise ratio.

The results of calculations of speech intelligibility in accordance with instrumentation-calculation method as well as the results of speech intelligibility determination by means of expert method (Table 1.3) were given in the work by Khorev and Makarov. The authors made a conclusion that noise speech-like interference provides relative reduction of speech intelligibility by 15…30% in comparison with “white” noise given that energetics are equal. At the same time, the article describes several ways of speech-like signal generation, each one of which is of different efficiency.

Kunitsyn I.V. and Lobashev A.C. also assume that it is speech-like interference that is more efficient compared to other types of noise interference. However, practicability of articulation tests usage during intelligibility evaluation is cast doubt upon. It is suggested to use the methods of mathematical modeling for efficiency check of different types of acoustic interferences. The authors emphasized the way of speech-like interference generation and came to a conclusion that speech-like interference generated with the same voice that informative signal was generated with possesses the best properties (Figure 1.2) .

Table 1.3 – Values of signal-to-noise ratios at which the required speech data protection efficiency is provided [8]

Noise type	Word intelligibility W, %	Signal-to-noise ratio qi in octave bands					Signal-to-noise ratio in frequency band 175-5600 Hz
		250	500	1000	2000	4000	
White noise	20	0,8	-2,2	-10,7	-18,2	-24,7	-10
	30	3,1	0,1	-8,4	-15,9	-22,4	-7,7
	40	5,1	2,1	-6,4	-13,9	-20,4	-5,7
Pink noise	20	-5,9	-5,9	-11,4	-15,9	-19,4	-8,8
	30	-3,7	-3,7	-9,2	-13,7	-17,2	-6,7
	40	-1,9	-1,9	-7,4	-11,9	-15,4	-4,9
Noise speech-like interference	20	-3,9	-7,9	-12,9	-15,9	-16,9	-9
	30	-1,7	-5,7	-10,7	-13,7	-14,7	-6,8
	40	0,1	-3,9	-8,9	-11,9	-12,9	-5



Qin – signal-to-noise ratio at correlator input;
Qout – signal-to-noise ratio at correlator output
Figure 1.2 – Result of statistical signal processing at correlator output

During the working process, the comparison of speech signal intelligibility in the conditions of influence by different types of interference was made. To this end, the recorded articulation phrase tables were played by the acoustic system and were recorded at certain points of the room using the microphone Octava MK-39, audio card ASUS Xonar Essence ST and personal computer with Cool Edit Pro software installed.

Interfering signal of white noise type and several types of speech-like signal (Russian speech) were reproduced simultaneously with the tables. After that intelligibility was determined by means of expert analysis (by listening to the recorded signal).

Nowadays the methods of speech-like signal synthesis are divided into the following categories (Figure 1.3).

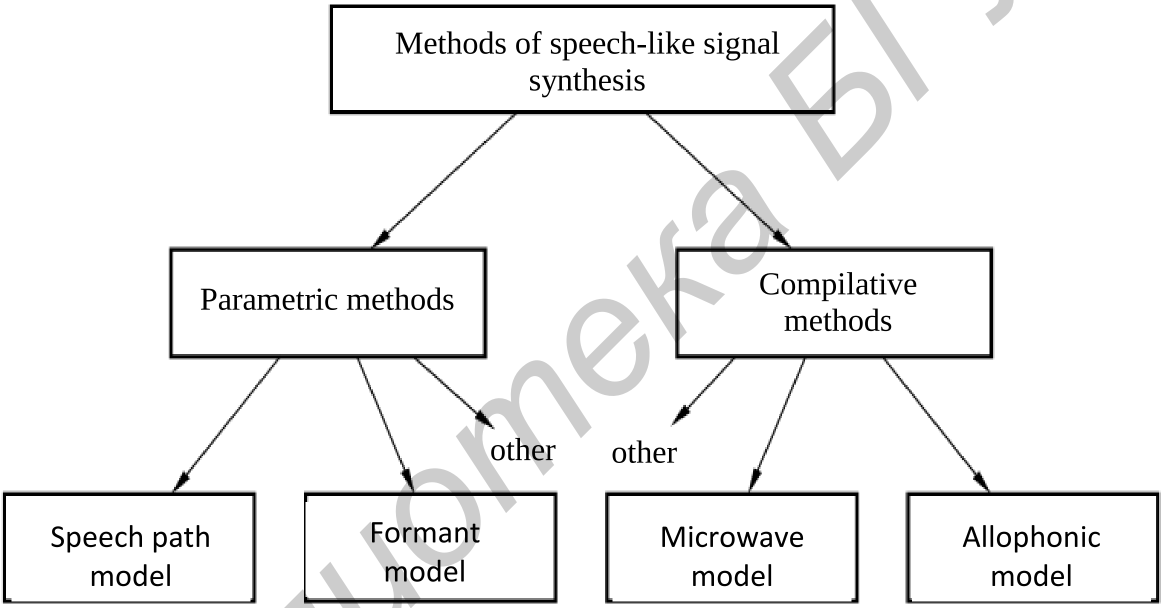


Figure 1.3 – **Classification of the method of speech-like signals generation**

1.3 The algorithm of speech-like interference synthesis for talks protection by means of speakers' allophones compilation

It should be noted that the method of speech-like signal (SLS) compilative synthesis is based on the usage of previously created bases of the allophones singled out not from the speech of talks participants but from the speech of other speakers and consequently the properties of generated SLS will differ from the properties of participants' speech what can reduce reliability of speech data (SD) protection. Due to this fact, the algorithm of SLS synthesis directly from the speech of talks participants without prior creation of participants' allophones bases is needed.

The idea of SLS generation from the allophones singled out directly from the speech of the speakers participating in a protected talk is used as the basis of the developed SLS synthesis algorithm .he idea of SLS generation from allophones consists in the following .

1. A word is composed of the syllables taking into account their frequency and the positions of stressed and pretonic syllables are determined. Another syllable is formed using the model which allows taking account of previous syllable.

2. A syntagma is composed of the words.

3. A phrase is composed of the syntagmas.

4. A phonopassage is composed of the phrases.

5. The words of an orthographic text are transformed into the words of a phonemic text.

6. Consecutively arranged letters of a phonemic text are transformed into allophones sequence and are vocalized. Herewith, the account is taken of pauses duration between syntagmas, phrases and phonopassages as well as generated fundamental pitch frequency (FPF).

Syllables formation, taking into account previous syllable, is realized in the following way. Statistically representative text is divided into syllables according to the rules:

– syllable must contain a vowel;

– consonants standing before the first vowel of the word belong to the first syllable;

– consonants standing after the last vowel of the word belong to the last syllable;

– the first one of the consonants standing between two vowels belongs to the previous syllable, the rest of them belong to the following syllable;

– one consonant standing between two vowels belongs to the following syllable;

– soft sign belongs to the preceding consonant and is not separated from it.

The number of the same syllables is taken into account in the process of text division into syllables. Thereafter, a glossary of the most frequently used syllables that constitute 90% of the text is composed. Then the text is divided again into syllables taking into account the glossary obtained according to the rules described above. The essence of this stage consists in conditional probability analysis of occurrence of one or another syllable from the glossary taking into account preceding syllable or its absence. The results are presented in a tabular form.

A syllable to be chosen from the glossary for word composition is determined by the number randomly generated and equally distributed from 0 to 1. Distribution used during syllable search is determined by the preceding, i.e. previously chosen syllable, which points to a certain line of conditional probabilities table. Based on the rule of each letter transformation into a phoneme, the words of an orthographic text which are formed in such a way are transformed into the words of a phonemic text. The words composed of phonemes according to the rule of each phoneme transformation into an allophone are transformed into allophone sequence.

Significantly, phonemic text formation is carried out in compliance with statistical language peculiarities. Vocalization of a phonemic text is carried out on the basis of formed sequence of allophones, utterances of which are contained in allophone bases formed directly from the speech of talks participants.

Algorithm of automatic allophone bases formation directly from the speech of talks participants for speech-like interferences analysis consists in segmentation found during speech detection for phonetic units and their allophonic classification according to minimum distance criterion.

Speech detection consists in identification and separation of speech from acoustic environment. Speaker verification allows separating speech of one talk participant from speech of other talk participant to create allophone base from his very speech what also allows separating different speakers' speeches .Besides, verification can result in determination that allophone base for this speaker was previously created and there is no need in creating a new one.

In the process of segmentation, speech separated during detection is divided into the sections of uniform vibrations corresponding to phonetic units by finding interphonemic transitions. The method of speech segmentation uses the analysis of spectrum change function as measures of correlation between consecutive windows of signal analyzed. Herewith, the distance between parameter vectors circumscribing signal windows.

When classifying phonetic units, it is determined which class of phonemes each of them belongs to. The method of speech classification consists in calculation of the coefficients of difference between the analyzed phonetic unit and the base of samples of phoneme realization options conditioned by certain phonetic surroundings of these phonemes presented by cepstral coefficients based on correlation matrix. Decision-making about analyzed phonetic unit belonging to one or another group of phonemes and classification of this unit as a specific phoneme is carried out by finding a sample, duration of which is closest to the analyzed phonetic unit and to which the smallest value of difference coefficient corresponds. It results in creation of speaker's allophones base containing utterances of phonetic units and information about them.

Nevertheless, SLS synthesis directly from the speech of the speakers participating in talks requires availability of at least one allophone base created in advance which will be used in the beginning of the talk when allophone bases of this talk participants haven't been yet formed. Usage of incomplete allophone bases of each of the speakers, which were created in advance and will be completed during the talk, is also possible. Such a necessity allows suggesting multifunctional algorithm of SLS synthesis, which combines the ability of SLS generation based on already existing allophone bases created in advance with the possibility of adding SLS generated directly from the speech of talks participants to these bases. Application of this synthesis algorithm will allow starting SLS generation at a proper point of time without prior registration of talks participants due to the usage of allophone bases created in advance from the speech of random speakers or after prior registration of talks participants using allophones separated directly from their speech. Moreover, SLS generation is possible in case of using the bases created in advance, gradually mixing into them SLI generated from allophones, which are generated directly from the speech of these talks participants, during the talks.

Signal generated in such a way is overlaid on the speech of confidential talk participants. The obtained mixture of information and masking signals can be picked up by secret voice recorders or microphones and other intercepting devices but it will be impossible to restore the meaning of talk.

CHAPTER 2
DESIGNING OF THE BENCH-SCALE PLANT FOR EXPLORATION OF THE INFLUENCE OF THE TYPE OF THE NOISE SIGNAL ON THE SPEECH INTELIGIBILITY

2.1 Creation and record of articulation tables

Test acoustic signals were generated based on articulation tables from STB GOST P 50840-2000 .

Speech intelligibility value is the criterion of speech data security for speech signals . Initial data necessary for the analysis of acoustic speech signals security are sound pressure level of speech modeling signals, sound proofing of standard components of building envelopes, sound pressure level of background acoustic noise and industrial vibrations levels of filling elements of premises structure. Obviously, speech intelligibility rate can be also used for efficiency evaluation of closing of technical channels of speech data leakage.

Experimental evaluation of speech intelligibility should be carried out in compliance with STB GOST P 50840-2000 using phrase articulation tables. To that end, a team of speakers and auditors who take special training is created.
The team of speakers consisting of five men and five women was created. Recording of acoustic signals representing test phrases from the Table was performed using the microphone Octava MK-319, audio card ASUS Xonar Essence ST and personal computer with installed software Sound Forge 9.0 in acoustically muffled room. Processing of the audio files obtained was carried out in Sound Forge 9.0 and Cool Edit Pro software environment.

2.2 Design of a bench-scale plant for determination of speech intelligibility when it is influenced by interfering signal

Experimental bench-scale plant which allows determining speech intelligibility including that going through the components of building envelope and is influenced by interfering signal was designed during working process. This plant presents the analog of a small anechoic acoustic chamber.

Bench-scale plant description. Experimental plant is made of two parts of metal tube with wall thickness of 6 mm (inside diameter – 0.26 m, parts length – 0.8 and 0.4 m) either of which is tightly welded up at outer faces and ring flanges with rubber pad for fixation of the sample components of building envelope are welded at inner faces (Figure 2.1).

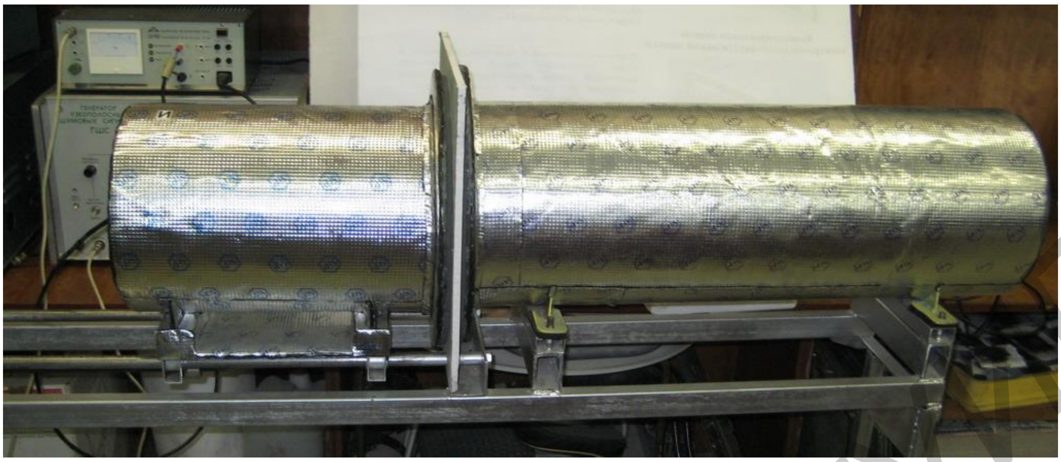


Figure 2.1 – Bench-scale plant

Both parts of the tube are fixed on metal bed, one – stationary on vibration insulating rubber pads, the other moves by means of worm gear drive. There is a microphone M-101 with microphone preamplifier BPM-101 installed in immobile part of the plant. The microphone is hung on rubber thread to reduce the influence of tube vibration on microphone. JUMP speaker of peak/nominal power of 130/65 W is installed in mobile part of the plant.

Vibration insulating material STP Bimast-Bomb of 4 mm thickness and 6.0 kg/m^2 specific weight based on bitumen and mastic compound with addition of anti-adhesive paper and aluminum foil was applied to internal and external plant surfaces to reduce indirect sound transmission to low-level chamber through metal bed. This material is vibration absorber (vibration damping device). Vibration damping device is stuck onto the surface being processed during operation process what makes it much heavier consequently reducing its vibration (converting into thermal energy). Noise-absorbing material based on foam rubber “Comfort S-Max” of 6 mm thickness efficient in noise absorbing within frequency range of 500 – 6 000 Hz to 32 dB was applied onto internal plant surface as the second layer. This material is highly elastic within wide temperature range (from -40 to +105°C), has long lifetime and its technical characteristics remain unchanged in the course of time (up to 20 years), it has low thermal conductivity (0,036 W/m2), belongs to flammability group G1 (doesn’t support combustion for it has the property of self-extinguishability in case of fire).

For the purpose of, Vibration insulating mounts (Figure 2.2) were used to eliminate vibration transmission from floor to metal bed. The mounts are structurally based on materials made from natural rubber and possess the following properties:

– housing material: galvanized steel;
– vibration insulation material: rubber-metal component;
– operating temperature: from –40 to +80 °C;
– working load range: from 250 to 1500 kg;
– mount length: 186 mm;
– mount width: 121 mm;
– mount height: 42 mm;
– mount weight: 1,5 kg;
– stud diameter: M16.

Figure 2.2 – Vibration insulating mount

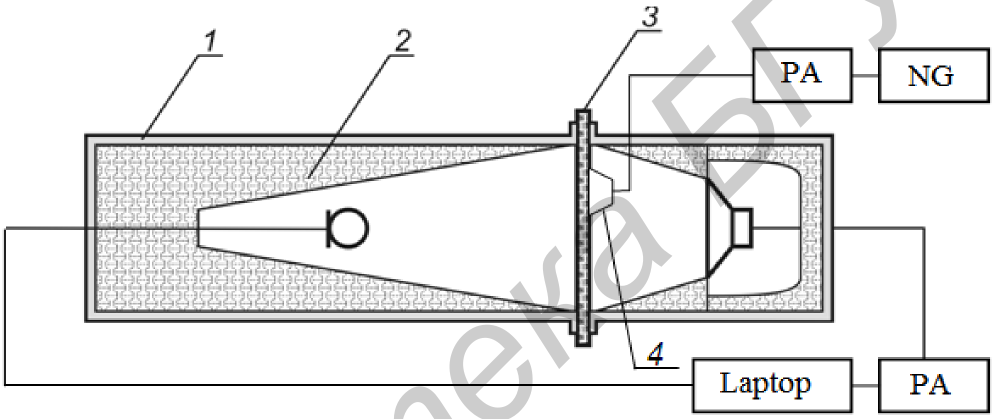
Mounting of vibration insulating mounts is carried out under metal bed of the plant at predetermined points using bolts and nuts. Due to the possibility of adjusting seated height, vibration insulating mounts allow performing compensation of skews emerging in case when foundation provided is not flat enough.

Acoustic design of "closed box" type (hermetically sealed closed acoustic system housing) was used to get flat amplitude-frequency response of the radiant plant part. Acoustic system of such a design is notable for simple structure and acceptable transient characteristics which are resulting from high resilience of 'diffusor – internal housing volume' oscillatory system.

After calculation of Thiele/Small parameters of the loudspeaker being used, such an internal volume of plant housing behind the loudspeaker was picked so that resonant frequency of the acoustic system made 95 Hz and doesn't fall within frequency range studied (200 – 7000 Hz).

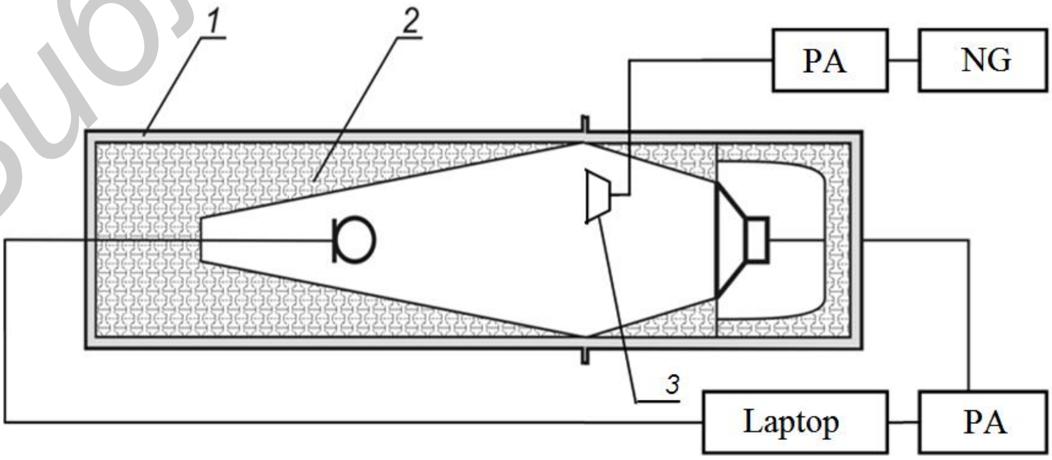
Bench-scale plant operating principle. Previously recorded test acoustic signals presenting test phrases from STB GOST 50840-2000 are reproduced using a laptop connected to plant loudspeaker through power amplifier LV-103.

Envelope component, on mobile side of which a vibration oscillator of interfering signal connected to noise generator through power amplifier is fixed, is placed between plant flanges to determine speech intelligibility going through the components of building envelope and being influenced by interfering signal (Figure 2.3). When it is direct acoustic signal that is studied, building envelope components are not applied, acoustic signals are distributed directly from mobile part of the plant to immobile one. At the same time, acoustic transducer (loudspeaker) placed in mobile part of the plant serves as the source of interfering signal (Figure 2.4).



1 – metal tube; 2 – sound-absorbing material; 3 – building envelope component; 4 – vibration oscillator; NG – noise generator; PA – power amplifier

Figure 2.3 – Schematic view of experimental plant (vibroacoustic channel study)



1 – metal tube; 2 – sound-absorbing material; 3 – acoustic transducer (loudspeaker); NG – noise generator;
PA – power amplifier.

Figure 2.4 – Schematic view of experimental plant (direct acoustic signal study)

Next, the signal presenting a blend of test and interfering signals is recorded to the laptop using the microphone installed in immobile part of the plant and is analyzed by the group of auditors.

CHAPTER 3
DEVELOPMENT OF METHODS AND HARDWARE-SOFTWARE COMPLEX FOR ASSESSING THE SPEECH DATA PROTECTION UNDER THE EFFECT OF SPEECH-LIKE NOISE

3.1 Designing of methods and hardware-software complex of evaluation of speech data security when it’s influenced by speech-like interference

Methods and hardware-software complex of evaluation of speech data security when it’s influenced by speech-like interference were developed on the basis of previously required data.

The methods include the following steps: test articulation tables recording; speech-like interference generation; noise masking of test signals with interfering one; noisy signal recording; determination of signal-to-noise ratio of test and interfering signals; calculation of errors and measurement confidence interval; determination of intelligibility and evaluation of speech signal security. Security evaluation is carried out using speech intelligibility value. Hardware-software complex includes software for speech signals recording and processing, equipment for sound recording and reproduction, multichannel sound level meter/analyzer.

Methods of security evaluation of speech data security when it’s influenced by speech-like interference:

1. Evaluation program.

1.1. Test unit.

Test unit is sound record of speech signal.

1.2. Not less than 40 sound records are presented for evaluation. Each record should contain 50 phrases masked with interfering signal. One half of total number of records is made with female voice and another half – with male voice.

1.3. Test objective.

Test objective is determination of speech signal intelligibility when it’s influenced by acoustic noise.

1.4. Test conditions and procedure.

1.4.1. Test site and equipment characteristics.

1.4.2. Test weather conditions.

Evaluation is carried out in standard climate conditions:

– ambient air temperature from plus 15 to plus 25°C;

– relative air humidity from 55 to 80% at a temperature of plus 25°C;

– atmosphere pressure from 84 to 107 kPa (from 630 to 800 mmHg).

1.4.3. Maintenance requirements.

It is necessary to carry out all the installation works, switching measuring devices on and off in compliance with operating documentation requirements as well as with other regulatory documentation on safety measures.

1.5. Scope of testing.

1.5.1. Duration of preliminary articulation training of articulation team (speakers and auditors) 4 hours.

1.5.2. Working hours of articulation team do not exceed 4 per day.

1.5.3. Total operating time during tests must make not less than 16 hours.

2. The method of influence evaluation of interfering signal type on speech data security.

2.1. Speech intelligibility determination.

2.1.1. Preliminary training of the speakers.

2.1.2. Creation and recording of articulation tables.

2.1.3. Generation of different types of interfering signals.

2.1.4 Preliminary training of the auditors.

2.1.5 Phrase intelligibility J is found as average for measurement cycle according to the formula:

$$J = \frac{1}{N}\sum_{i=1}^{N} Ji \qquad (3.1)$$

where:

J – phrase intelligibility at normal rate, %;

J_i – single measurement result, %; is calculated as a percentage of correctly received phrases;

N – number of single measurements.

2.1.6 J_i value is determined by listening of sound records containing test articulation tables combined with acoustic interference by the auditors and comparing the phrases obtained in such a way with noise-free tables. Reconciliation of initial and obtained data is carried out by the auditor or the checker of appropriate competence.

2.2. Errors calculation.

2.2.1. Excluding gross errors influence on final result using “three sigmas” criterion.

2.2.2. Confidence interval [$\pm S_X$] determination with given reliability value of 0.95 using the formula:

$$S_X = \sqrt{\frac{\sum_{i=1}^{n}\left(\langle N \rangle - N_i\right)^2}{n(n-1)}}, \qquad (3.2)$$

where:
<*N*> – arithmetic average of phrase intelligibility;
Ni – *I* value of realization;
n – total number of measurments.

2.3. Signal-to-noise ratio determination.

2.3.1. Test tables reproduction by means of acoustic system.

2.3.2. Test tables recording using the microphone in the point of possible placement of sound-recording equipment.

2.3.3. Determination of mean-square value of the amplitude of the sound record made.

2.3.4. Interfering signal reproduction by means of acoustic system.

2.3.5. Interfering signal recording using the microphone in the point of possible placement of sound-recording equipment.

2.3.6. Determination of amplitude mean-square value of sound record made.

2.3.7. Signal-to-noise ratio determination using the formula

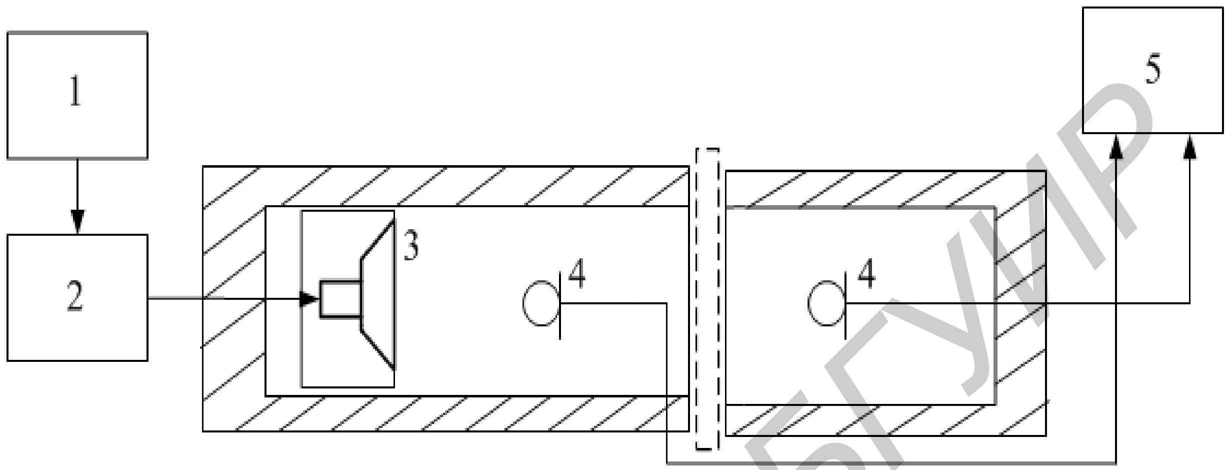
$$\mathrm{SNR(dB)} = 10\log_{10}\frac{P_{signal}}{P_{noise}} = 20\log_{10}\frac{A_{signal}}{A_{noise}} \qquad (3.3)$$

where:
P – average power of test signal or interference;
A – mean-square value of test signal or interference amplitude

2.4. Equipment connection schemes

The measurements were taken using small anechoic chamber made by BSUIR. Such a camera was placed in sound-proof room to reduce external acoustic noises influence.

2.4.1. Measuring equipment calibration is carried out in the following way. In case of indoor, the value of amplifier output voltage shall be adjusted so that noise value on the display of acoustic noise analyzer operating in sound level meter mode makes 90 dB (±5 %) for test tables (Figure 3.1).



1 – personal computer with special software; 2 – power amplifier; 3 – acoustic system;
4 – microphone; 5 – sound level meter/analyzer Manom

Figure 3.1 – Equipment connection scheme when measuring test acoustic signal level

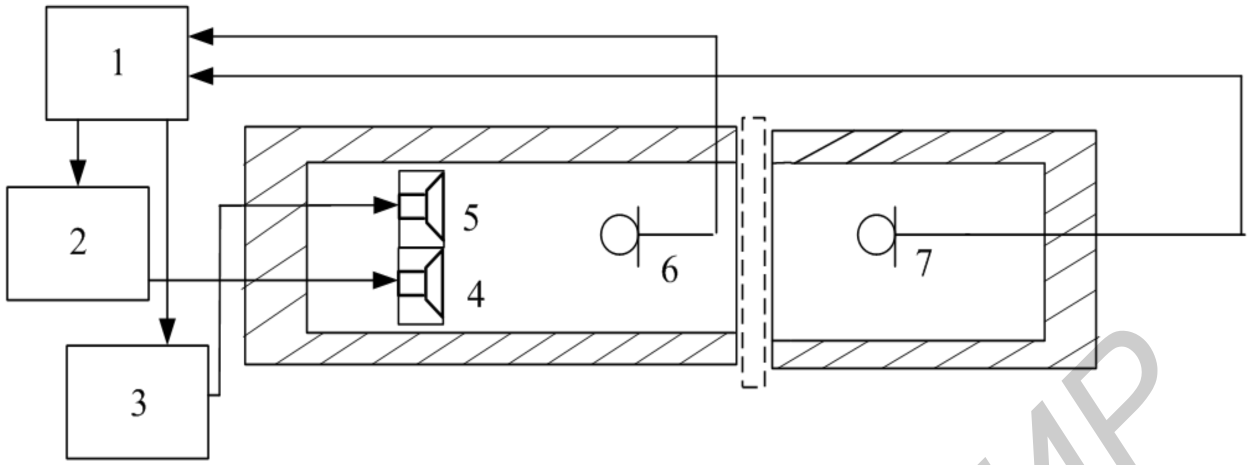
2.4.2. Speech intelligibility study transmitted through direct acoustic channel is carried out according to the scheme in Figure 3.2.

2.4.2.1. Test acoustic signals are reproduced by means of personal computer 1 connected to acoustic signal source 4 through the amplifier 2.

2.4.2.2. Interfering acoustic signal is reproduced by means of generator of personal computer 1 connected to interfering acoustic signal source 5 through the amplifier 3.

2.4.2.4. The mixture of test acoustic and interfering acoustic signals is received by microphone 7 and is recorded in the form of digital sound record to personal computer 1.

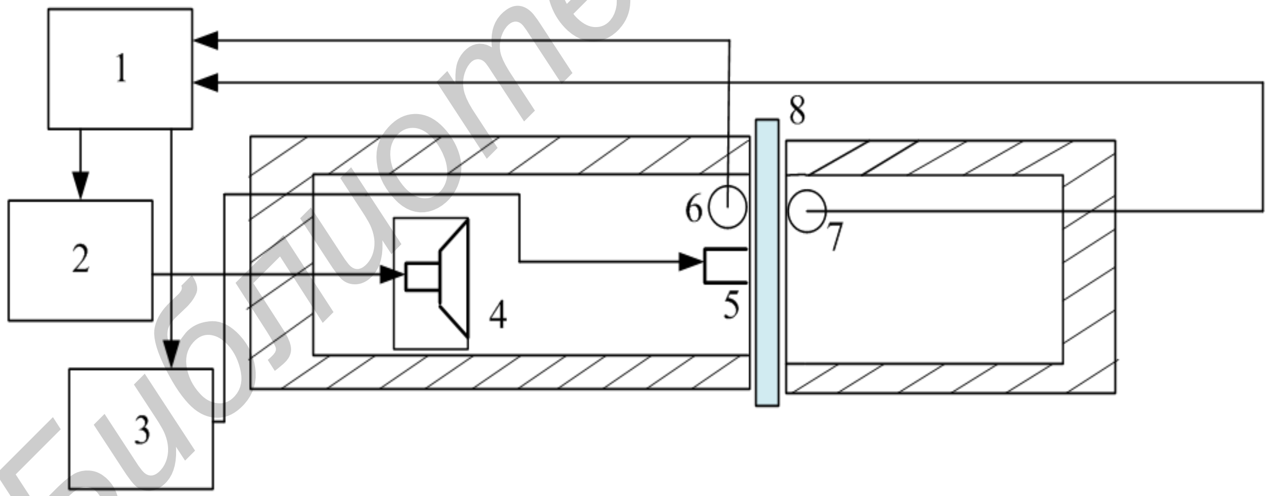
2.4.2.5. Intelligibility is determined in compliance with STB GOST P 50840-2000 .



1 – personal computer with special software; 2,3 – power amplifier; 4,5 – acoustic system; 6,7 – microphone;

Figure 3.2 – Scheme of equipment connection when studying direct acoustic channel of speech data leakage

2.4.3. Intelligibility study of speech transmitted through vibroacoustic channel is carried out according to the scheme in Figure 3.3.



1 – personal computer with special software; 2,3 –power amplifier; 4 – acoustic system; 5 – interfering vibration signal source; 6,7 – accelerometer; 8 – building structure component

Figure 3.3 – Scheme of equipment connection when studying vibration channel of speech data leakage

2.4.3.1. Test acoustic signal are produced by means of personal computer 1 connected to the source of acoustic test signal through the amplifier 2.

2.4.3.2. Interfering acoustic signal is reproduced by means of generator of personal computer 1 connected to the source of interfering vibration signal 5 through the amplifier 3.

2.4.3.3. Signal-to-noise ratio of test and interfering signals is determined according to p. 2.3.

2.4.3.4. The mixture of test acoustic and interfering acoustic signals is received by electronic stethoscope 7 and is recorded in the form of digital sound record to personal computer 1.

2.4.3.5. Intelligibility is determined in compliance with STB GOST P 50840-2000.

2.5. Interfering signal efficiency evaluation

The measurements are taken for different signal-to-noise ratios and different types of interfering signals. Efficiency comparison of different interfering signal types is carried out comparing speech intelligibility value at given signal-to-noise ratio.

3.2 Approbation of experimental methods and hardware-software complex. Study of further development prospects and practical usage of the results obtained

Hardware-software complex implementing experimental methods belongs to testing equipment for speech intelligibility determination in adverse conditions, when influenced by noise, and is intended for determination of efficiency of interfering acoustic signals generated by the devices of active speech data protection.

The complex allows comparing efficiency of speech-like interfering signals obtained in different ways for the systems of active data protection. The complex described above was used to study acoustic properties of air-bubble panel for data protection. Air-bubble panel can be used as integrated protective structure to provide data protection against leakage through acoustic, optical and electromagnetic channels. Intelligibility of speech going through this protective structure was determined.

Possible field of application of the developed methods can study of speech-like interferences efficiency for foreign languages, particularly, for the languages for which formant methods of speech intelligibility determination were not developed. Instrumental and computation method based on the results of experimental studies of N.B. Pokrovsky can be an example of such a method.

Development of speech-like interferences synthesis for data protection of the speech uttered in foreign language (Kazakh, Japanese) is suggested as a direction of further study.

CONCLUSION

Speech intelligibility value was chosen to be protection criterion as a result of literature sources analysis. Method of articulation measurements was chosen for intelligibility determination. Articulation team consisting of five men and five women was created. Test acoustic signals were recorded.

Comparative analysis of expert and objective methods of speech intelligibility determination was carried out. Calculation of speech intelligibility for optically transparent compounds of building envelope was carried out using the example of objective formant method of N.B. Pokrovsky in compliance with the standard STB GOST P 50840-2000.

Bench-scale plant which allows determining speech intelligibility, including speech going through building envelope components, under the influence of interfering signal was designed. Instrumental errors of measuring equipment and measurement errors were determined. The choice of the points in the room for placement of acoustic signals receivers was grounded.

The algorithm of speech-like interference synthesis, which is based on compilation of the segments of speakers' utterance records according to a phonemic text taking into account language statistical peculiarities, was suggested. Evaluation of speech intelligibility under the influence of interfering signal was done. White noise as well as speech-like signal generated in several ways was used as interfering signal. It was shown that speech-like correlated interference is the most efficient.

The methods and hardware-software complex for assessing speech data security under the influence of acoustic interferences was developed.