Assessment of Speech Intelligibility by Formant-Modulation Method

Arkadiy Prodeus

Acoustics and Acoustoelectronics Department, National Technical University of Ukraine (Kyiv Polytechnic Institute),
9/16, Ac. Yangelia Str., 03056 Kyiv, Ukraine

aprodeus@gmail.com

Abstract- **Detailed description of speech intelligibility assessment by proposed formant-modulation method is presented in the paper. Analytical expressions for expectation and variance of estimators of a modulation coefficient and effective signal-to-noise ratio are obtained. Two situations which are most interesting in engineering applications are considered: a predominant noise disturbance and a predominant reverberant disturbance. Convenience in engineering applications equations, which permit finding efficient test signal duration for required measurement exactness, has been obtained. Computer modelling and experimental testing have confirmed efficiency of the formant-modulation method.**

Keywords- Speech Intelligibility; Formant-Modulation Method; Assessment; Measuring Exactness

I. INTRODUCTION

The new instrumental method of speech intelligibility assessment, which had been offered in works [1, 2], was called "formant-modulation" as it was supposed that new method would join better features of formant [3] and MTF [4, 5] methods. The idea of calculation of probability of correct understanding of speech elements (phonemes, syllables, words, phrases) was saved from formant method, and the idea of application of the modulated test signal, which allows taking into account both noise and reverberant disturbances, was borrowed from an MTF method.

Main purpose of the paper is verification of efficiency and assessment exactness of the proposed method.

II. SHORT DESCRIPTION OF FORMANT-MODULATION METHOD

When they make an acoustic examination of apartment with the use of instrumental (objective) methods of speech intelligibility evaluation, they radiate a test signal $x(t)$ in a certain point of apartment (announcer location), and in other point of apartment (listener location) they accept a signal $y(t)$, which is exposed to special processing for calculating one or few numerical parameters, which characterize speech intelligibility of this apartment [3–6].

When an apartment has impulse response $h(v)$, which describes reverberation, and when there is interfering noise $n(t)$ in the same apartment, they watch a signal $y(t)$ in the reception point:

$$y(t) = \int_{-\infty}^{\infty} h(v)x(t-v)dv + n(t). \tag{1}$$

At the use of formant-modulation method [1, 2], they calculate articulation intelligibility A in the same way as in a formant method [3]:

$$A = \sum_{k=1}^{K} p_k \cdot P_k(E_k), \tag{2}$$

where p_k is formant probability distribution on frequency bands; $P_k(E_k)$ is speech perception coefficient; E_k is effective perception level of speech signal in k-th frequency band Δf_k.

A difference between formant-modulation and formant methods consists in difference of signal-to-noise assessment. When using formant method, they use a test signal $x(t)$, which is stationary stochastic process and power spectrum of which is like long term speech spectrum, and they estimate E_k in accordance with equation:

$$E_k \approx q_k = 10\lg(D_{sk}/D_{nk}), \tag{3}$$

where q_k is partial signal-to-noise ratio, D_{sk} and D_{nk} are respectively variances of stationary signal and noise in k-th frequency band. When using formant-modulation method, they estimate E_k in another way, which is used in modulation

method and which permits taking into account presence of reverberation disturbance [4, 5].

In accordance with this way, they use non-stationary stochastic process test signal $x(t)$, variance of which $D_x(t)$ is modulated on harmonic law with modulation frequency F :

$$x(t) = \xi(t)\sqrt{f(t)} \, , \tag{4}$$

$$f(t) = 1 + \cos 2\pi F t \, . \tag{5}$$

$$D_x(t) = \left\langle x^2(t) \right\rangle - \left\langle x(t) \right\rangle^2 = D_\xi(1 + \cos 2\pi F t) \, , \tag{6}$$

where $\xi(t)$ is stationary stochastic process with variance D_ξ and with power spectrum, which is like long term speech spectrum; $\langle \cdot \rangle$ is expectation symbol. It is evident that modulation coefficient m of variance $D_x(t)$ is equal to 1 in this case.

Variance $D_y(t)$ of process $y(t)$ will also be modulated on harmonic law with the same frequency F , but as a result of action of reverberant and noise disturbances, modulation coefficient m of variance $D_y(t)$ will be less than 1. Truly, substituting $h(v) \approx \delta(v)$ in Eq. (1), where $\delta(v)$ is Dirac delta function, we get:

$$D_y(t) = D_\xi(1 + \cos 2\pi F t) + D_n \, , \tag{7}$$

$$m = D_\xi / (D_\xi + D_n) \, . \tag{8}$$

Signal-to-noise ratio value follows from Eq. (8):

$$SNR = 10\lg(D_\xi / D_n) = 10\lg[m/(1-m)] \, . \tag{9}$$

At the united action of noise and reverberation disturbances, analytical expression for the modulation coefficient m of variation $D_y(t)$ appears more difficult, as compared to Eq. (8), and depends not only on noise variance D_n , but also on standard reverberation time T_{60} , and even on modulation frequency F .

MTF method authors suggested using of a set of 14 modulation frequencies F_i for imitation of non-stationary character of the real speech signal variance. They also introduced concept of effective signal-to-noise ratio:

$$\overline{SNR}_k = \frac{1}{14}\sum_{i=1}^{14} SNR_{k,i} \, , \tag{10}$$

$$SNR_{k,i} = 10\lg \frac{m_k(F_i)}{1 - m_k(F_i)} \, . \tag{11}$$

where $m_k(F_i)$ is modulation coefficient of variance $D_y(t)$ of signal $y_k(t)$ in k -th frequency band [4, 5]. $SNR_{k,i}$ is calculated on formula, which describes the noise disturbance action, while the modulation coefficients $m_k(F_i)$ carry on themselves the seal of action by not only noise but also reverberation disturbance. Averaging of $SNR_{k,i}$ on the number i of modulation frequency F_i in Eq. (10) results in a double effect: influence of F_i on the results of measuring is repressed, and also variance of the estimation diminishes.

It becomes clear in this connection, why it is suggested in a formant-modulation method to estimate the effective perception level E_k of speech signal in accordance with equation:

$$E_k = \overline{SNR}_k \, . \tag{12}$$

As far as estimation of modulation coefficients $m_k(F_i)$, it is expedient to produce it by means of Fourier transformation because of harmonic modulation of variance $D_y(t)$ [5]:

$$\tilde{m}_k(F_i) = \frac{2|A_{ik}(F_i)|}{|A_{ik}(0)|} \, , \tag{13}$$

$$A_{ik}(f) = \frac{1}{T}\int_0^T y_{ik}^2(t)e^{-j2\pi ft}dt,$$

(14)

were $|\cdot|$ is absolute value symbol; T is duration of process $y_{ik}(t) = \int_{-\infty}^{\infty} h(v)x_{ik}(t-v)dv + n(t)$; $x_{ik}(t) = \xi_k(t)\sqrt{1+\cos 2\pi F_i t}$ is

modulated band-pass white noise in the k-th frequency band.

<center>III. ANALYSIS OF MEASURING EXACTNESS</center>

If we suppose that formants probability distribution on frequency p_k and perception coefficient $P_k(E_k)$ are known with high exactness, then it follows from Eqs. (2), (4), (5), (10)–(14), that a bias and variance of intelligibility estimator (2) are completely determined by statistical properties of magnitudes $A_{ik}(0)$ and $A_{ik}(F_i)$.

A. Prevailing Influence of Noise

In the beginning we shall make a statistical analysis of a pair of stochastic variables $A_{ik}(0)$ and $A_{ik}(F_i)$ in the supposition that noise disturbance is prevailing above reverberation disturbance. In this case model of a signal in k-th frequency channel can be presented as:

$$y_{ik}(t) = \xi_k(t)\sqrt{1+\cos(2\pi F_i t + \varphi_k)} + n_k(t),$$

(15)

where $\xi_k(t)$ and $n_k(t)$ are normal band-pass (in k-th frequency band Δf_k) white and statistically independent stationary stochastic processes with zero expectations; φ_k is unknown initial phase.

At first let's analyze expectation of magnitudes $A_{ik}(0)$ and $A_{ik}(F_i)$. Omitting, for simplification, indexes in Eqs. (14) and (15), we receive:

$$\langle A(f) \rangle =$$
$$= \frac{1}{T}\int_0^T \langle y^2(t)\rangle e^{-j2\pi ft}dt = (D_\xi + D_n)e^{-j\pi fT}Sa(\pi fT) + \frac{D_\xi}{2}e^{j\varphi}e^{-j\pi(f-F_i)T}Sa[\pi(f-F_i)T] + e^{-j\varphi}e^{-j\pi(f+F_i)T}Sa[\pi(f+F_i)T]$$

(16)

where $Sa(x) = \sin(x)/x$. At implementation of condition $T = r/F_i$, where r is arbitrary positive integer, it follows from (16):

$$\langle A(0) \rangle = D_\xi + D_n,$$

(17)

$$|\langle A(F_i) \rangle| = \langle |A(F_i)| \rangle = \frac{D_\xi}{2}.$$

(18)

As can be seen, equation $T = r/F_i$ is a condition of an unbiased estimator of articulation intelligibility (2) by means of formant-modulation method.

We will now find variance of magnitudes $A_{ik}(0)$ and $A_{ik}(F_i)$. In general case:

$$D\{|A(f)|\} = \frac{1}{T^2}\int_0^T\int_0^T \langle y^2(t_1)y^2(t_2)\rangle e^{-j2\pi f(t_1-t_2)}dt_1 t_2 - \frac{1}{T^2}\int_0^T\int_0^T \langle y^2(t_1)\rangle\langle y^2(t_2)\rangle e^{-j2\pi f(t_1-t_2)}dt_1 t_2,$$

(19)

$$\langle y^2(t_1)y^2(t_2)\rangle - \langle y^2(t_1)\rangle\langle y^2(t_2)\rangle = 2D_n^2\left[\frac{D_\xi}{D_n}R_\xi(t_2-t_1)f(t_1)f(t_2) + R_n(t_2-t_1)\right]^2,$$

(20)

where $R_\xi(\tau)$ and $R_n(\tau)$ are correlation coefficients of stationary stochastic processes $\xi(t)$ and $n(t)$ respectively.

Since

$$R_\xi(\tau) = R_n(\tau) = Sa(\pi\Delta f\tau)\cos 2\pi f_0\tau,$$

(21)

where Δf is frequency bandwidth of processes $\xi(t)$ and $n(t)$, f_0 is central frequency of this band-pass, after rather bulky

calculations, under the assumption $T\Delta f \gg 1$, it is possible to get next final formula:

$$D\big[|A(0)|\big] \approx D\big[|A(F_i)|\big] \approx \frac{D_n^2}{\Delta f T}\left(\frac{3q^2}{2} + 4q + 1\right),\tag{22}$$

where $q = D_\xi / D_n$ is signal-to-noise ratio.

It follows from Eq. (22) that Estimator (2), generated by a formant-modulation method, is consistent.

It will be useful for engineering applications next formulas for relative variances, which describe an assessment relative error:

$$\frac{D\big[|A(0)|\big]}{\langle A(0)\rangle^2} \approx \frac{1}{\Delta f T}\frac{\big(3q^2/2 + 4q + 1\big)}{(1+q)^2}, \qquad \frac{D\big[|A(F_i)|\big]}{\big\langle |A(F_i)|\big\rangle^2} \approx \frac{1}{\Delta f T}\frac{4\big(3q^2/2 + 4q + 1\big)}{q^2}\tag{23}$$

The diagrams of relative variances (23) for $\Delta f = 90$ Hz ($f_0 = 125$ Hz) and $T = 16$ s are represented in Fig. 1.

Investigations of cross correlation of magnitudes $A_{ik}(0)$ and $A_{ik}(F_i)$ are also accompanied with the rather bulky calculations, therefore we only bring final results.
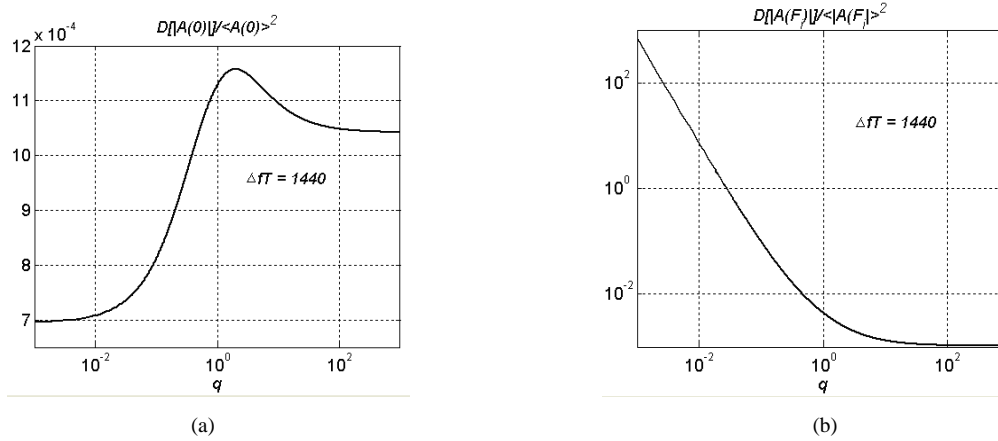

Fig. 1 Relative variances of $A_{ik}(0)$ (a) and $A_{ik}(F_i)$ (b)

For small ($q \ll 1$) signal-to-noise ratios, at implementation of conditions $T\Delta f \gg 1$ and $T = r/F_i$, the approximate equality may be written:

$$K\big\{A(0)A^*(F_i)\big\} \approx K\big\{A(F_i)A^*(0)\big\} \approx 0,\tag{24}$$

i.e. magnitudes $A_{ik}(0)$ and $A_{ik}(F_i)$ are practically uncorrelated (* is symbol of complex conjugate). And due to the central limit theorem it is possible to consider these magnitudes as statistically independent.

For large ($q \gg 1$) signal-to-noise ratios the approximate equality may be written:

$$K\big\{A(0)A^*(F_i)\big\} \approx K\big\{A(F_i)A^*(0)\big\} \approx \frac{D_\xi^2}{\Delta f T}.\tag{25}$$

It follows from Eqs. (25) and (22) that magnitudes $A_{ik}(0)$ and $A_{ik}(F_i)$ are statistically dependent, and factor of their cross correlation is close to magnitude 0.66:

$$R\big\{A(0)A^*(F_i)\big\} = \frac{K\big\{A(0)A^*(F_i)\big\}}{\sigma[A(0)]\sigma\big[|A^*(F_i)|\big]} \approx \frac{2}{3}.\tag{26}$$

Despite of correlation of magnitudes $|A(0)|$ and $|A(F_i)|$, the conclusion about a consistency of an Estimator (2), generated by a formant-modulation method, remains in a force.

In view of weak correlation of magnitudes $|A(0)|$ and $|A(F_i)|$, it follows from Eq. (13) that under condition $T = r/F_i$

$$\langle \tilde{m}_k(F_i) \rangle \approx \frac{D_\xi}{D_\xi + D_n} = \frac{q}{1+q} = m \, , \tag{27}$$

i.e. estimator of modulation coefficient is unbiased.

Expectation of modulation coefficient estimator square is:

$$\langle \tilde{m}_k^2(F_i) \rangle \approx 4 \frac{\langle |A(F_i)|^2 \rangle}{\langle |A(0)|^2 \rangle} = 4 \frac{D[A(F_i)] + \langle |A(F_i)| \rangle^2}{D[A(0)] + \langle |A(0)| \rangle^2} \, . \tag{28}$$

Taking into account Eqs. (17), (18) and (22), we can get from Eqs. (27) and (28) the expression for modulation coefficient estimator variance:

$$D[\tilde{m}] = 4 \frac{\frac{1}{\Delta fT}\left(\frac{3q^2}{2} + 4q + 1\right) + \frac{q^2}{4}}{\frac{1}{\Delta fT}\left(\frac{3q^2}{2} + 4q + 1\right) + (1+q)^2} - \frac{q^2}{(1+q)^2} \, . \tag{29}$$

It is easy to find from (29):

$$\lim_{q \to 0} D[\tilde{m}] \approx \frac{4}{\Delta fT} \, , \quad \lim_{q \to \infty} D[\tilde{m}] \approx \frac{4,5}{\Delta fT} \, , \quad D[\tilde{m}]\big|_{q=1} \approx \frac{6}{\Delta fT} \, ,$$

therefore it will be convenient to use in engineering applications the following simple formula:

$$D[\tilde{m}_k] \leq \frac{6}{\Delta f_k T} \, . \tag{30}$$

Due to unbiasedness of an estimator $\tilde{m}$ (under condition $T = r/F_i$), signal-to-noise ratio estimator

$$\tilde{SNR} = 10 \lg \frac{\tilde{m}}{1 - \tilde{m}} \, . \tag{31}$$

is unbiased too under the same condition.

As we can see from (30), variance $D[\tilde{m}_k]$ is small for $\Delta f_k T \gg 1$, so we can get next approximately true equation:

$$D[\tilde{SNR}_k] \approx \left(\frac{df(m)}{dm}\right)^2 D[\tilde{m}_k] \, ; \tag{32}$$

$$f(m) = SNR = 10 \lg \frac{m}{1 - m} \, .$$

Taking into account (30) and

$$\frac{df(m)}{dm} = \frac{4,34}{m(1-m)} = \frac{4,34\left(1 + 10^{-0,1 \cdot SNR}\right)^2}{10^{-0,1 \cdot SNR}} \, ,$$

we derive suitable for engineering applications equation:

$$D[\tilde{SNR}_k] \leq \frac{113\left(1 + 10^{-0,1 \cdot SNR}\right)^4}{10^{-0,2 \cdot SNR}} \cdot \frac{1}{\Delta f_k T} \, . \tag{33}$$

It may be shown for Eq. (3) [10]:

$$D[\tilde{q}_k] \approx \frac{38}{\Delta f_k T} \, . \tag{34}$$

When comparing Eqs. (33) and (34), we can suggest that formant method may be preferable for case of prevailing influence of noise. Computer modelling shows truth of the suggestion. Relative error of formant method word intelligibility

estimator (Fig 2,a) is about 0.15-0.20 of the formant-modulation method error (Fig 2,b). Formant method estimation time is about 1/14 of formant-modulation method estimation time.
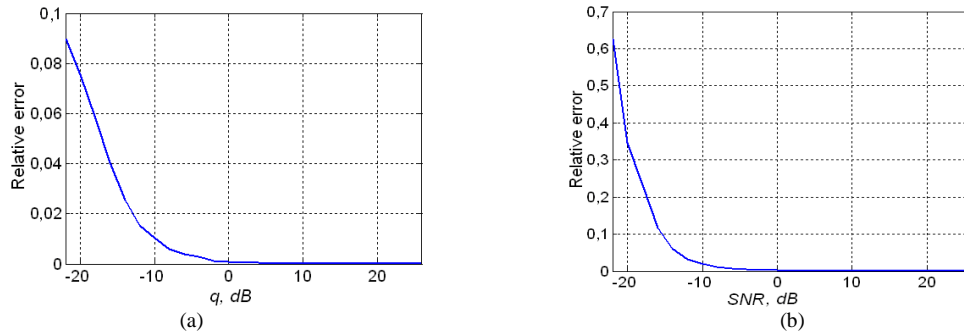

Fig. 2 Relative errors of word intelligibility estimators for formant method (a) and formant-modulation method (b)

B. Prevailing Influence of Reverberation

When considering the case of prevailing influence of reverberation, the k-th frequency channel signal pattern can be presented as

$$y_{ik}(t) = \int_{-\infty}^{\infty} h_{ek}(v)x_i(t-v)dv, \tag{35}$$

where $x_i(t) = \xi(t)\sqrt{f_i(t)}$; $h_{ek}(v) = \int_{-\infty}^{\infty} h(v)h_{k0}(z-v)dv$; $h(v)$ is room impulse response and $h_{k0}(v)$ is k-th band-pass filter impulse response.

Expectation of $A_{ik}(F_i)$ is:

$$\langle A(F_i) \rangle = \frac{1}{T} \int_0^T \langle y_{ik}^2(t) \rangle e^{-j2\pi F_i t} dt, \tag{36}$$

$$\langle y_{ik}^2(t) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_{ek}(t-\tau_1)h_{ek}(t-\tau_2)\langle \xi(\tau_1)\xi(\tau_2) \rangle \sqrt{f_i(\tau_1)f_i(\tau_2)} d\tau_1 d\tau_2. \tag{37}$$

Process $\xi(t)$ in Eq. (37) is band-pass ($0 \le f \le \Delta f$) white noise with variance D_ξ, therefore

$$\langle \xi(t)\xi(t+\tau) \rangle = D_\xi \cdot Sa(2\pi\Delta f\tau) \approx \frac{D_\xi}{2\Delta f}\delta(\tau), \tag{38}$$

where $\delta(\tau)$ is Dirac delta function. We get from Eqs. (37) and (38), under condition $T = r/F_i$:

$$\langle A_k(F_i) \rangle \approx \frac{D_\xi}{2\Delta f} \int_{-\infty}^{\infty} h_{ek}^2(v)e^{-j2\pi F_i v} dv. \tag{39}$$

It follows from Eq. (39) in the specific case $F_i = 0$:

$$\langle A_k(0) \rangle \approx \frac{D_\xi}{2\Delta f} \int_{-\infty}^{\infty} h_{ek}^2(v) dv. \tag{40}$$

Now let's analyze the variance of $A_k(F_i)$. We can get after bulky transforms:

$$D\{|A_k(F_i)|\} =$$
$$= \frac{2}{T}\left(\frac{D_\xi}{2\Delta f}\right)^2 \left[\int_{-\infty}^{\infty} |H_{ek}(f)|^2 |H_{ek}(f+F_i)|^2 df + \frac{1}{4}\int_{f_1} H_{ek}(f_1)H_{ek}(f_1+2F_i)\left(H_{ek}^*(f_1+F_i)\right)^2 df_1 + \frac{1}{4}\int_{f_1} H_{ek}^2(f_1)H_{ek}^*(f_1-F_i)H_{ek}^*(f_1+F_i)df_1 \right], \tag{41}$$

$$D\{|A_k(0)|\} = \frac{2}{T}\left(\frac{D_\xi}{2\Delta f}\right)^2 \left[\int_{-\infty}^{\infty}|H_{ek}(f)|^4 df + \frac{1}{4}\int_f |H_{ek}(f)|^2|H_{ek}(f+F_i)|^2 df + \frac{1}{4}\int_f |H_{ek}(f)|^2|H_{ek}(f-F_i)|^2 df\right], \qquad (42)$$

where $H_{ek}(f) = \int_{-\infty}^{\infty} h_{ek}(v)e^{-j2\pi fv}dv$.

It will be more convenient to use in engineering application the expression for upper bound of variance:

$$D\{|A_k(F_i)|\} \approx D\{|A_k(0)|\} \le\le \frac{3}{T}\left(\frac{D_\xi}{2\Delta f}\right)^2 \int_{-\infty}^{\infty}|H_k(f)|^4 df = \frac{3}{T}\left(\frac{D_\xi}{2\Delta f}\right)^2 \int_{-\infty}^{\infty}B_{h_{ek}}^2(\tau)d\tau, \qquad (43)$$

$$B_{h_{ek}}(\tau) = \int_{-\infty}^{\infty} h_{ek}(t)h_{ek}(t+\tau) = \int_{-\infty}^{\infty}|H_{ek}(f)|^2 e^{j2\pi f\tau}df . \qquad (44)$$

Under the assumption of week correlation of magnitudes $|A(0)|$ and $|A(F_i)|$, it can be gotten from Eqs. (13), (39)–(40) the formulas for expectation and variance of modulation coefficient estimator:

$$\langle \tilde{m}_{ki}\rangle \approx \left|2\int_0^{\infty} h_{ek}^2(t)e^{-j2\pi F_i t}dt \middle/ \int_0^{\infty} h_{ek}^2(t)dt\right| . \qquad (45)$$

$$D[\tilde{m}_k(F_i)] < \frac{24}{\Delta f_k T} . \qquad (46)$$

It follows from Eq. (46) that estimator of modulation coefficient is consistent. When comparing Eqs. (30) and (46), we can see that upper bound of variance for reverberation case is much more than for noise case one. It is necessary to make additional research in future for getting more precise result.

There is obvious similarity between Eq. (45) and known Schroeder formula [9].

IV. COMPUTATIONAL MODELLING RESULTS

In order to reduce duration of computational modelling, it is expedient to use band-pass processes ξ_k instead of process $\xi(t)$, where ξ_k are results of filtering of $\xi(t)$ by means of octave filter bank. So, we had used expression

$$x_{ik}(t) = \xi_k(t)\sqrt{f_i(t)} \qquad (47)$$

instead of Eq. (4) for numerical modelling.

Multiplying $x_{ik}(t)$ by specially chosen coefficients a_k, we provide a similarity of long-term spectra of test and speech signals. Components $n_k(t)$ of noise $n(t)$ are generated similarly, with an only difference that coefficients b_k are used in place of coefficients a_k. It allows generating of required colored noise. Another feature of schematic representation of noise generation is absence of modulation blocks, where multiplying on $\sqrt{f_i(t)}$ is made.

Modulation coefficients estimation is implemented in accordance with Eqs. (13)–(14), and effective signal-to-noise ratio estimation is executed in accordance with Eqs. (10)–(12). Articulation intelligibility estimation is fulfilled in accordance with Eq. (2). The obtained estimation is recalculated in wordy intelligibility [3].

At first we shall consider results of modelling of a situation, when reverberation disturbance is absent, and the noise disturbance acts only. The evaluation results of a wordy intelligibility obtained by a formant-modulation method for various integrated signal-to-noise ratios and for various colored noises are indicated in Fig. 3,a. Similar diagrams for formant method are shown in Fig. 3,b. As we can see, both methods lead to practically identical results, despite of essential difference of test signals and estimation methods. It means that the developed computer model of the measuring system realizing a formant-modulation method is efficient and ensures deriving correct outcomes in case of a prevailing action of a noise disturbance.

Now we shall consider results of modelling of situation when reverberant disturbance is prevailing. Diagrams of wordy intelligibility for room with reverberation time $T_{60} = 0,6\,s$ are shown in Fig. 4,a, and analogous diagrams for room with reverberation time $T_{60} = 1\,s$ are shown in Fig. 4,b. As could be waited, increasing of reverberation time leads to speech intelligibility reducing. This reduction is most noticeable for middle and small signal-to-noise ratios. The reduction of speech

intelligibility is less noticeable for signal-to-noise ratios over 5-7 dB. It is also interesting that the degree of intelligibility reduction depends on color of noise disturbance: there is minimum reduction for brown noise and maximum reduction for white noise. That fact may be explained by means of some low-frequency filters as key block of reverberation process generation model [4, 5].
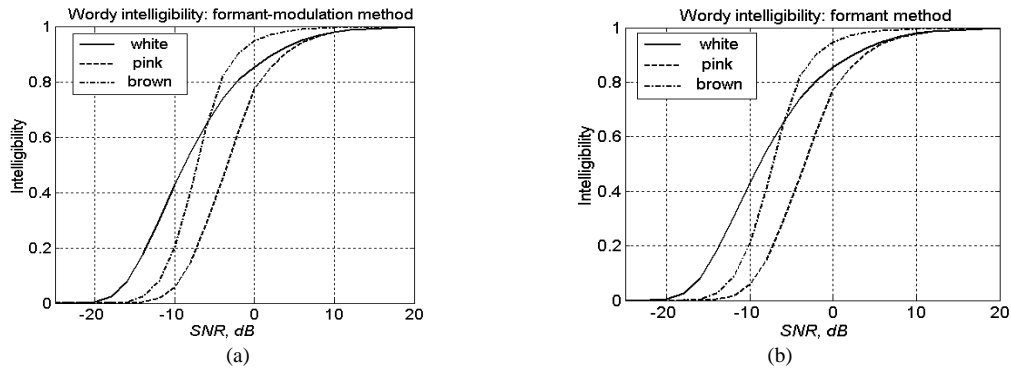

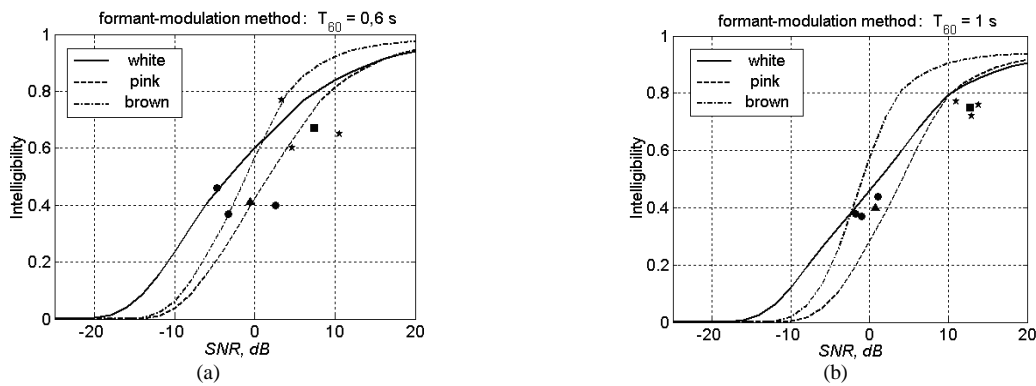Fig. 3 Intelligibility estimators for noise disturbance: formant-modulation(a) and formant (b) methods

Fig. 4 Intelligibility estimators: $T_{60} = 0.6\,s$ (a) and $T_{60} = 1\,s$ (b)

V. ARTICULATION TESTING RESULTS

The articulation testing was realized in two rooms of National Technical University of Ukraine "Kyiv Polytechnic Institute": small (80 m³) laboratory with reverberation time $T_{60} \approx 0.6\,s$ and large (1690 m³) lecture hall with reverberation time $T_{60} \approx 1\,s$. The articulation tests were conducted in the correspondence with the recommendations of GOST 25902-83 [7]. Three speakers and three auditors participated in testing, and each speaker read on 5 syllable tables (with 50 syllables in each table) from GOST 50840-95 [8] in each session.

Two experiments were realized in each room. A background noise was used as noise disturbance in the first experiment. A white noise, generated by computer and emitted by acoustic system, was added to a background noise in the second experiment. Emitting loudspeaker was placed between speaker and auditors, on a distance of 1.5 to 2 m from auditors. Undirected microphone GT57 was used for reception of a signal and the accepted signal was recorded on the second computer disk.

The features of recorded noise and speech signals are pointed in the Table 1: P_n is integrated level of a noise disturbance; P_s is integrated level of a speech signal; $SNR = P_s - P_n$ is integrated signal-to-noise ratio; S is syllable intelligibility.

The magnitudes S appropriated to outcomes of the separate speakers in Experiments 1 and 3 (background noise), are represented in Fig. 4 by asterisks, and the outcomes of their averaging are represented by a small square. The magnitudes S appropriated to outcomes of the separate speakers in Experiments 2 and 4 (background noise + added noise), are represented in Fig. 4 by circles, and the outcomes of their averaging are represented by a small triangle.

TABLE 1 FEATURES OF SIGNALS AND NOISES

T_{60}, s	Noise	P_n, dB	P_s, dB	SNR, dB	S
0.6	background	-29	-21.6	7.4	0.67
0.6	background +added	-21	-21.6	-0.6	0.41
1	background	-41	-28.3	12.7	0.75
1	background +added	-29	-28.3	0.7	0.40

It is obvious that there is a good enough concordance of model and experimental research results.

VI. CONCLUSIONS

Efficiency of proposed formant-modulation method is confirmed by analytical and experimental investigation results of the paper. Equation $T = r/F_i$, where r is arbitrary positive integer, is a condition of an unbiased intelligibility estimator for formant-modulation method. Formant-modulation estimator is consistent, e.g. its variance tends to zero when estimation time T tends to infinity.

Analytical researches and computer modelling show that formant-modulation estimator is worse compared to formant method for case of prevailing influence of noise: relative error of formant method word intelligibility estimator is about 0.15-0.20 of the formant-modulation method error and formant method estimation time is about 1/14 of formant-modulation method estimation time. Analytical researches show that upper bound of variance for reverberation case is much more than one for noise case. This result is preliminary and it is necessary to make additional research in future for getting more precise result.

REFERENCES

[1] A. Prodeus, "On Some Evolution Features of Objective Methods of Speech Intelligibility Measurements," Electronics and Communication, thematic issue "Electronics and Nanotechnology," vol. 2, pp. 217-223, 2010.

[2] A. Prodeus, "On Possibility of Advantages Join of Formant and Modulation Methods of Speech Intelligibility Evaluation," in *Proceedings of the VI International Conference MEMSTECH, Lviv, Polyana*, pp. 254-259, 2010.

[3] N.B. Pokrovskiy, *Calculation and Measurement of Speech Intelligibility*. Moscow, USSR: Sviazizdat, 1962.

[4] H.J.M. Steeneken and T. Houtgast, "Basics of the STI-measuring method" [Online]. Available: http://www.steeneken.nl/sti/

[5] H.J.M. Steeneken and T. Houtgast, "RASTI: A Tool for Evaluating Auditoria" [Online]. Available: http://www.bksv.com/doc/

[6] A.N. Prodeus, V.S. Didkovskiy and M.V. Didkovskaia, *Acoustic Examination of Speech Communications Channels*. Kyiv, Ukraine: Imex-LTD, 2008.

[7] GOST 25902-83. *Auditoriums. Methods of determination of speech intelligibility*. Moscow, USSR: Izdatelstvo standartov, 1983.

[8] GOST 50840-95. *Speech Transmission on Communication Channels. Methods of Evaluation of Quality, Intelligibility and Cognizance*. Moscow, USSR: Izdatelstvo standartov, 1996.

[9] K. Jacob, "Correlation of Speech Intelligibility Tests in Reverberant Rooms with Three Predictive Algorithms," *J. Audio Eng. Soc.*, vol. 37, iss. 12, pp. 1020-1030, Dec. 1989.

[10] A. Prodeus, L. Dronzhevskaia, V. Klimkov and D. Shagitova. "Formant and Formant-Modulation Methods of Speech Intelligibility Assessment. Part 2. Assessment Exactness and Speed," *Electronics and Communications,* vol. 64, iss. 6, pp. 16-24, 2011.

Arkadiy N. Prodeus is professor at the Acoustics and Acoustoelectronics Department, National Technical University of Ukraine "Kyiv Polytechnic Institute" (NTUU "KPI"). He received an Engineering Diploma from the NTUU "KPI", Kyiv, Ukraine, PhD from the State Scientific Research Institute of Hydrodevices, Kyiv, and a DSc from the NTUU "KPI", Kyiv. He is the first author of the book *Expert Systems in Medicine* published in 1998. He is also the first author of the book *Acoustic Examination of Speech Communications Channels* published in 2008 and the first author of the book *Computer Processing of Acoustic Signals* published in 2010. He has published numerous scientific papers and patents.