

А.М. Продеус, В.О. Клімков
 Національний технічний університет України
 “Київський політехнічний інститут”,
 кафедра акустики та акустoeлектроніки

ПРО ВПЛИВ РЕВЕРБЕРАЦІЙНОЇ ПЕРЕШКОДИ НА ТОЧНІСТЬ ВИМІРЮВАННЯ ФУНКЦІЇ РОЗПОДІЛУ РІВНІВ МОВНОГО СИГНАЛУ

© Продеус А.М., Клімков В.О., 2010

Виконано аналітичні й експериментальні дослідження впливу ревербераційної перешкоди на точність вимірювання функції розподілу рівнів мовного сигналу.

Ключові слова: ревербераційна перешкода, функція розподілу, мовний сигнал.

Analytical and experimental investigations of influence of reverberation disturbances on measuring exactness of speech signal levels distributing function has been made.

Keywords: reverberation disturbances, distributing function, speech signal.

1. Вступ

Нині серед методів акустичної експертизи приміщень і ліній зв'язку найпопулярнішими є два об'єктивні (інструментальні) методи – формантний і модуляційний [1]. Кожний із цих методів має певні переваги й недоліки.

Важливою перевагою формантного методу є вибір показника якості у вигляді середньої ймовірності сприйняття формант:

$$A = \sum_{k=1}^K p_k \cdot w_k(E_k), \quad (1)$$

де p_k – ймовірність перебування формант у k -й смузі частот:

$$p_k = F_1(f_{вк}) - F_1(f_{нк}), \quad (2)$$

де $F_1(f)$ – функція розподілу формант за частотою; $w_k(E_k)$ – коефіцієнт сприйняття мови; E_k – відношення сигнал-шум у k -й частотній смузі. Оскільки формантну розбірливість можна перерахувати в ймовірність правильного розпізнавання звуків, складів, слів і навіть фраз, такий вибір є надзвичайно зручним, тому що за використання суб'єктивних методів оцінки розбірливості мови показником якості також слугує ймовірність правильного розпізнавання звукосполучень, складів, слів або фраз.

Разом з тим, формантний метод має деякі недоліки, головними серед яких є неоднозначність визначення коефіцієнта сприйняття для російської мови й тісно пов'язані із цією обставиною проблеми визначення коефіцієнта сприйняття для української мови, а також принципові труднощі врахування дії ревербераційної перешкоди [1].

Чудовою властивістю модуляційного методу є можливість одночасно враховувати вплив шумової й ревербераційної перешкод. Досягається це використанням спеціального тестового сигналу у вигляді модульованого шуму, з подальшим вимірюванням коефіцієнта модуляції й

перерахуванні його в так зване “ефективне” відношення сигнал-шум, що дає змогу врахувати вплив не тільки шумової, але й ревербераційної перешкоди.

Недоліком модуляційного методу є те, що мірою розбірливості є не ймовірність правильного сприйняття мови, а індекс STI (Speech Transmission Index), так що для зіставлення результатів модуляційного й суб'єктивного методів необхідно використовувати спеціальні допоміжні графіки [1].

У роботах [2, 3] запропоновано об'єднати переваги обох методів, а саме: вимірювати формантну розбірливість відповідно до співвідношення (1) з тією лише відмінністю, що як відношення сигнал-шум E_k використовувати “ефективне” відношення сигнал-шум, вимірюване відповідно до технології модуляційного методу.

Ефективність нового “формантно-модуляційного” методу, як, втім, і класичного форматного методу, значною мірою залежить від якості попередніх досліджень функцій і параметрів, що входять у співвідношення (1). Зокрема, ефективність нового методу залежить від точності попередніх вимірювань функції розподілу рівнів мовного сигналу $F(x)$, оскільки коефіцієнт сприйняття $w(E)$ пов'язаний з функцією $F(x)$ простим співвідношенням [1]:

$$w(E) = 1 - F(-E).$$

У роботі [4] отримано попередні оцінки характеру й ступеня впливу шумової й ревербераційної перешкод на точність вимірювань функції $F(x)$. Мета цієї роботи полягає у поглибленому дослідженні впливу ревербераційної перешкоди на точність вимірювань функції $F(x)$.

2. Аналітичний опис впливу ревербераційної перешкоди

Як вказувалося в роботі [4], за математичну модель сигналу, спотвореного ревербераційною перешкодою, звичайно вибирають співвідношення згортки мовного сигналу $s(t)$ з імпульсною характеристикою (ІХ) приміщення $h(t)$:

$$z(t) = \int_{-\infty}^{\infty} s(v-t)h(v)dv. \quad (3)$$

Однак для аналітичних досліджень статистичних властивостей рівня мовного сигналу замість моделі (3) зручнішою є модель у вигляді адитивної суміші

$$z(t) = s(t) + r(t) \quad (4)$$

мовного сигналу $s(t)$ (“прямий сигнал”) і ревербераційної перешкоди $r(t)$ (сукупність відбитих сигналів).

Процеси $s(t)$ й $r(t)$ загалом є корельованими. Разом з тим, для досить великих приміщень, коли інтервал часу τ між приходом у точку приймання прямого і першого відбитого сигналів перевищує середню тривалість фонему, тобто $\tau > 0,1 \dots 0,2$ с, процеси $s(t)$ й $r(t)$ можна вважати практично некорельованими в кожний момент часу t . У цьому випадку оцінку рівня ζ сигналу $z(t)$ можна подати приблизно у такому вигляді:

$$\begin{aligned} \tilde{\zeta}(t) &= \frac{1}{T} \int_{t-T}^t z^2(t)dt = \frac{1}{T} \int_{t-T}^t [s^2(t) + 2s(t)r(t) + r^2(t)]dt \approx \\ &\approx \frac{1}{T} \int_{t-T}^t s^2(t)dt + \frac{1}{T} \int_{t-T}^t r^2(t)dt = \tilde{D}_s(t) + \tilde{D}_r(t) \end{aligned} \quad (5)$$

де $\tilde{D}_s(t)$ й $\tilde{D}_r(t)$ – оцінки дисперсій сигналу й ревербераційної перешкоди, відповідно.

Внаслідок некорельованості у кожний момент часу t процесів $s(t)$ і $r(t)$ щільність розподілу $W'_{\zeta}(x')$ процесу $\zeta(t)$ являє собою згортку [5]:

$$W'_{\zeta}(x') = \int_{-\infty}^{\infty} W'_{\tilde{D}_r}(y') W'_{\tilde{D}_s}(x'-y') dy', \quad (6)$$

де $W'_{\tilde{D}_s}(x')$ й $W'_{\tilde{D}_r}(x')$ – щільності розподілу оцінок $\tilde{D}_s(t)$ і $\tilde{D}_r(t)$, відповідно. Штрихи при символах змінних $x' = x/x_0$ і $y' = y/x_0$ символах функцій W' означають, що в співвідношенні (6) фігурують розподіли нормованих, за середнім значенням x_0 рівня мовного сигналу, рівнів сигналу й перешкоди.

Для функції розподілу $F'_{\zeta}(x')$ процесу $\zeta(t)$ з (6) неважко одержати:

$$F'_{\zeta}(x') = \int_{-\infty}^{\infty} W'_{\tilde{D}_r}(y') F'_{\tilde{D}_s}(x'-y') dy'. \quad (7)$$

Якщо обмежитися розглядом тільки першого відбиття (що відповідає реверберації в приміщеннях з доволі великим фондом звукопоглинання), тоді

$$W'_{\tilde{D}_r}(y') \approx a W'_{\tilde{D}_s}(ay'), \quad (8)$$

де a – відношення середніх значень рівня сигналу й перешкоди (відношення сигнал-шум у разях за потужністю), і співвідношення (7) набуває вигляду:

$$F'_{\zeta}(x') \approx a \int_{-\infty}^{\infty} W'_{\tilde{D}_s}(ay') F'_{\tilde{D}_s}(x'-y') dy'. \quad (9)$$

Значення a можна зв'язати з індексом чіткості C_{50} , що характеризує відношення енергій прямого й відбитого звуків мовного сигналу [3]:

$$C_{50} = 10 \lg \frac{\int_{0,05}^{0,05} p^2(t) dt}{\int_{0,05}^{\infty} p^2(t) dt} = 10 \lg a = \text{SNR}, \quad (10)$$

де $p(t)$ – тиск; 0,05 – значення фіксованого моменту часу, с; SNR – відношення сигнал-перешкода, дБ.

Приймаючи $p^2(t) = p \cdot e^{-\alpha t}$, одержимо:

$$C_{50} = 10 \lg(e^{0,05\alpha} - 1). \quad (11)$$

Оскільки $\alpha = \frac{13,8}{T_{60}}$, де T_{60} – час реверберації, з (10) з урахуванням (11) отримуємо:

$$a = e^{\frac{0,69}{T_{60}}} - 1. \quad (12)$$

У табл. 1 наведено кілька значень залежності $a = a(T_{60})$, описуваної співвідношенням (12), що дають змогу дійти попереднього висновку, що за часу реверберації $T_{60} < 0,3$ с ревербераційна перешкода не повинна істотно позначитися на формі функції розподілу рівня мовного сигналу.

Залежність $a = a(T_{60})$

T_{60}	0,1	0,3	0,7	1	2	3
a	991,27	8,97	1,67	0,99	0,41	0,26
C_{50}	30	9,53	2,23	0	-3,87	-5,85

Для перевірки справедливості цього висновку виконаємо обчислення відповідно до співвідношень (9)–(12), вважаючи, що розподіл рівня мовного сигналу моделі можна апроксимувати гамма-розподілом [6]:

$$W(x) = \begin{cases} 0, & x \leq 0; \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), & x > 0 \end{cases} \quad (13)$$

Справді, приймаючи у співвідношенні (12) $\alpha = \beta = 0,7$ (математичне очікування випадкової величини X у такому разі дорівнює $m_x = \frac{\alpha}{\beta} = 1$), одержимо графік $W(x)$ (рис. 1, а), що мало відрізняється від оцінки функції $W(x)$ реального мовного сигналу (рис. 1, б) [1].

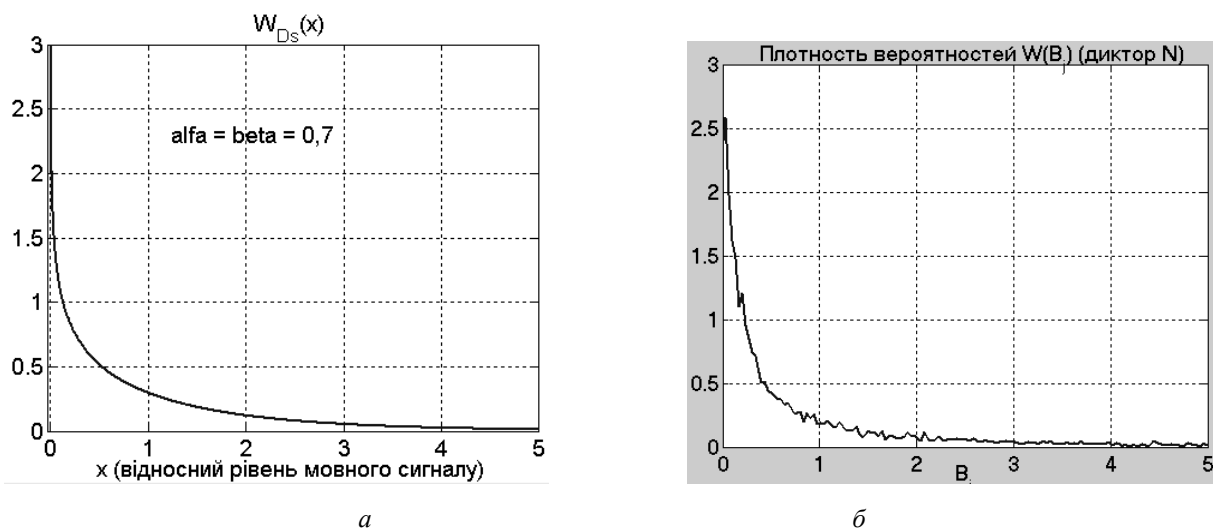


Рис. 1. Гамма-розподіл для $\alpha = \beta = 0,7$ (а) і оцінка функції $W(x)$ для реального мовного сигналу (б)

Результати обчислень відповідно до співвідношень (9)–(13) наведено на рис. 2. Очевидно, для $T_{60} \approx 0,1 \dots 0,2$ с впливом реверберації справді можна практично знехтувати (рис. 2, а, б), оскільки викликана цим впливом зміна функції розподілу не перевищує 0,03. Якщо $T_{60} \approx 0,3$ с, зміна функції розподілу, викликана впливом реверберації, досягає 0,1 (рис. 2, в). З подальшим зростанням T_{60} функція розподілу рівнів мовного сигналу усе сильніше зміщується праворуч, причому зрушення більшою мірою є вираженим в області малих та середніх рівнів сигналу (рис. 2, г, д). Якщо $T_{60} \approx 3$ с, зміна функції розподілу рівнів мовного сигналу досягає 0,5 (рис. 2, е).

Варто враховувати, що подані на рис. 2 результати мають наближений характер через допущення, що передували виведенню співвідношень (4) і (9), а також через наближеність апроксимації функції розподілу мовного сигналу співвідношенням (13).

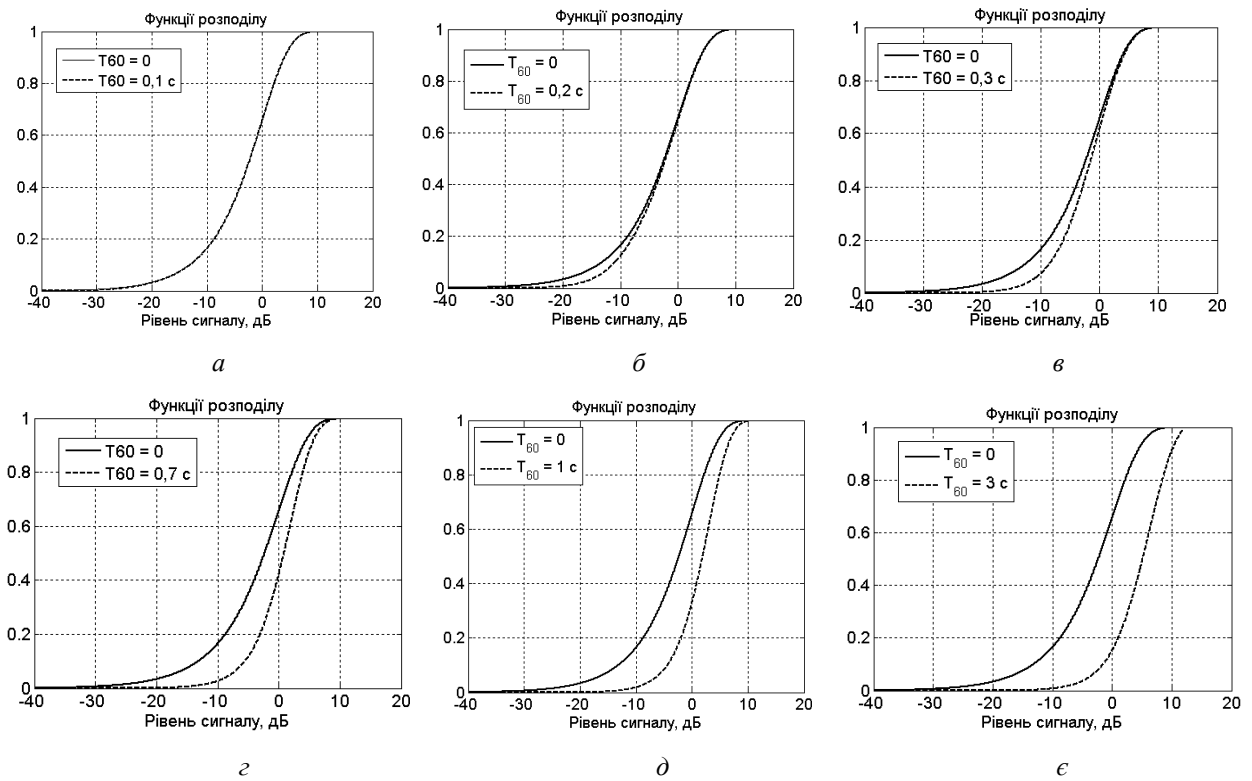


Рис. 2. Результати розрахунку $F'_\zeta(x')$ для різних T_{60}

3. Експериментальні дослідження

Для точнішої оцінки ступеня впливу ревербераційної перешкоди на вид функції $F(x)$ використаємо співвідношення (3), згорнувши мовний сигнал $s(t)$ з експериментально отриманими імпульсними характеристиками $h(t)$ приміщень із різним часом реверберації T_{60} .

Для детальнішого аналізу вихідний мовний сигнал $s(t)$ і мовний сигнал $z(t)$, спотворений ревербераційною перешкодою, зазнавав фільтрації гребінкою із семи октавних фільтрів із центральними частотами 125, 250, 500, 1000, 2000, 4000 і 8000 Гц (алгоритм фільтрації описано в [1]). Це дало змогу оцінити функцію розподілу мовного сигналу $F(x)$ для кожної з вказаних октавних смуг частот і, завдяки цьому урахувати частотні властивості ревербераційної перешкоди.

Говорячи про частотні властивості ревербераційної перешкоди, зазначимо, що доволі типовою є тенденція до зменшення часу реверберації зі зростанням частоти (рис. 3).

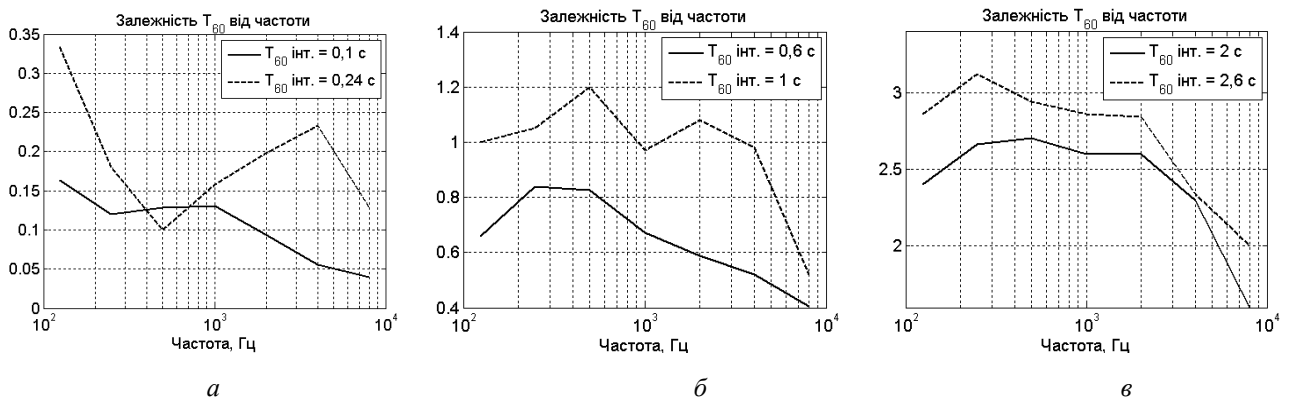


Рис. 3. Залежність часу реверберації від частоти

Оцінювання функції розподілу мовного сигналу $F(x)$ виконувалося за алгоритмами, описаними в роботі [1]. Результати оцінювання наведено на рис. 4, 5, де суцільною лінією показана оцінка $F(x)$ для мовного сигналу за відсутності ревербераційної перешкоди ($T_{60} = 0$), а пунктирною й штрихпунктирною лініями – оцінки $F(x)$ для мовного сигналу, спотвореного ревербераційною перешкодою.

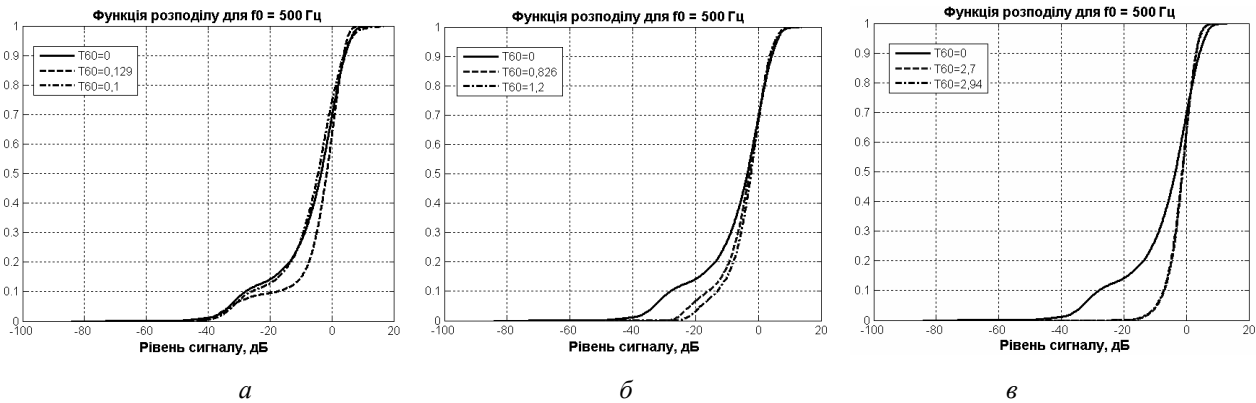


Рис. 4. Функція розподілу для $f_0 = 500$ Гц

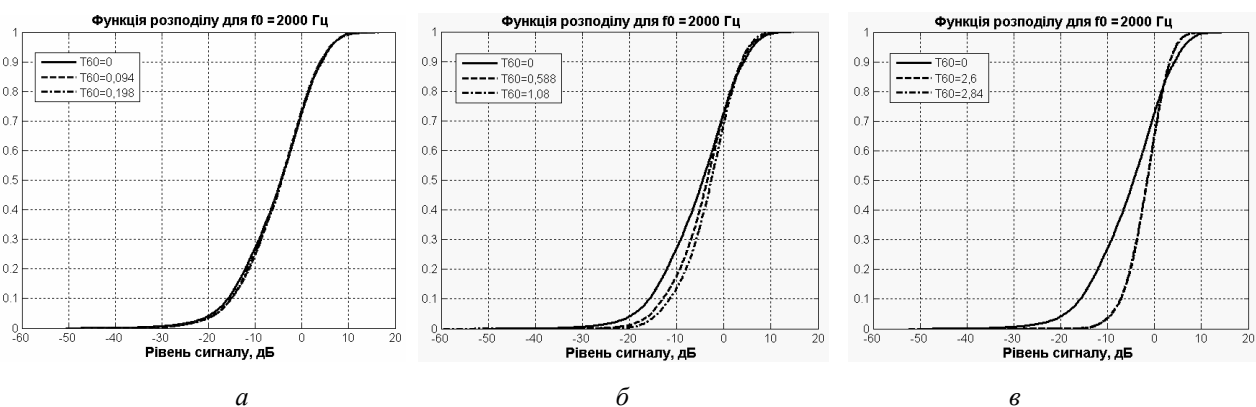


Рис. 5. Функція розподілу для $f_0 = 2000$ Гц

Порівнюючи результати аналітичних і експериментальних досліджень, зазначимо, що якісно їхні результати узгоджуються непогано. Практично у всіх випадках виконується правило: оцінка $F(x)$ зміщується праворуч тим сильніше, чим більшим є час реверберації, тобто чим меншим є відношення сигнал-шум для ревербераційної перешкоди. Крім того, експериментальні дослідження підтверджують результати аналітичних досліджень, згідно з якими для $T_{60} \approx 0,1 \dots 0,2$ с впливом реверберації можна практично зехтувати.

Разом з тим, наші експериментальні дослідження свідчать, що для часу реверберації $0,1 \dots 3$ с вплив ревербераційної перешкоди на оцінку функції $F(x)$ позначається переважно на рівнях мовного сигналу, що є меншими за середній рівень мовного сигналу.

Висновки

Виконано аналітичні й експериментальні дослідження впливу ревербераційної перешкоди на точність вимірювань функції розподілу $F(x)$ рівнів мовного сигналу.

Показано, що оцінка $F(x)$ зміщується праворуч (у бік більших рівнів сигналу) тим сильніше, чим більший час реверберації, тобто чим менше відношення сигнал-шум для ревербераційної перешкоди. Для часу реверберації $T_{60} < 0,3$ с зсув оцінки $F(x)$ є настільки незначним, що впливом ревербераційної перешкоди можна практично знехтувати.

Показано також, що для приміщень із часом реверберації понад 0,3 с вплив ревербераційної перешкоди проявляється в появі помітного негативного зміщення оцінки функції $F(x)$ для негативних значень аргументу функції $F(x)$, і незначного позитивного зміщення в області позитивних значень аргументу функції $F(x)$. Інакше кажучи, “права” гілка оцінки функції $F(x)$ порівняно мало піддається впливу ревербераційної перешкоди.

Отримані результати є доволі важливими як для теорії, так і для практики акустичної експертизи приміщень і каналів зв'язку, оскільки дають підстави для обґрунтованих висновків як про якість спеціальних заглушених приміщень (“акустичних кімнат”), так і про вірогідність оцінок коефіцієнта сприйняття $w(E)$, сформованих з використанням таких приміщень.

1. Дідковський В.С., Дідковська М.В., Продеус А.М. Акустична експертиза каналів мовної комунікації: Монографія. – К: Імекс-Лтд, 2008. – 420 с.
2. Продеус А.М. Про деякі особливості розвитку об'єктивних методів вимірювань розбірливості мови // Електроніка й зв'язок, тематичний випуск “Електроніка й нанотехнології”. – 2010. – № 2. – С. 217–223.
3. Prodeus A. On Possibility of Advantages Join of Formant and Modulation Methods of Speech Intelligibility Evaluation // Proceedings of the VI International Conference MEMSTECH 2010. – Lviv, Polyana, 2010. – P. 254–259.
4. Дідковський В.С., Продеус А.М., Рудь Д.П. Про точність вимірювань функції розподілу рівнів мовного сигналу. // Електроніка й зв'язок, тематичний випуск “Електроніка й нанотехнології”. – 2010. – № 3. – С. 152–159.
5. Вентцель Е.С. Теорія ймовірностей. – М.: Наука, 1969. – 575 с.
6. Корн Г., Корн Т. Довідник по математиці для науковців і інженерів. – М.: Наука, 1970. – 720 с.