

Суб'єктивне оцінювання розбірливості мови на тлі шуму та реверберації

Продеус А. М., д.т.н. проф., ORCID [0000-0001-7640-0850](https://orcid.org/0000-0001-7640-0850)

e-mail om.prodeus@aae.kpi.ua

Вітик А. В., ORCID [0000-0002-6652-4152](https://orcid.org/0000-0002-6652-4152)

e-mail andrvt8@gmail.com

Дворник О. О., ORCID [0000-0003-4735-2225](https://orcid.org/0000-0003-4735-2225)

e-mail alexanderdvornyk@gmail.com

Котвицький І. В., ORCID [0000-0003-2800-0080](https://orcid.org/0000-0003-2800-0080)

e-mail igorktvzk@gmail.com

Чайка О. С., ORCID [0000-0002-2501-595X](https://orcid.org/0000-0002-2501-595X)

e-mail sanyarouge@gmail.com

Ярошенко М. О., ORCID [0000-0001-6327-148X](https://orcid.org/0000-0001-6327-148X)

e-mail yaroshenom13@gmail.com

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського», kpi.ua

Київ, Україна

Реферат—В даній роботі представлено результати суб'єктивного оцінювання, здійснюваного шляхом артикуляційних випробувань, розбірливості односкладових звукосполучень на тлі шуму та реверберації. Оцінювання здійснювалося за допомогою спеціально розробленого програмного забезпечення, що дозволило автоматизувати й таким чином суттєво полегшити та пришвидшити процедуру артикуляційних випробувань. За результатами випробувань маскувальна здатність білого шуму виявилася кращою за таку для коричневого шуму при відношенні сигнал-шум, менших за мінус 5 дБ, що не повністю узгоджується із попередніми прогнозними оцінками. Крім того, виявилось, що слухання мови, спотвореної шумом, через акустичні монітори може призводити до суттєвого підвищення (до 0,85-0,93) оцінок розбірливості мови, порівняно із випадком слухання через навушники (0,1-0,3). Аналогічні результати одержано для ревербераційної завади: для часу реверберації 2,7 с розбірливість збільшилася із 0,65 до 0,94. Даний феномен можна в значній мірі пояснити дією ранніх відбиттів звуку в приміщеннях, наявністю двох джерел випромінювання та бінауральним прослуховуванням. Додатковими причинами можуть бути особливості психофізичного стану слухачів та розробленої автоматизованої системи артикуляційних випробувань.

Бібл. 13, рис. 5.

Ключові слова — розбірливість мови; суб'єктивне оцінювання; артикуляційні випробування; шум; реверберація; односкладові звукосполучення.

I. Вступ

На сьогодні існує два підходи до вимірювання розбірливості мови в лініях зв'язку: суб'єктивний та об'єктивний (інструментальний) [1], [2].

Останнім часом, з огляду на економію часових, фінансових та людських ресурсів, значну увагу приділяють об'єктивному підходу, який дозволяє автоматизувати процедуру вимірювань. Проте, незважаючи на недоліки суб'єктивного підходу, його продовжують використовувати, оскільки одержані за його допомогою результати в подальшому можуть бути використані для калібрування об'єктивних систем вимірювання.

Тому на сьогодні актуальним є завдання автоматизації артикуляційних випробувань (рис. 1), що

дозволить максимально пришвидшити та полегшити процедуру таких випробувань.

Російський стандарт ГОСТ Р 50840-95 припускає використання комп'ютерної техніки та регламентує проведення артикуляційних випробувань за схемою рис. 1б. На жаль, українські інженери не можуть скористатися цим стандартом за відсутністю відповідного програмного забезпечення й тому змушені користуватися застарілим стандартом СРСР ГОСТ 16600-72.

Кроком на шляху до створення вітчизняного державного стандарту, що відповідає сучасним вимогам, можна вважати створення та випробування перших артикуляційних таблиць українських слів [3], [4]. Одним із недоліків вказаних праць є відсутність про-



ївський політехнічний інститут імені Ігоря Сікорського». Для записів використовувалися зовнішня звукова карта PRESONUS AudioBox USB, мікрофон Superlux ECM 999 та програмний звуковий редактор Audacity. Запис виконувався із частотою дискретизації 22050 Гц та бітовою глибиною 16 біт. Звукосполучення зачитувалися із застосуванням словесного оточення. Наприклад, звукосполучення «няк» зачитувалося як «Запишіть няк тепер».

Прослуховування відбувалося для чотирьох ситуацій:

- чиста мова;
- мова, спотворена шумом;
- мова, спотворена реверберацією;
- мова, спотворена спільною дією шуму та реверберації.

В першому випадку слухачами почергово прослуховувалися записи звукосполучень 3-х артикуляційних таблиць, кожна з яких містила 50 звукосполучень.

В другому випадку прослуховувалися звукосполучення, спотворені адитивним шумом із відношеннями сигнал-шум -10, 0 та +10 дБ. При цьому використовувалися моделі білого, рожевого та коричневого шумів, маскувальна властивість яких вважається досить добре вивченою [1].

В третьому випадку моделювалася, шляхом згортки сигналів чистої мови із імпульсними характеристиками різних приміщень, реверберована мова для часів реверберації від 0,3 до 2,7 с, а в четвертому випадку розглядалася спільна дія рожевого шуму та реверберації.

Моделювання сигналів, спотворених адитивним шумом та реверберацією, виконувалося за загальним алгоритмом:

$$y(t) = x(t) \otimes h(t) + n(t) \quad (1)$$

де $x(t)$ є мовний сигнал, $h(t)$ є імпульсна характеристика приміщення, $n(t)$ є адитивний шум, \otimes є символ згортки. В окремому випадку дії виключно шумової завади алгоритм (1) можна спростити:

$$y(t) = x(t) + k \cdot n(t), \quad (2)$$

де $k = 10^{0.05(SNR-SNR_0)}$, $SNR = 10 \lg(D_x/D_n)$ є «початкове» значення відношення сигнал-шум для чистої мови $x(t)$ та синтезованого «початкового» шуму $n(t)$, D_x та D_n є дисперсіями сигналу $x(t)$ та шуму $n(t)$ відповідно, SNR_0 є бажане SNR для суміші (2).

В окремому випадку сумісної дії шуму та реверберації алгоритм (1) дещо ускладнюється з метою врахування особливостей структури імпульсної характеристики приміщення:

$$h(t) = h_e(t) + h_l(t) \quad (3)$$

$$h_e(t) = \begin{cases} h(t), & t \in 0 \dots 50 \text{ ms}; \\ 0, & t \notin 0 \dots 50 \text{ ms}, \end{cases}$$

$$h_l(t) = \begin{cases} h(t), & t > 50 \text{ ms}; \\ 0, & t \leq 50 \text{ ms}, \end{cases}$$

де $h_e(t)$ описує ранні відбиття, а $h_l(t)$ описує пізню реверберацію. Тоді «початкове» відношення сигнал-шум $SNR = 10 \lg D_x / D_n$, де D_x є дисперсія сигналу, $x_e(t) = x(t) \otimes h_e(t)$.

Прослуховування виконувалося через навушники 26 студентами середнім віком 22 роки. Три студенти із зазначеного загалу виконали роботу із відхиленнями від завдання, а саме, прослуховуючи сигнали через акустичні монітори (або через комп'ютерні колонки, які в подальшому також будемо називати акустичними моніторами). Тому одержані результати було вирішено розділити на 2 групи: прослуховування через навушники та прослуховування через акустичні монітори.

В процесі артикуляційних випробувань слухачі мали зафіксувати, за допомогою комп'ютерної клавіатури, почуте звукосполучення. Даний спосіб фіксації почутого принципово відрізняється від такого, що пропонується в стандарті 50840-95. Це зроблено свідомо з метою максимального наближення процедури й результатів артикуляційних випробувань до неавтоматизованого варіанту випробувань, який доречно вважати еталонним з огляду на його первинний характер.

Обробка одержаних результатів полягала в оцінюванні розбірливості мови шляхом підрахунку частки правильно розпізнаних звукосполучень. При цьому правильно розпізнаним вважалося таке зафіксоване звукосполучення, символічний склад якого повністю співпадав із таким хоча б для одного із двох текстових варіантів артикуляційних таблиць.

III. РЕЗУЛЬТАТИ ВИПРОБУВАНЬ ТА ЇХ ОБГОВОРЕННЯ

На рис. 4 представлено графіки усереднених оцінок розбірливості мови, спотвореної шумом та реверберацією, для випадку прослуховування через навушники.

Аналогічні результати для випадку прослуховування через акустичні монітори представлено на рис. 5.

Аналіз графіків рис. 4 та 5 свідчить про те, що маскувальні властивості шуму та реверберації залежать не тільки від відношення сигнал-шум, забарвленості шуму та часу реверберації, але й можуть суттєво залежати від способу прослуховування.

Так, у випадку шумової завади, при переході від прослуховування через навушники до прослуховування через акустичні монітори розбірливість піднялася із 0,1-0,3 до 0,85-0,93 при відношенні сигнал-шум мінус 10 дБ (стандартне відхилення оцінок розбірливості було близьким в середньому до 5%). Це означає, що відношення сигнал-шум піднялося на



17-20 дБ при прослуховуванні сигналів через акустичні системи, порівняно із прослуховуванням через навушники.

Одержаний приріст відношення сигнал-шум дивує, оскільки є надто великим, хоча сам факт зростання розбірливості є очікуваним. Дійсно, при прослуховуванні зашумленого звуку, записаного в режимі «моно», через акустичну систему, що складається із 2-х випромінювачів звуку, на вході кожного вуха сигнали прямого звуку від кожного випромінювача є практично когерентними, тоді як шуми від тих же випромінювачів можна вважати некогерентними.

Тому відношення сигнал-шум (для дії прямого звуку) на вході кожного вуха має збільшитися на 3 дБ порівняно із випадком використання навушників.

Крім того, значну роль відіграють ранні відбиття звуку в приміщеннях, аналіз ролі яких [9], [10], [11], [12] свідчить, що за рахунок дії ранніх відбиттів ефективно відношення сигнал-шум може підвищуватися ще на 6-9 дБ.

Нарешті, виграш ще приблизно на 2-3 дБ можна пояснити бінауральним прослуховуванням [12]. Просте підсумовування дії зазначених чинників дає нам виграш у 12-15 дБ. Це вже досить близько до одержаного нами виграшу 17-20 дБ, хоча не вистачає ще приблизно 3-5 дБ.

Однією із імовірних причин цієї, досить значної, різниці видається наявність особливостей організації експериментальних досліджень, а саме:

- неповне врахування особливостей української мови при створенні артикуляційних таблиць;
- обмеженість вибірки дикторів;
- недостатньо висока якість звукового варіанту артикуляційних таблиць;
- недостатня кількість текстових варіантів артикуляційних таблиць;
- особливості інтерфейсу програмного забезпечення.

Виходячи з цього, доцільно в майбутньому усунути зазначені недоліки й перевірити справедливості припущення, що прослуховування зашумлених сигналів через пару акустичних моніторів у звичайних приміщеннях не може, як правило, призводити до виграшу у відношенні сигнал-шум, вищого за 12-15 дБ.

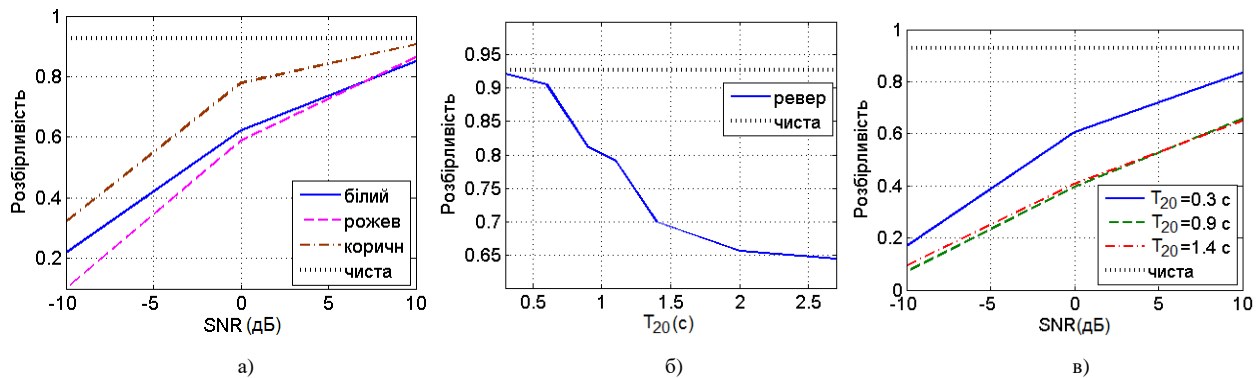


Рис. 4 Оцінки розбірливості при слуханні через навушники: шум (а), реверберація (б), рожевий шум+реверберація (в).

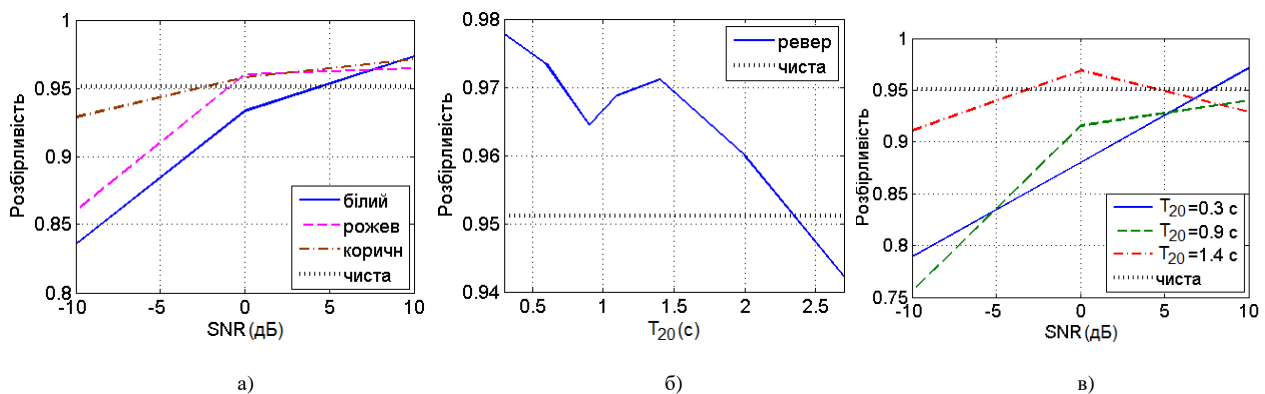


Рис. 5 Оцінки розбірливості при слуханні через акустичну систему: шум (а), реверберація (б), рожевий шум+реверберація (в)



Разом із тим слід припустити, що ще одним важливим чинником може бути психофізичний стан слухачів. Піддрунтям для такого припущення є той факт, що при пізніших повтореннях експерименту із прослуховування сигналів через акустичні монітори в одній із кімнат, де було одержано аномально високу розбірливість мови, нові оцінки розбірливості хоча й були підвищеними, проте не перевищували 0,3-0,4 для відношення сигнал-шум мінус 10 дБ. Обговорення із слухачем можливих причин цього факту показало, що перше прослуховування, де було одержано аномально високу оцінку розбірливості, відбувалося в стані значного емоційного підйому. Так чи інакше, але гіпотеза щодо зазначеного чинника має право на існування.

Нарешті, в подальшому доцільно також детально проаналізувати акустичні характеристики кімнат, в яких одержано зазначений вище аномально великий приріст розбірливості мови, з метою виявлення можливих особливостей геометрії цих кімнат.

Повертаючись до ситуації прослуховування сигналів через навушники, зазначимо, що графіки рис. 4а досить добре узгоджуються із відомими результатами [1], [2], [7], [13] в області середніх та високих значень SNR (0-10 дБ). Проте в області малих значень SNR (менших за мінус 5 дБ) маскувальна здатність білого шуму виявилася кращою за таку для коричневого шуму. Це суперечить попереднім прогностичним оцінкам [1], [2], тому доречно також в подальшому дослідити причини зазначеної невідповідності.

Що стосується дії ревербераційної завади, підтверджено очікування щодо зниження розбірливості мови із ростом часу реверберації. Разом із тим, бачимо, що прослуховування через монітори також значно підвищило розбірливість мови. Так, розбірливість зросла із 65% до 94% для часу реверберації 2,7 с, а для часу реверберації 0,3-2 с розбірливість навіть перевищила таку для чистої мови на 1-3%. Зазначимо, що в даному випадку стандартне відхилення оцінок розбірливості було досить великим й близьким до 10%, що пояснюється малою кількістю слухачів.

Аналіз оцінок розбірливості для випадку спільної дії шуму та реверберації також свідчить про суттєве зростання розбірливості мови при прослуховуванні її через акустичні монітори (стандартне відхилення оцінок розбірливості не перевищувало 10%).

Підсумовуючи, зазначимо, що за результатами експериментальних досліджень розроблену автоматизовану систему артикуляційних випробувань можна визнати працездатною. Що стосується її якості, перевірка виявила ряд певних недоліків й допомогла сформулювати відповідні рекомендації із їх усунення:

- при створенні артикуляційних таблиць слід більше уваги приділяти фонетичним особливостям української мови;
- перед записом звукових варіантів артикуляційних таблиць необхідно не тільки

відповідним чином налаштовувати апаратно-програмне забезпечення, але й ретельно інструктувати та тренувати дикторів, застерігаючи їх від підкреслення звукосполучень за допомогою пауз або збільшенням гучності;

- кількість текстових варіантів артикуляційних таблиць доцільно збільшити із двох до трьох, що полегшить врахування не тільки неоднозначності слухового сприйняття окремих фонем, але й варіанти введення із клавіатури почутих звукосполучень;
- прослуховування звукосполучень має виконуватися у випадковому порядку, що зменшить ризик запам'ятовування звукосполучень;
- необхідною є можливість корекції введених із клавіатури звукосполучень, оскільки стомлення слухача призводить до збільшення частоти помилкових натиснень клавіатури;
- доцільним є обмеження обсягу візуальної інформації, яка надається слухачеві на моніторі після прослуховування кожної артикуляційної таблиці, що дозволить зменшити ризик поліпшення результатів, зумовленого аналізом слухачем своїх помилок.

В майбутньому доцільно також розглянути ряд інших важливих питань. Зокрема, це використання професійно-спрямованих наборів артикуляційних таблиць, де вимагається використання не складів, а слів і словосполучень. Крім того, для практичного застосування є актуальним та доцільним створення комбінованих таблиць, де були б враховані особливості як української, так і російської та англійської мов, оскільки сучасна українська мова зазнає їх суттєвий вплив, особливо в таких напрямках, як інформаційні технології, проектування радіоелектронної апаратури та інших.

Зазначимо також, що, хоча питання автоматизації артикуляційних випробувань є важливим з точки зору економії фінансових, матеріальних, людських та часових ресурсів, іншим принципово важливим питанням є достовірність одержаних результатів, які в подальшому можуть бути використані для калібрування об'єктивних систем оцінювання розбірливості мови.

Висновки

Прослуховування мовленнєвих сигналів, спотворених шумом та реверберацією, виконане із застосуванням запропонованої автоматизованої системи артикуляційних випробувань, свідчить, з однієї сторони, про працездатність та досить високу якість розробленої системи. З іншої сторони, в процесі артикуляційних випробувань виявлено феномен аномально високої розбірливості мови при прослуховуванні спотворених сигналів через акустичні системи



в приміщеннях. Крім того, при малих (менших за мінус 5 дБ) відношеннях сигнал-шум виявлено неповну відповідність одержаних функціональних залежностей попереднім прогнозним оцінкам.

Аналіз можливих причин феномену аномального підвищення розбірливості мови при використанні акустичних моніторів показав, що найбільш вагомим чинником є ранні відбиття звуку в приміщеннях. Додатковий внесок доречно віднести на рахунок таких чинників, як використання двох джерел випромінювання звуку, бінауральний режим слухання, психоемоційний стан слухача й, нарешті, певну недосконалість організації артикуляційних випробувань. З огляду на останній чинник вироблено рекомендації стосовно подальшої модернізації розробленої системи артикуляційних випробувань з метою підвищення достовірності артикуляційних випробувань.

ПОДЯКА

Автори статті дякують магістрам кафедри акустики та акустoeлектроніки за участь у розробці і випробуваннях автоматизованої системи артикуляційних випробувань, а також за участь в обговоренні одержаних результатів.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] V. S. Didkovskiy, M. V. Didkovskaya, and A. N. Prodeus, Akusticheskaja ekspertiza kanalov rechevoj kommunikacii. Monografija [Acoustic examination of the channels of speech communication. Monograph], Kyiv: Imex-LTD, 2008. ISBN 978-966-8861-85-7
- [2] A. N. Prodeus, V. S. Didkovskiy, and M. V. Didkovskaya, Akusticheskaja ekspertiza i korekciya kommunikacionnyh kanalov. [Acoustic examination and correction of communication channels. Monograph], Saarbrucken: LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co. KG, 2017. ISBN: 978-3-330-04591-0
- [3] O. Arkhypova, V. Zhuravlov, and V. Kumeiko, "Artikulyatsiini tablytsi sliv ukrainskoi movy [Articulation tables of words of the Ukrainian language]," vol. 19, no. 2, pp. 13-17, 2009. URL: <http://ela.kpi.ua/handle/123456789/9689>
- [4] A. E. Arhipov, and E. A. Arhipova, "Analiz i obrabotka dannyh artikulyacionnyh ispytaniy [Analysis and processing of articulatory test data]," *Scientific and Practical Journal "Information Protection"*, no. 4, pp. 34-42, 2012. URL: http://aprodeus.narod.ru/Teaching/ZAI/Art_2012_Arhipova_Analis_artikul_tabl.pdf
- [5] V. G. Mihajlov, "Diagnosticheskie artikulyacionnye tablicy [Diagnostic articulation tables]," *Acoustic Journal*, vol. 48, no. 5, pp. 705-712, 2002. URL: http://www.akzh.ru/pdf/2002_5_705-712.pdf
- [6] "ANG-2200. Acoustic Noise Generator. Owner's Guide," Research Electronics, Intl. 455 Security Place Algood, TN 38506, USA. [Online]. Available: https://reiusa.net/wp-content/uploads/2017/11/ANG_Manual_revG.pdf
- [7] A. M. Prodeus, A. V. Vityk, and D. Yu. Didenko, "Sub'iektyvne otsiniuvannia yakosti ta rozbirlyvosti movlennievkykh syhnaliv, spotvorenykh syntezovanyym shumamy [Subjective evaluation of quality and intelligibility of speech distorted by synthesized noise]," *Microsystems, Electronics and Acoustics*, vol. 22, no. 6, pp. 45-63, 2017. DOI: [10.20535/2523-4455.2017.22.6.101929](https://doi.org/10.20535/2523-4455.2017.22.6.101929)
- [8] A. Prodeus, K. Bukhta, P. Morozko, O. Serhienko, I. Kotvytskyi, and I. Shherbenko, "14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)," in *Automated System for Subjective Evaluation of the Ukrainian Speech Intelligibility*, Lviv-Slavske, Ukraine, February 20-24, 2018 (pre-print).
- [9] J. Bradley, H. Sato, M. Picard, "On the importance of early reflections for speech in rooms," *Journal of Acoustical Society of America*, vol. 113, no. 6, pp. 3233-3244, June 2003. DOI: [10.1121/1.1570439](https://doi.org/10.1121/1.1570439)
- [10] J. S. Bradley, "Speech intelligibility studies in classrooms," *Journal of the Acoustical Society of America*, vol. 80, no. 3, pp. 846-854, 1986. DOI: [10.1121/1.393908](https://doi.org/10.1121/1.393908)
- [11] W. Yang and J. S. Bradley, "Effects of room acoustics on the intelligibility of speech in classrooms for young children," *Journal of the Acoustical Society of America*, vol. 125, no. 2, pp. 922-933, 2009. DOI: [10.1121/1.3058900](https://doi.org/10.1121/1.3058900)
- [12] I. Arweller, J. Buchholz, and T. Dau, "Proceedings of 2nd International Symposium on Auditory and Audiological Research (ISAAR 2009)," in *Speech intelligibility enhancement by early reflections*, Elsinore, Denmark, August 2009. URL: [http://orbit.dtu.dk/en/publications/speech-intelligibility-enhancement-by-early-reflections\(368caba2-dca2-461f-bc45-d8615f1dd675\).html](http://orbit.dtu.dk/en/publications/speech-intelligibility-enhancement-by-early-reflections(368caba2-dca2-461f-bc45-d8615f1dd675).html)
- [13] Yu. K. Makarov, and A. A. Horev, "K ocenke effektivnosti zashity akusticheskoy (rechevoj) informacii [To assess the effectiveness of protection of acoustic (speech) information]," *Information Security*, no. 4, 2005. URL: [http://www.vrsystems.ru/stati/k_ocenke_effektivnosti_zashiti_akusticheskoi_\(rechevoj\)_informacii.htm](http://www.vrsystems.ru/stati/k_ocenke_effektivnosti_zashiti_akusticheskoi_(rechevoj)_informacii.htm)

Надійшла до редакції 09 квітня 2018 р.

УДК 004.934

Субъективное оценивание разборчивости речи на фоне шума и реверберации

Продеус А. Н., д.т.н. проф., ORCID [0000-0001-7640-0850](https://orcid.org/0000-0001-7640-0850)
e-mail om.prodeus@ae.kpi.ua



Витык А. В., ORCID [0000-0002-6652-4152](https://orcid.org/0000-0002-6652-4152)
 e-mail andrvt8@gmail.com
 Дворник А. А., ORCID [0000-0003-4735-2225](https://orcid.org/0000-0003-4735-2225)
 e-mail alexanderdvornyk@gmail.com
 Котвицкий И. В., ORCID [0000-0003-2800-0080](https://orcid.org/0000-0003-2800-0080)
 e-mail igorktvzk@gmail.com
 Чайка А. С., ORCID [0000-0002-2501-595X](https://orcid.org/0000-0002-2501-595X)
 e-mail sanyarouge@gmail.com
 Ярошенко М. А., ORCID [0000-0001-6327-148X](https://orcid.org/0000-0001-6327-148X)
 e-mail yaroshenkom13@gmail.com

Национальный технический университет Украины
 "Киевский политехнический институт имени Игоря Сикорского", kpi.ua
 Киев, Украина

Реферат—В данной работе представлены результаты субъективного оценивания, проведенного путем артикуляционных испытаний, разборчивости односложных звуко сочетаний на фоне шума и реверберации. Оценивание осуществлялось с помощью специально разработанного программного обеспечения, что позволило автоматизировать и, таким образом, существенно облегчить и ускорить процедуру артикуляционных испытаний. По результатам испытаний маскировочная способность белого шума оказалась лучше таковой для коричневого шума при отношении сигнал-шум, меньших минус 5 дБ, что не вполне согласуется с предыдущими прогнозными оценками. Кроме того, оказалось, что прослушивание речи, искаженной шумом, через акустические мониторы может приводить к существенному повышению (до 0,85-0,93) оценок разборчивости речи, по сравнению со случаем прослушивания через наушники (0,1-0,3). Аналогичные результаты получены для реверберационной помехи: для времени реверберации 2,7 с разборчивость увеличилась с 0,65 до 0,94. Данный феномен может быть в значительной степени объяснен действием ранних отражений звука в помещениях, наличием двух источников излучения и бинауральным прослушиванием. Дополнительными причинами могут быть особенности психофизического состояния слушателей и разработанной автоматизированной системы артикуляционных испытаний.

Библ. 13, рис. 5.

Ключевые слова - разборчивость речи; субъективное оценивание; артикуляционные испытания; шум; реверберация; односложные звуко сочетания.

UDC 004.934

Subjective Evaluation of the Speech Intelligibility on the Background of Noise and Reverberation

A. M. Prodeus, Dr.Sc.(Eng.) Prof., ORCID [0000-0001-7640-0850](https://orcid.org/0000-0001-7640-0850)

e-mail om.prodeus@aae.kpi.ua

A. V. Vityk, ORCID [0000-0002-6652-4152](https://orcid.org/0000-0002-6652-4152)

e-mail andrvt8@gmail.com

O. O. Dvornyk, ORCID [0000-0003-4735-2225](https://orcid.org/0000-0003-4735-2225)

e-mail alexanderdvornyk@gmail.com

I. V. Kotvytskii, ORCID [0000-0003-2800-0080](https://orcid.org/0000-0003-2800-0080)

e-mail igorktvzk@gmail.com

O. S. Chaika, ORCID [0000-0002-2501-595X](https://orcid.org/0000-0002-2501-595X)

e-mail sanyarouge@gmail.com

M. O. Yaroshenko, ORCID [0000-0001-6327-148X](https://orcid.org/0000-0001-6327-148X)

e-mail yaroshenkom13@gmail.com

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", kpi.ua
 Kyiv, Ukraine



Abstract—In this paper, the results of speech intelligibility subjective assessment of Ukrainian speech monosyllabic sound combinations against background noise and reverberation through articulation tests are presented. The evaluation was carried out with the help of specially developed software that allowed automating and thus greatly facilitating and accelerating the procedure of articulation tests. Special text and sound tables of the Ukrainian speech monosyllables of the consonant-vowel-consonant (CVC) type were developed for the tests. The recordings of pure signals were made in the muffled room of the Acoustics and Acoustoelectronics Department of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". The recording was performed at a frequency of 22050 Hz and a bit depth of 16 bits. Speech monosyllables were read using a verbal environment. Listening was done for four situations: pure language; speech distorted by noise; speech distorted by reverberation; speech distorted by the combined effect of noise and reverberation. In the first case, speech monosyllables of 3 articulation tables were listened, each of which contained 50 monosyllables. In the second case, speech distorted by the additive noise with the signal-to-noise ratios minus 10 dB, 0 dB and plus 10 dB was listened. In this case, models of white, pink and brown noises were used, the masking property of which is rather well-studied. In the third case, the reverberant speech for reverberation times from 0.3 to 2.7 s was modeled by convolution of pure speech signals with impulse characteristics of various rooms, and in the fourth case the joint action of pink noise and reverberation was considered. It turned out that the masking ability of white noise exceeds one for brown noise for signal-to-noise ratios (SNR) less than minus 5 dB, which is not entirely consistent with previous preliminary predictive estimates. Signals listening had been made by two ways, namely through the headphones and by means of acoustic monitors. The SNR varied in the range -10...+10 dB, and the reverberation time varied in the range 0.3...2.7 s. In addition, it turned out that listening to speech distorted by noise through acoustic monitors could lead to a significant increase in the intelligibility of speech, compared to the case of listening through headphones. With a signal-to-noise ratio of minus 10 dB, the values of speech intelligibility scores increased from 0.1-0.3 to 0.85-0.93. Similar results were obtained for the reverberation action: the speech intelligibility increased from 0.65 to 0.94 for the reverberation time of 2.7 s. This speech intelligibility growth can be explained, firstly, by the action of early reflections. Secondly, usage of two loudspeakers as sound sources and also binaural listening can be also considered as reasons of speech intelligibility increase. Nevertheless, these reasons for increasing the intelligibility of speech are hardly the only ones. Therefore, it is necessary to further study the causes of the revealed significant increase in speech intelligibility. Among such reasons can be both features of the developed automated system of articulation tests, and features of the psychophysical state of listeners.

Ref. 13, fig. 5.

Keywords - speech intelligibility; subjective assessment; articulatory test; noise; reverberation; monosyllabic words.

