

# Оцінка розбірливості мови за коефіцієнтом міжвушної кореляції мовного сигналу

Деркач<sup>f</sup> Н. М.

e-mail [natashadirkach@gmail.com](mailto:natashadirkach@gmail.com)

Вдовенко<sup>f</sup> М. В., ORCID [0000-0003-1666-3389](https://orcid.org/0000-0003-1666-3389)

e-mail [mv.vdovenko@aae.kpi.ua](mailto:mv.vdovenko@aae.kpi.ua)

Луцьова С. А., доц. к.ф.-м.н., ORCID [0000-0003-0683-1211](https://orcid.org/0000-0003-0683-1211)

e-mail [svetlana\\_lunyova@yahoo.com](mailto:svetlana_lunyova@yahoo.com)

Факультет Електроніки [fel.kpi.ua](http://fel.kpi.ua)

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського» [www.kpi.ua](http://www.kpi.ua)

Київ, Україна

**Реферат** — У даній статті аналізується можливість оцінки розбірливості мови за коефіцієнтом міжвушної кореляції. Обґрунтовується ідея використання мовленнєвої програми для об'єктивної оцінки розбірливості мови в приміщенні. Дається порівняння суб'єктивної та об'єктивізованої оцінки розбірливості мови з оцінкою за допомогою коефіцієнта міжвушної кореляції, одержаного на основі вимірювань бінауральних сигналів в місцях розташування слухачів.

**Ключові слова** — міжвушна кореляція; розбірливість мови; крос-кореляційна функція; стандартні смуги частот; локалізація джерела звуку.

## I. ВСТУП

На сьогоднішній день існує значна кількість об'єктивних методів оцінки розбірливості мови у приміщеннях та акустичних трактах. В усіх цих методах у ролі досліджуваного сигналу використовують широкосмуговий шум. Як показує практика, не дивлячись на достатній набір слухових частот, широкосмуговий шум має значні відмінності від мови у сприйнятті їх людиною [1].

На відміну від шуму мовний сигнал має складнішу структуру. Фізичною основою мовного звучання є акустичні коливання повітряного струменя, які формуються органами артикуляції. Однак не тільки акустика звучання визначає те, як людина чує мову. Фонетична система мови справляє визначальний вплив на механізми і результат сприйняття мови на слух. Звідси - помилки при сприйнятті усної мови в різних ситуаціях спілкування, якщо загальний контекст цієї мови не дозволяє однозначно відновити її словесний склад [2].

Основними параметрами якості мови вважають складову і фразову розбірливість мови, якість мови в порівнянні з якістю мови еталонного тракту та якість мови в реальних умовах роботи. Дослідити та оцінити такі параметри за допомогою широкосмугового шуму є достатньо складно, оскільки при його використанні ми не маємо ні фраз, ні складів. Отже, для більш достовірної об'єктивної оцінки розбірливості мови доцільно використовувати мовленнєвий сигнал.

Метою статті є дослідження можливості використання коефіцієнтів міжвушної кореляції, одержаних шляхом вимірювань бінауральних мовленнєвих сигналів, для оцінки розбірливості мови в приміщенні.

## II. АНАЛІТИЧНИЙ ОГЛЯД ЛІТЕРАТУРИ ТА ПОСТАНОВКА ПРОБЛЕМИ

Розбірливість мови може бути погіршена поганою акустикою приміщення. Це може статися в результаті спотворення самого мовного сигналу, оскільки багато версій мови (відбиті звукові хвилі з часовими затримками) підсумовуються вухом, що викликає як спектральне забарвлення, так і часове розмиття. Відлуння може посилювати маскування сигналу, перешкоджаючи процесам, за допомогою яких слухова система долає таке маскування [3,4]. Відносна важливість цих ефектів залежить від типу слухової ситуації [5].

При розрахунку акустичної якості приміщення, як правило, базуються на таких статистичних даних, як час реверберації T60 та індекс чіткості C50.

Якщо стандартний час реверберації можливо попередньо розрахувати (а у збудованому приміщенні і виміряти), то індекс чіткості C50 тільки вимірюють на слухачьких місцях. Але, не зважаючи на добре розроблену методику таких вимірювань і зручність знаходження індексів чіткості, цей показник не завжди відповідає розбірливості мови.

Авторами пропонується додатково вимірювати ще один об'єктивний параметр, який характеризує



сприйняття звуку слухачем, а саме функцію міжвушної кореляції (IAFC) [6]. За максимальним значенням IAFC, так званим коефіцієнтом міжвушної кореляції (IACC), спробуємо оцінити розбірливість мови.

### III. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

Експеримент проводився в залі з розмірами 14,3x18,5x6,25 м, об'ємом  $V \approx 1653 \text{ м}^3$  (конференц-зал факультету електроніки Київського політехнічного інституту ім. Ігоря Сікорського), який відповідає залу середніх розмірів.

Дослідження проводилися для стереофонічного звучання (ширина бази стереопари – 4,5м. Геометрія розташування точок вимірів представлена на рис.1: відстані від лінії бази до першого ряду точок вимірювання - 3м, до другого ряду – 6,5м і до третього - 9м. Відстані між точками вимірювання по горизонталі: 2,25; 4,5; 7м.

**Експеримент 1.** В оцінюванні розбірливості мови суб'єктивним методом приймали участь 3 диктори (два чоловічих та один жіночий голоси) та аудиторі (у кількості 30 осіб). Експеримент проводився згідно ГОСТ Р 50840-95.

**Експеримент 2.** При оцінюванні розбірливості мови за коефіцієнтом міжвушної кореляції використовувалися: гучномовці Radiotekhnika S-30B, підсилювач потужності Park Audio V4-900 MkII, звукова карта M-Audio Fast Track Pro, манекен голови людини з вмонтованими вимірювальними мікрофонами Behringer ECM8000. Досліджувані точки обрані відповідно до місць розташування аудиторів. Вимірювання бінауральних сигналів проводилися згідно ISO 3382-1:2009.

### IV. ОБГОВОРЕННЯ РЕЗУЛЬТАТІВ

**Експеримент 1.** Виходячи з вимог ГОСТ 50840-95 для проведення артикуляційних випробувань, диктори почергово зачитували складові артикуляційні таблиці. Аудиторі, що знаходилися на слухацьких місцях, записували у відповідний бланк почуті склади та виставляли суб'єктивну оцінку розбірливості мови у даному приміщенні за 5-ти бальною шкалою (оцінка «5» відповідала відмінній розбірливості, «4» – добрій, «3» – задовільній. «2» – слабкій, 1 – недостатній, 0 – абсолютно нерозбірливий). Експеримент проводився декілька разів з різними групами аудиторів, результати усереднені і представлені у таблиці 1.

Суб'єктивна оцінка обрахована у відсотках до максимальної оцінки – 5 балів (100%). Об'єктивізована оцінка обраховується у відсотках артикуляції і дорівнює відсотку вірно зрозумілих складів від загальної кількості прочитаних складів.

Значення об'єктивізованої оцінки коливається у межах 81,6-64%, що знаходиться у межах доброї (76-85% - з напругою уваги, без перепитувань і повторів) і задовільної (61-75% з напругою уваги, рідкими перепитуваннями і повторами) розбірливості. Причому, на більшості слухацьких місць отримана задовільна оцінка розбірливості (середня оцінка розбірливості – 74%).

Що стосується суб'єктивної слухацької оцінки, то вона виставлена у межах 80-40%, що відповідає 4-2 балам. Більшість слухацьких місць мають оцінки «задовільно» (3 бали або 60%). Таке сприйняття мови характеризується підвищеною напругою слухача та необхідністю постійних перепитувань та уточнень.

Отже, суб'єктивні та об'єктивізовані характеристики практично збігаються, і зал у цілому можна характеризувати як зал із задовільною розбірливістю мови. Для конференц-залу, тобто мовного залу, така розбірливість мови є недостатньою.

Однак, слід врахувати, що складові артикуляційні таблиці відрізняються від фразового мовлення більш ускладненою розбірливістю. При зв'язному тексті розбірливість мови поліпшують інтонація, дикція, пристосованість слухачів до термінології і певного стилю мовного матеріалу та інше.

**Експеримент 2.** Пара гучномовців випромінює мовленнєве повідомлення (уривок з роману Михайла Булгакова «Майстер і Маргарита» тривалістю 20с).

Звукові бінауральні сигнали записані в усіх точках приміщення у відповідності з рис.1, окрім точок 3, 7 та 11. На основі вимірюваних функцій обчислені значення коефіцієнту міжвушної кореляції (IACC) [6], представлені у таблиці 2.

Найвищі значення коефіцієнту міжвушної кореляції співпадають з центральним розміщенням слухацьких крісел у залі відносно положення гучномовців (т.1, т.5). Це узгоджується зі значенням міжвушної часової затримки, яка є найменшою у цих точках і складає близько 0,3мс.

Зазначимо, що значення коефіцієнту міжвушної кореляції для мовленнєвого сигналу майже у два рази менші ніж для імпульсного сигналу [6]. Такий спад пояснюється складною структурою мови на відміну від широкопasmового шуму.

Таким чином, на основі порівнянь значень IACC (табл.2) з результатами артикуляційних випробувань (табл.1) можна встановити відповідність значень IACC певній степені розбірливості мови: значення  $IACC > 0,5$  – розбірливість добра;  $0,3 < IACC < 0,5$  - задовільна;  $0,2 < IACC < 0,3$  – слабка;  $IACC < 0,2$  – незадовільна.

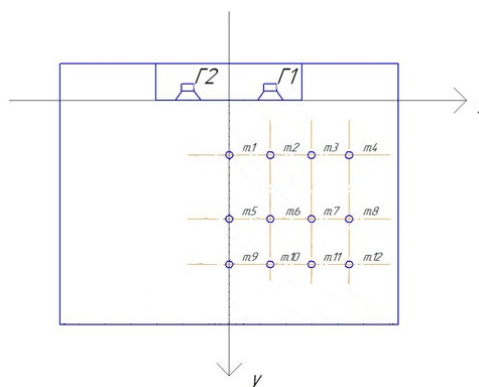


Рис. 1 Геометрія розташування стереосистеми та точок вимірювання в залі риклад назви рисунку

ТАБЛИЦЯ 1 ОЦІНКА РОЗБІРЛИВОСТІ МОВИ СУБ'ЕКТИВНИМ (СУБ.) ТА ОБ'ЕКТИВНИМ (ОБ.) СПОСОБАМИ, У ВІДСОТКАХ

Ряд	Суб.	Об.	Суб.	Об.	Суб.	Об.	Суб.	Об.
Перший	т.1		т.2		т.4		Середня по ряду	
	80	81,6	60	79,2	60	76,4	67	79
Середній	т.5		т.6		т.8		Середня по ряду	
	80	81,6	60	69,8	40	61	60	71,8
Останній	т.9		т.10		т.12		Середня по ряду	
	60	73,6	60	70,6	40	70,4	53	71,6
Загальна оцінка по залу							60	74

ТАБЛИЦЯ 2 ЗНАЧЕННЯ КОЕФІЦІЄНТІВ МІЖВУШНОЇ КОРЕЛЯЦІЇ (ІАСС) У ТОЧКАХ ВИМІРЮВАННЯ ДЛЯ МОВНОГО СИГНАЛУ

№ точки вимірювань	1	2	4	5	6	8	9	10	12
ІАСС	0,44	0,34	0,42	0,53	0,35	0,2	0,34	0,27	0,28

### ВИСНОВКИ

Для оцінки розбірливості мови у залі можливо використовувати значення коефіцієнтів міжвушної кореляції мовного сигналу, одержаних шляхом вимірювання бінауральних сигналів на слухацьких місцях.

На основі порівняння результатів артикуляційних випробувань у залі зі значеннями коефіцієнтів міжвушної кореляції можна встановити шкалу значень ІАСС, яка відповідає певній градації розбірливості мови.

### ПЕРЕЛІК ПОСИЛАНЬ

- [1] S. Luniova, V. Didkovs'kyu, O. Pedchenko "Akustyka movotvorennia [Acoustics of speech formation]," Lambert Academic Publishing, 2018, ISBN: 978-613-7-32891-0.
- [2] G. Ye. Kedrova, V. V. Potapov, Ye. B. Omel'yanova, A. M. Yegorov, and M. V. Volkova, "Russkaya fonetika: Elektronnyy uchebnyk [Russian Phonetics: An Electronic Textbook]," Otrasleyoy Fond algoritmov i programm M, 2002.
- [3] T. May, S. van de Par, and A. Kohlrausch, "Binaural Localization and Detection of Speakers in Complex Acoustic Scenes," in *The Technology of Binaural Listening*, J. Blauert, Ed. Springer-Verlag Berlin Heidelberg, 2013, pp. 397–425. DOI: [10.1007/978-3-642-37762-4\\_15](https://doi.org/10.1007/978-3-642-37762-4_15).
- [4] J. Blauert and W. Cobben, "Some Consideration of Binaural Cross Correlation Analysis," *Acta Acust. united with Acust.*, vol. 39, no. 2, pp. 96–104, 1978.
- [5] M. Lavandier and J. F. Culling, "Speech segregation in rooms: Monaural, binaural, and interacting effects of reverberation on target and interferer," *J. Acoust. Soc. Am.*, vol. 123, no. 4, pp. 2237–2248, 2008, DOI: [10.1121/1.2871943](https://doi.org/10.1121/1.2871943)
- [6] M. V. Vdovenko and S. A. Luniova "Vyznachennya oblasti stereofonichnoho zvuchannya dzherel informatsiynykh syhnaliv [Determination of the Are of Stereo Sound of Sources of Information Signals]," *Microsystemi, elektronika ta akustika*, 2018, pp. 58-65, DOI: [10.20535/2523-4455.2018.23.6.148691](https://doi.org/10.20535/2523-4455.2018.23.6.148691)
- [7] V.V. Furduev "Akusticheskiye osnovy veshchaniya [Acoustic broadcast basics]," Moscow, USSR: Gossudarstvennoye izdatel'stvo literatury po voprosam svyazi i radio, 1980, p.318.
- [8] J. Blauert, "Prostranstvennyy slukh [Spatial hearing]," Moscow, USSR: Energiya, 1979.
- [9] Y. A. Kovalgin, "Stereofonia [Stereophony]," Moscow, USSR: Radio i svyaz, 1989.
- [10] I. Titze, Principles of voice production. Englewood Cliffs, N.J.: Prentice Hall, 1994, ISBN 978-0-13-717893-3.



УДК 534.134

# Оценка разборчивости речи по коэффициенту междушной корреляции речевого сигнала

Деркач Н. М.

e-mail [natashadirkach@gmail.com](mailto:natashadirkach@gmail.com)

Вдовенко М. В., ORCID [0000-0003-1666-3389](https://orcid.org/0000-0003-1666-3389)

e-mail [mv.vdovenko@aae.kpi.ua](mailto:mv.vdovenko@aae.kpi.ua)

Лунёва С. А., доц. к.ф.-м.н., ORCID [0000-0003-0683-1211](https://orcid.org/0000-0003-0683-1211)

e-mail [svetlana\\_lunyova@yahoo.com](mailto:svetlana_lunyova@yahoo.com)

Факультет электроники [fel.kpi.ua](http://fel.kpi.ua)

Национальный технический университет Украины

«Киевский политехнический институт имени Игоря Сикорского» [www.kpi.ua](http://www.kpi.ua)

Киев, Украина

**Реферат** — У данной статье анализируется возможность оценки разборчивости речи по коэффициенту междушной корреляции. Обосновывается идея использования речевой программы для объективной оценки разборчивости речи у помещения. Дается сравнение субъективной и объективизированной оценки разборчивости речи с оценкой с помощью коэффициента междушной корреляции, полученного на основе измерений бинауральных сигналов у местах расположения слушателей.

**Ключевые слова** — междушная корреляция; разборчивость речи; кросс-корреляционная функция; стандартные полосы частот; локализация источника звука.



# The assessment of speech intelligibility by the coefficient of inter ear correlation speech signal

N. M. Derkach<sup>f</sup>

e-mail [natashadirkach@gmail.com](mailto:natashadirkach@gmail.com)

M. V. Vdovenko<sup>f</sup>, ORCID [0000-0003-1666-3389](https://orcid.org/0000-0003-1666-3389)

e-mail [mv.vdovenko@aae.kpi.ua](mailto:mv.vdovenko@aae.kpi.ua)

S. A. Luniova, Assoc.Prof. PhD, ORCID [0000-0003-0683-1211](https://orcid.org/0000-0003-0683-1211)

e-mail [svetlana\\_lunyova@yahoo.com](mailto:svetlana_lunyova@yahoo.com)

Faculty of Electronics [fel.kpi.ua](http://fel.kpi.ua)

National technical university of Ukraine

«Igor Sikorsky Kyiv polytechnic institute» [www.kpi.ua](http://www.kpi.ua)

Kyiv, Ukraine

**Abstract** — This article analyzes the possibility of evaluating the intelligibility of speech in terms of the coefficient of interaural correlation. The idea of using a speech program for objectively assessing the intelligibility of speech in the premises is substantiated. A comparison is made of the subjective and objectivized evaluation of the intelligibility of the speech with the estimation by means of the coefficient of the correlation obtained on the basis of measurements binaural signals in the locations of listeners. Today, there is a significant amount of objective methods for assessing the readability of speech in rooms and acoustic paths. In all these methods, the role of the test signal uses broadband noise. As practice shows, despite the sufficient set of auditory frequencies, broadband noise has significant differences from the language human perception.

Unlike noise, the speech signal has a more complex structure. The physical basis of speech is the acoustic oscillations of air jet, which are formed by the organs of articulation. However, not only the acoustics of sounding determines how a person hears the tongue. The phonetic language system has a decisive influence on the mechanisms and the result of speech perception. Hence - errors in the perception of oral speech in different situations of communication, if the general context of this language does not allow unequivocally restore its verbal composition.

The main parameters of language quality are considered to be component and phrasal legibility of the language, the quality of the language in comparison with the quality of the language of the reference path and the quality of speech in real working conditions. Exploring and estimating such parameters with broadband noise is quite difficult, since when using it, we have no phrases or syllables. Consequently, for a more reliable objective assessment of legibility of speech it is expedient to use a speech signal. The purpose of the article is to investigate the possibility of using interaural correlation obtained by measuring binaural speech signals, in order to assess the readability of the language in the room.

In order to assess the readability of speech in the hall, it is possible to use the values of the interaural correlation obtained by measuring binaural signals in listening places.

By comparing the results of articulation tests in the room with the values of the coefficients between the ear correlations, one can establish a scale of IASS values that accord to a certain degree of intelligible speech.

**Keywords** — *interaural correlation; speech intelligibility; cross-correlation function; standard frequency bands; sound source localization.*

