

Міністерство освіти і науки України
Національний технічний університет України «Київський політехнічний
інститут імені Ігоря Сікорського»

Кваліфікаційна наукова праця на правах рукопису

Моторнюк Дар'я Євгенівна

УДК 621.391.83

ДИСЕРТАЦІЯ
ОЦІНКА АКУСТИЧНИХ ПАРАМЕТРІВ ПРИМІЩЕННЯ ЯК КАНАЛУ
МОВЛЕННЄВОЇ КОМУНІКАЦІЇ

17 Електроніка та телекомунікації

171 Електроніка

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело

_____ / Моторнюк Д. Є.

Науковий керівник Продеус Аркадій Миколайович, доктор технічних наук,
професор

Київ – 2021

АНОТАЦІЯ

Моторнюк Д. Є. Оцінка акустичних параметрів приміщення як каналу мовленнєвої комунікації. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 171 «Електроніка». – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», МОН України, Київ, 2021.

Дисертаційна робота присвячена дослідженню впливу акустичних характеристик приміщення, ранніх і пізніх відбиттів звуку на розбірливість мовленнєвого сигналу, автоматизації оцінювання акустичних характеристик приміщення, автоматизації проведення артикуляційних випробувань

Зміст дисертаційного дослідження викладено у чотирьох розділах, у яких представлені та обґрунтовані основні результати роботи.

У вступі обґрунтовано актуальність дисертаційної роботи, сформульовано мету та перераховано задачі дослідження, описано методи дослідження, надана інформація про наукову новизну та практичне значення одержаних результатів.

Перший розділ присвячено огляду публікацій з моделювання імпульсної характеристики приміщення, впливу реверберації на мовленнєвий сигнал, методів компенсації впливу реверберації та методів вимірювання часу реверберації. В існуючих роботах з впливу приміщення на мовленнєвий сигнал бракує аналізу зв'язку між якістю та розбірливістю мовлення. Окрім цього, існуюча методологія проведення суб'єктивних тестів дуже ресурсомістка і потребує автоматизації.

У другому розділі розглянуто вимірювання імпульсної характеристики і часу реверберації прямими методами (з використанням тестового сигналу). Запропоновано покращення алгоритму автоматичного оцінювання часу реверберації прямим методом, що полягає у автоматичному, віконному з перекриттям, аналізі записаного відгуку на збуджуючий сигнал. Запропонований алгоритм був апробований при дослідженні акустичних

параметрів учбових приміщень кафедри акустичних та мультимедійних електронних систем НТУУ «Київський політехнічний інститут імені Ігоря Сікорського». Крім того, розроблено рекомендації щодо коригування характеристик акустичної апаратури непрофесійного рівня якості при проведенні вимірювань якості та розбірливості мовлення із використанням штучної голови. Ефективність зазначених рекомендацій була перевірена експериментально, шляхом оцінювання акустичних характеристик університетських аудиторій малого, середнього та великого розміру, а також шляхом перевірки узгодження отриманих результатів із літературними даними.

У третьому розділі представлено результати модернізації відомого алгоритму сліпого вимірювання часу реверберації. Складність завдання полягає в тому, що при таких вимірюваннях в розпорядженні інженера є лише запис мовленнєвого сигналу, спотвореного реверберацією. В дисертаційній роботі запропоновано таке покращення зазначеного алгоритму, яке дозволило розширити діапазон вимірювання часу реверберації розглянутим методом. Достовірність та точність отриманих результатів перевірена із використанням комп'ютерного моделювання, шляхом порівняння результатів алгоритму-прототипу та модернізованого алгоритму.

В четвертому розділі дисертаційної роботи представлено результати розробки апаратно-програмної системи для автоматизації артикуляційних випробувань при суб'єктивному оцінюванні розбірливості мовлення. Використання розробленої системи дозволило зіставити результати оцінювання якості мовленнєвих сигналів із результатами оцінювання розбірливості мовлення та зробити висновки щодо принципової можливості використання простих, з точки зору об'єму обчислень, оцінок якості сигналів замість складних оцінок розбірливості мовлення. Представлено результати експериментального порівняння розбірливості мовлення при діотичному та дихотичному прослуховуванні мовленнєвих сигналів. Виконано порівняння впливу ранньої та пізньої реверберації на результати суб'єктивного оцінювання розбірливості

мовлення. Представлено результати комп'ютерного моделювання впливу ранньої реверберації на розбірливість мовлення в залежності від структури імпульсної характеристики приміщення.

Результати досліджень дозволили створити уточнені аналітичні та комп'ютерні моделі дії акустичних завад у вигляді шуму, ранніх та пізніх відбиттів звуку в приміщенні на якість мовленнєвих сигналів та на розбірливість мовлення. Виявлення негативної дії ранніх відбиттів звуку на розбірливість мовлення в середній частині приміщень дозволяє обґрунтувати необхідність прийняття відповідних дій для нейтралізації виявленого ефекту.

Результати, представлені у дисертації, можуть бути використані для експертизи та корекції акустичних характеристик приміщень шляхом розрахунку та вимірювання якості та розбірливості мовленнєвих сигналів у приміщеннях, що проектуються, будуються або піддаються реновації. Володіння інформацією щодо характеру та ступеня впливу шуму, ранніх відбиттів звуку та пізньої реверберації на якість та розбірливість мовлення в різних точках приміщень дозволить підвищити точність акустичної експертизи. Автоматизація вимірювань акустичних характеристик приміщення, в тому числі якості та розбірливості мовлення, дозволить підвищити точність вимірювань, знизивши при цьому витрати матеріальних, часових, фінансових та людських ресурсів.

Викладені у дисертації нові теоретичні та практичні результати досліджень знайшли застосування у освітньому процесі кафедри акустичних та мультимедійних електронних систем за спеціальністю 171 Електроніка, освітньою програмою “Акустичні електронні системи та технології обробки акустичної інформації” Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”.

В дисертаційній роботі отримано наступні наукові результати:

- вперше вказано умови, за яких об'єктивне оцінювання якості мовленнєвих сигналів, що маскуються шумовою завадою, може бути застосовано замість суб'єктивного оцінювання розбірливості мовлення;
- вперше вказано на можливість зниження розбірливості мови в окремих частинах приміщень через погіршення якості мовленнєвого сигналу, яке спричинене дією ранніх відбиттів звуку;
- вперше запропоновано при автоматизації оцінювання розбірливості українського мовлення артикуляційним методом, спосіб фіксації сприйнятих слухачем звукових сполучень, що використовуються в якості елементів мовлення;
- вдосконалено алгоритми оцінювання часу реверберації прямим та сліпим методами, що дозволяє розширити границі застосування зазначених алгоритмів та підвищити ступінь автоматизації акустичних вимірювань;
- розроблено рекомендації щодо коригування характеристик акустичної апаратури непрофесійного рівня якості при проведенні вимірювань якості та розбірливості мовлення із використанням штучної голови.

Практичне значення отриманих в дисертаційній роботі результатів полягає у наступному:

- показано можливість підвищення достовірності результатів артикуляційних випробувань та підвищення ступеня автоматизації таких випробувань, спрямованих на суб'єктивну оцінку розбірливості українського мовлення в приміщеннях;
- встановлено принципову можливість використання мір якості мовленнєвого сигналу замість мір розбірливості мовлення, що дозволяє спростити процедуру оцінювання розбірливості мови;
- розроблено програмно-апаратну систему для артикуляційних випробувань каналів передачі української мови в умовах дії шумової завади та

реверберації, що дозволяє автоматизувати такі випробування, підвищити їх достовірність та суттєво знизити ресурсомісткість випробувань;

– вдосконалено алгоритми оцінювання часу реверберації прямим та сліпим методами, що дозволяє розширити границі застосування зазначених алгоритмів та підвищити ступінь автоматизації акустичних вимірювань.

Ключові слова: імпульсна характеристика приміщення, реверберація, мовленнєвий сигнал, спотворення мовленнєвого сигналу, суб'єктивна розбірливість, автоматизація процедур оцінювання.

Список публікацій здобувача

1. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, and D. Motorniuk, "Automated Subjective Assessment of Speech Intelligibility under Diotic and Dichotic Listening," *Lecture Notes on Data Engineering and Communications Technologies*, 2021, 48, pp. 1–16, January 2021, Springer, Cham, ISSN 2367-4512, e-ISSN 2367-4520, DOI: 10.1007/978-3-030-43070-2_1
Scopus: <https://www.scopus.com/sourceid/21100975545>

2. A. Prodeus, M. Didkovska, K. Kukharicheva, D. Motorniuk, "Two Simplified Models of Early Sound Reflections in a Room", *Electronics and Control Systems*, No. 3(65), 2020, pp. 79-87. DOI:10.18372/1990-5548.65.14991, ISSN 1990-5548 (Published: 2020-11-27)

3. O. Dvornyk, A. Prodeus, D. Motorniuk, M. Didkovska, "Hardware and Software System "Artificial Head". Part 2. Evaluation of Speech Intelligibility in Classrooms," *Microsystems, Electronics and Acoustics*, vol. 22, no. 3, 2020, pp. 48-55, DOI: 10.20535/2523-4455.me.209928 (Опубліковано: Dec 28, 2020), ISSN 2523-4447, e-ISSN 2523-4455

4. O. Dvornyk, A. Prodeus, M. Didkovska, D. Motorniuk, "Artificial Software Complex "Artificial Head". Part 1. Adjusting the Frequency Response of the Path," *Microsystems, Electronics and Acoustics*, vol. 22, no. 1, 2020, pp. 56-64, DOI: 10.20535/2523-4455.me.198431, ISSN 2523-4447, e-ISSN 2523-4455

5. Ю.С. Костючок, Л.С. Мартинович, Д.Е. Моторнюк, В.А. Нечитайло, А.В. Храпачевский, А.Н. Продеус. Акустическая паспортизация учебных помещений. - ISSN 1811-4512. Electronics and Communications, 2016, Vol.21, No.2(91). – P.63-70

6. A. Prodeus, M. Didkovska, D. Motorniuk, O. Dvornyk, "The Effects of Noise, Early and Late Reflections on Speech Intelligibility" Proc. 2020 IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO), pp. 488-492, DOI: 10.1109/ELNANO50318.2020.9088854

7. A. Prodeus, M. Didkovska, K. Kukharicheva, D. Motorniuk, "Modeling the Influence of Early Sound Reflections on Speech Intelligibility," Proc. of 2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 20-23 Oct. 2020, Kyiv, Ukraine. DOI: 10.1109/MSNMC50359.2020.9255657

8. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, D. Motorniuk, and A. Khrapachevskyi, "Objective and Subjective Assessment of the Quality and Intelligibility of Noised Speech," Proc. of 2018 IEEE International Scientific-Practical Conference "Problems of Infocommunications. Science and Technology" (PIC S&T`2018), October 9-12, 2018, Kharkiv, Ukraine, pp. 71-74. DOI: 10.1109/INFOCOMMST.2018.8632125.

9. А. М. Продеус, І. В. Котвицький, М. В. Дідковська, В. С. Дідковський, К. А. Кухарічева, Д. Є. Моторнюк, О. О. Дворник Спосіб виявлення кліпування мовного та музичного сигналів // Патент UA 144291 U, МПК G01R 23/20, опубл. 25.09.2020.

SUMMARY

Motorniuk D.E. Assessment of an acoustic parameters of the room as a speech communication channel. – Qualifying scientific work on the rights of the manuscript.

Thesis for the degree of Philosophy Doctor, in specialty 171 “Electronics”. – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 2021.

The dissertation is devoted to the study of the influence of acoustic characteristics of the room, early and late sound reflections on the intelligibility of the speech signal, automation of evaluation of acoustic characteristics of the room, automation of articulation tests on the subjective intelligibility of speech.

The content of the dissertation research is presented in four sections, in which the main results of the work are presented and substantiated.

The introduction substantiates the relevance of the dissertation, formulates the purpose and lists the objectives of the study, describes the research methods, provides information about the scientific novelty and practical significance of the results.

The first section is devoted to a review of publications on modeling the room impulse response, the impact of reverberation on the speech signal, methods of compensating for the impact of reverberation and methods of measuring reverberation time. The considered works on the influence of the room on the speech signal lack the analysis of the connection of objective indicators of speech quality with subjective intelligibility. In addition, the proposed methodology for conducting subjective tests is very resource-intensive and requires automation.

The second section considers the measurement of impulse response and reverberation time by direct methods (using a test signal). It is proposed to improve the algorithm for automatic estimation of reverberation time by the direct method, which consists in automatic windowed with overlap analysis of the recorded response to the test signal. The proposed algorithm was used in the study of acoustic parameters of the classrooms of the Department of Acoustic and Multimedia Electronic Systems of NTUU "Igor Sikorsky Kyiv Polytechnic Institute". In addition, recommendations

have been developed for adjusting the characteristics of non-professional quality acoustic equipment when measuring the quality and intelligibility of speech using an artificial head. The effectiveness of these recommendations was tested experimentally, by evaluating the acoustic characteristics of university classrooms of small, medium and large size, as well as by checking the consistency of the results with the literature.

The third section presents the results of improvement of the known algorithm for blind reverberation time estimation. The complexity of the problem is that in such measurements, the engineer has only a record of the speech signal distorted by reverberation. In the dissertation such improvement of the specified algorithm which has allowed to expand the range of measurement of reverberation time by the considered method is offered. The reliability and accuracy of the obtained results were verified using computer modeling, by comparing the results of the prototype algorithm and the improved algorithm.

The fourth section of the work presents the results of hardware and software system development for automation of articulation tests in the subjective assessment of speech intelligibility. The use of the developed system allowed to compare the results of speech signal quality assessment with the results of speech intelligibility assessment and to draw conclusions about the fundamental possibility of using simple, in terms of computational volume, signal quality assessments instead of complex speech intelligibility assessments. The results of experimental comparison of speech intelligibility during diotic and dichotic listening of speech signals are presented. A comparison of the influence of early and late reverberation on the results of speech intelligibility subjective assessment is performed. The results of computer modeling of the influence of early reverberation on speech intelligibility depending on the structure of the room impulse response are presented.

The research results allowed to create refined analytical and computer models of the influence of acoustic interference in the form of noise, early and late reflections of sound in the room on the quality of speech signals and speech intelligibility. Discovery of the negative effect of early sound reflections on speech intelligibility in

the middle of the rooms allows to justify the need to take appropriate action to neutralize the discovered effect.

The results presented in the dissertation can be used for examination and correction of room acoustic characteristics by calculating and measuring the quality and intelligibility of speech signals in the rooms that are designed, built or renovated. Having information on the nature and extent of noise, early sound reflections and late reverberation on the speech quality and intelligibility in different parts of rooms will increase the accuracy of acoustic examination. Automation of measurements of room acoustic characteristic, including speech quality and intelligibility, will increase the accuracy of measurements, while reducing the cost of material, time, financial and human resources.

The new theoretical and practical research results presented in the dissertation were used in the educational process of the Department of Acoustic and Multimedia Electronic Systems in the specialty 171 Electronics, educational program "Acoustic Electronic Systems and Acoustic Information Processing Technologies", National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

The following scientific results were obtained in the dissertation:

- for the first time, the conditions under which the subjective assessment of the quality of speech signals masked by noise interference can be used instead of the speech intelligibility subjective assessment are indicated;
- for the first time the possibility of speech intelligibility reduction in some parts of the rooms due to the deterioration of the speech signal quality, which is caused by the action of early sound reflections is indicated;
- for the first time in the automation of Ukrainian speech intelligibility assessment by articulation method, a method of recording the perceived by the listener sound combinations, used as elements of speech, proposed;
- the algorithms for reverberation time estimation by direct and blind methods have been improved, which allows to expand the application limits of these algorithms and increase the degree of acoustic measurements automation;

– developed recommendations for adjusting the characteristics of acoustic equipment of non-professional quality level for measuring the speech quality and intelligibility using an artificial head.

The practical significance of the results obtained in the dissertation is as follows:

– the possibility of increasing the reliability of articulation tests results and increasing the degree of automation of such tests, aimed at the subjective assessment of Ukrainian speech intelligibility in the rooms is demonstrated;

– the principal possibility of speech signal quality measures usage instead of speech intelligibility measures is established, which allows to simplify the procedure of speech intelligibility assessment;

– a software and hardware system for articulation tests of Ukrainian speech transmission channels under conditions of noise interference and reverberation has been developed, which allows to automate such tests, increase their reliability and significantly reduce the resource consumption;

– the algorithms for reverberation time estimation by direct and blind methods have been improved, which allows to expand the application limits of these algorithms and increase the degree of acoustic measurements automation.

Keywords: room impulse response, reverberation, speech signal, distortion of speech signal, subjective intelligibility, estimation procedures automation.

List of applicant's publications

1. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, and D. Motorniuk, “Automated Subjective Assessment of Speech Intelligibility under Diotic and Dichotic Listening,” *Lecture Notes on Data Engineering and Communications Technologies*, 2021, 48, pp. 1–16, January 2021, Springer, Cham, ISSN 2367-4512, e-ISSN 2367-4520, DOI: 10.1007 / 978-3-030-43070-2_1

Scopus: <https://www.scopus.com/sourceid/21100975545>

2. A. Prodeus, M. Didkovska, K. Kukharicheva, D. Motorniuk, "Two Simplified Models of Early Sound Reflections in a Room", *Electronics and Control Systems*,

No. 3 (65), 2020, pp. 79-87. DOI: 10.18372/1990-5548.65.14991, ISSN 1990-5548 (Published: 2020-11-27)

3. O. Dvornyk, A. Prodeus, D. Motorniuk, M. Didkovska, "Hardware and Software System "Artificial Head" Part 2. Evaluation of Speech Intelligibility in Classrooms," *Microsystems, Electronics and Acoustics*, vol. 22, no. 3, 2020, pp. 48-55, DOI: 10.20535/2523-4455.me.209928 (Published: Dec 28, 2020), ISSN 2523-4447, e-ISSN 2523-4455

4. O. Dvornyk, A. Prodeus, M. Didkovska, D. Motorniuk, "Artificial Software Complex "Artificial Head" Part 1. Adjusting the Frequency Response of the Path," *Microsystems, Electronics and Acoustics*, vol. 22, no. 1, 2020, pp. 56-64, DOI: 10.20535/2523-4455.me.198431, ISSN 2523-4447, e-ISSN 2523-4455

5. Yu. S. Kostyuchok, L. S. Martinovich, D. E. Motorniuk, V. O. Nechytailo, A. V. Khrapachevskyi, A. M. Prodeus. Acoustic certification of classrooms. – ISSN 1811-4512. *Electronics and Communications*, 2016, Vol.21, No. 2 (91). – P.63-70

6. A. Prodeus, M. Didkovska, D. Motorniuk, O. Dvornyk, "The Effects of Noise, Early and Late Reflections on Speech Intelligibility" *Proc. 2020 IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO)*, pp. 488 - 492, DOI: 10.1109/ELNANO50318.2020.9088854

7. A. Prodeus, M. Didkovska, K. Kukharicheva, D. Motorniuk, "Modeling the Influence of Early Sound Reflections on Speech Intelligibility," *Proc. of 2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC)*, 20-23 Oct. 2020, Kyiv, Ukraine. DOI: 10.1109/MSNMC50359.2020.9255657

8. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, D. Motorniuk, and A. Khrapachevskyi, "Objective and Subjective Assessment of the Quality and Intelligibility of Noised Speech," *Proc. of 2018 IEEE International Scientific-Practical Conference "Problems of Infocommunications. Science and Technology" (PIC S&T`2018)*, October 9-12, 2018, Kharkiv, Ukraine, pp. 71-74.

DOI: 10.1109/INFOCOMMST.2018.8632125.

9. A. M. Prodeus, I. V. Kotvitskyi, M. V. Didkovska, V. S. Didkovskyi, K. A. Kukharicheva, D. E. Motorniuk, O. O. Dvornyk Method of detecting clipping of speech and music signals // Patent RU 144291 U, IPC G01R 23/20, publ. 09/25/2020

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	16
ВСТУП.....	17
1. АКУСТИЧНІ ХАРАКТЕРИСТИКИ ПРИМІЩЕННЯ, МОДЕЛЮВАННЯ ІМПУЛЬСНОЇ ХАРАКТЕРИСТИКИ ПРИМІЩЕННЯ	25
1.1. Приміщення як канал мовленнєвої комунікації.....	25
1.2. Вплив шуму, пізніх і ранніх відбиттів на розбірливість мови.....	27
1.3. Імпульсна характеристика приміщення та її моделі	29
1.4. Реверберація як різновид спотворення мовленнєвого сигналу	33
1.5. Використання спектра пізньої реверберації для компенсації впливу пізньої реверберації в частотній області.....	38
1.6. Прямі методи вимірювання часу реверберації.....	40
1.7. Сліпі методи вимірювання часу реверберації	43
Висновки до розділу 1	48
2. ПРЯМІ МЕТОДИ ВИМІРЮВАННЯ АКУСТИЧНИХ ХАРАКТЕРИСТИК	49
2.1. Вимірювання акустичних характеристик приміщення за умови наявності фонової завади.....	49
2.2. Вимірювання бінауральних імпульсних характеристик приміщення	53
Висновки до розділу 2	66
3. МОДЕЛЮВАННЯ АЛГОРИТМУ СЛІПОГО ВИМІРЮВАННЯ ЧАСУ РЕВЕРБЕРАЦІЇ І ПРИДУШЕННЯ РЕВЕРБЕРАЦІЇ.....	67
3.1. Комп'ютерне моделювання алгоритму вимірювання часу реверберації методом максимальної правдоподібності.....	67
3.2. Оптимізація алгоритму вимірювання часу реверберації методом максимальної правдоподібності	70
3.3. Оптимізація параметрів алгоритму придушення реверберації	72
Висновки до розділу 3	74

4. АВТОМАТИЗАЦІЯ ОБ'ЄКТИВНОГО І СУБ'ЄКТИВНОГО ОЦІНЮВАННЯ ЯКОСТІ І РОЗБІРЛИВОСТІ МОВЛЕННЄВОГО СИГНАЛУ	75
4.1. Автоматизація артикуляційних випробувань для суб'єктивного оцінювання розбірливості мовлення.....	75
4.2. Порівняння об'єктивних показників якості мови з суб'єктивною оцінкою розбірливості.....	80
4.3. Аналіз впливу бінаурального ефекту на розбірливість мовлення у приміщенні	87
4.4. Порівняння впливу ранньої та пізньої реверберації на суб'єктивну оцінку розбірливості мовлення.	96
4.5. Моделювання впливу ранніх відбиттів на розбірливість мови	105
4.6 Акустична паспортизація навчальних приміщень.....	114
4.7. Автоматизоване оцінювання розбірливості мовлення у приміщенні в залежності від розташування слухача.....	120
4.8 Спосіб виявлення кліпування мовленнєвого сигналу	128
Висновки до розділу 4	131
ВИСНОВКИ	133
Список літератури.....	136
ДОДАТОК А. Список опублікованих праць за темою дисертації.....	150
ДОДАТОК Б. Графіки відповідності ЧР виміряного методом МП дійсному ЧР.....	152
ДОДАТОК В. Графіки залежності абсолютної похибки вимірювання від часу реверберації	154
ДОДАТОК Г. Приклад запропонованого акустичного паспорту навчальної аудиторії	156
ДОДАТОК Д. Акт про використання результатів дисертаційної роботи	157

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

- BSD, Bark Spectral Distortion – барк-спектральне спотворення
- CEP – відстань кепструма
- DMOS, Degradation Mean Opinion Scale – шкала деградації середньої суб'єктивної оцінки
- EDT, Early Decay Time – час приходу ранніх відбиттів
- ERB, Early Reflection Benefit – користь ранніх відбиттів
- IS – відстань Ітакура-Сайто
- LLR – логарифмічний коефіцієнт правдоподібності
- LSD, Log-Spectral Distance – логарифмічна спектральна відстань
- MLS, Maximum Length Sequence – послідовність максимальної довжини, M-послідовність
- PESQ, Perceptual Evaluation of Speech Quality – перцептивна оцінка якості мовлення
- RIR, Room Impulse Response – імпульсна характеристика кімнати
- SNR, Signal-to-Noise Ratio – відношення сигнал/завада
- STI, Speech Transmission Index – коефіцієнт передачі мовлення
- WSS – зважена спектральна відстань
- AT – акустичний тракт
- АЧХ – амплітудно-частотна характеристика
- ВКФ – взаємно-кореляційна функція
- ГМ – підсистема «гучномовець-мікрофон»
- ІХ – імпульсна характеристика
- ЛЧМ – лінійно-частотна модуляція
- МП – максимальна правдоподібність
- ЧР – час реверберації
- ШГ – штучна голова

ВСТУП

Актуальність роботи. Дисертаційна робота направлена на об'єктивізацію та автоматизацію оцінювання якості приміщень, призначених для мовленнєвих презентацій. Суттєвим недоліком суб'єктивного оцінювання якості приміщень є високі вимоги до людських, часових та фінансових ресурсів. Такого недоліку в значній мірі позбавлені об'єктивні (інструментальні) методи й системи оцінювання, чим і пояснюється незмінний інтерес науковців та інженерів до таких методів і систем. Проте впровадження об'єктивних методів та систем оцінювання якості приміщень утруднене існуванням низки проблем.

Суттєвою проблемою при застосуванні методів об'єктивного оцінювання є вибір міри якості приміщення як каналу передачі мовленнєвої інформації. Складність вибору полягає в існуванні двох груп мір: мір якості мовленнєвих сигналів та мір розбірливості мовлення, спотвореного в приміщенні, яке є каналом передачі інформації. Іншою проблемою є різний характер завад, таких як шум, реверберація, відлуння. Ситуація ускладнюється також існуванням різних способів випромінення мовленнєвих сигналів (диктор, що в процесі мовлення змінює своє місцеположення, або один чи низка гучномовців, розміщених в різних місцях приміщення). Нарешті важливим є спосіб сприйняття мовленнєвого сигналу (сприймання мовленнєвого сигналу слуховою системою людини, що знаходиться в приміщенні, де відбувається мовленнєва презентація, або запис та обробка сигналу системою автоматичного розпізнавання мовлення).

Оскільки якість та розбірливість мовленнєвого сигналу є поняттями тісно пов'язаними, хоча й не тотожними, виникає питання про принципову можливість використання мір якості замість мір розбірливості, й навпаки.

Нагальною проблемою є автоматизація суб'єктивного оцінювання розбірливості мовленнєвих сигналів, оскільки від наявності та достовірності еталонних результатів суб'єктивного оцінювання залежить принципова можливість практичного використання об'єктивних мір розбірливості

мовлення. Ліцензованих систем такого типу в Україні немає, що робить актуальною розробку відносно простих, недорогих та надійних систем автоматизації оцінювання розбірливості українського мовлення.

Іншою проблемою є надзвичайно велика кількість об'єктивних мір якості сигналів, що значно утруднює вибір конкретної міри. Хоча кількість об'єктивних мір розбірливості мовлення є суттєво меншою, проте труднощі вибору конкретного методу оцінювання розбірливості мовлення пов'язані із необхідністю попереднього визначення типу завад, а також із наявними вимогами до процедури вимірювань, такими як скритність, швидкість, точність вимірювань.

Значний внесок у дослідження та розв'язання завдань прогнозування та вимірювання акустичних параметрів приміщень зробили такі іноземні та радянські вчені, як Коллард Д., Флетчер Х., Штейнберг Д., Бредлі Д., Шредер М., Рашевський Я.І., Каргашин В.Л., Макаров Ю.К., Хорев А.А., Покровський М.Б., Сапожков М.А., Биков Ю.С., М'ясніков Л.Л. Серед вітчизняних вчених дослідженням каналів передачі мовленнєвої інформації займалися: Касьянов Ю.І., Нужний С.М., Архіпов О.Є., Архіпова О.О., Журавльов В.М., Кумейко В.М., Продеус А.М., Найда С.А., Дідковський В.С.

Проте аналіз праць, що стосуються оцінки акустичних параметрів приміщення як каналу мовленнєвої комунікації, виявив низку прогалин в розв'язанні завдань зазначеного напрямку. Зокрема, це такі важливі з теоретичної та практичної точки зору питання як зв'язок між якістю та розбірливістю мовленнєвих сигналів, вплив відбиваючих поверхонь приміщення на розбірливість мовлення в різних точках приміщення, достовірність та ступінь автоматизації оцінок акустичних характеристик приміщення.

Таким чином, розробка нових та вдосконалення існуючих методів оцінки якості приміщення як каналу мовленнєвої комунікації є **актуальною** науково-

технічною задачею акустичної експертизи комунікаційних каналів, що має важливе прикладне значення.

Мета і задачі дослідження. *Метою дисертаційної роботи є розробка нових та вдосконалення відомих методів оцінки акустичних параметрів приміщення, що забезпечують підвищення достовірності та зменшення ресурсомісткості акустичної експертизи трактів передачі мовленнєвої інформації.*

Об'єктом дослідження є процес оцінювання акустичних характеристик та параметрів приміщень, призначених для мовленнєвих презентацій.

Предметом дослідження є вплив акустичних параметрів приміщення на якість та розбірливість мовленнєвого сигналу.

Для досягнення поставленої мети необхідно було вирішити такі завдання:

1. Виконати аналітичний огляд сучасних методів оцінювання акустичних параметрів приміщення, в тому числі, прямих та сліпих методів вимірювання часу реверберації.

2. Дослідити можливість вдосконалення прямого та сліпого методів автоматичного вимірювання часу реверберації з метою розширення границь їх застосування при збереженні точності вимірювань.

3. Розробити апаратно-програмну систему автоматизованого оцінювання розбірливості української мови, спотвореної шумом та реверберацією, а також іншими чинниками, притаманними каналам передачі акустичної інформації.

4. Виробити рекомендації щодо коригування характеристик акустичної апаратури непрофесійного рівня якості при проведенні вимірювань імпульсної характеристики приміщення та її параметрів для подальшого прийняття рішення стосовно розбірливості мовлення, спотвореного реверберацією, в різних точках приміщення.

5. Встановити умови припустимості застосування простих в обчислювальному сенсі об'єктивних мір якості мовленнєвих сигналів замість складних об'єктивних мір розбірливості мовлення.

6. Дослідити залежність оцінок розбірливості мовлення від способів сприйняття мовленнєвих сигналів, таких як діотичне та діхотичне прослуховування, при спотворенні мовленнєвих сигналів шумом та реверберацією.

Методи дослідження. Для досягнення поставленої мети використано методи аналітичного та комп'ютерного моделювання імпульсних характеристик приміщення та мовленнєвих сигналів, спотворених шумом та реверберацією. При проведенні експериментальних досліджень, корекцію частотної характеристики вимірювального тракту виконано методами оберненої фільтрації та регуляризації. При вимірюваннях імпульсних характеристик приміщень застосовано метод взаємно-кореляційної обробки вхідного та вихідного сигналів лінійної системи. Для автоматизації обчислень оцінок акустичних характеристик та параметрів, автоматизації проведення артикуляційних випробувань та для автоматизації вимірювань імпульсних характеристик із застосуванням апаратно-програмної системи «Штучна голова» використано середовище MATLAB.

Наукова новизна отриманих результатів.

У дисертації представлено наступні наукові результати:

1. Вперше вказано на існування об'єктивних оцінок якості мовленнєвих сигналів, які можуть бути використані замість оцінок розбірливості мовлення на тлі шумової завади, що дозволяє суттєво спростити процедуру вимірювання розбірливості мови.

2. Вперше вказано на доцільність розширення переліку параметрів, що входять до акустичного паспорту приміщення, включенням до цього переліку

оцінок розбірливості мови в різних точках приміщення, що дозволяє обґрунтувати рішення стосовно необхідності коригування акустики приміщення.

3. Вперше експериментально доведено принципову можливість автоматизованої фіксації та обробки звукових сполучень типу "приголосний-голосний-приголосний", сприйнятих слухачем при оцінюванні розбірливості українського мовлення артикуляційним методом, що дозволяє суттєво зменшити ресурсомісткість процедури артикуляційних випробувань.

4. Вдосконалено способи оцінювання часу реверберації прямим та сліпим методами, що дозволяє розширити границі застосування зазначених алгоритмів та підвищити ступінь автоматизації процедури вимірювань часу реверберації.

5. Розроблено рекомендації щодо коригування характеристик акустичної апаратури при проведенні вимірювань розбірливості мовлення із використанням штучної голови, що дозволяє виконувати вимірювання із використанням апаратури непрофесійного рівня якості.

Особистий внесок здобувача

Усі результати, наведені у дисертаційній роботі і винесені на захист, отримані за активної участі автора та опубліковано у спеціалізованих фахових виданнях.

У роботі [1], опублікованій в співавторстві, здобувачем особисто виконано наступне: виконано аналіз сучасного стану та тенденцій розвитку методів виявлення кліпування сигналів, що дозволило виявити прогалини в даному напрямі досліджень. У роботі [2], опублікованій в співавторстві, здобувачем особисто виконано наступне: виконано обробку результатів експериментальних досліджень, що дозволило отримати залежності результатів вимірювань від умов вимірювань. У роботі [3], опублікованій в співавторстві, здобувачем особисто виконано наступне: розроблено програмне забезпечення для автоматичної оцінки акустичних характеристик приміщень, організовано

проведення вимірювань акустичних параметрів низки університетських аудиторій та розроблено форму акустичного паспорту. У роботі [4], опублікованій в співавторстві, здобувачем особисто виконано наступне: виконано розрахунки при моделюванні ранніх відбиттів. У роботі [5], опублікованій в співавторстві, здобувачем особисто виконано наступне: виконано обчислення при моделюванні ранніх відбиттів. У роботах [6], [7], [8], [9], опублікованих у співавторстві, здобувачем особисто виконано обробку та аналіз результатів експериментальних досліджень.

Практичне значення одержаних результатів.

Результати досліджень дозволили створити уточнені аналітичні та комп'ютерні моделі впливу акустичних завад у вигляді шуму, ранніх та пізніх відбиттів звуку в приміщенні на якість мовленнєвих сигналів та на розбірливість мовлення. Виявлення негативної дії ранніх відбиттів звуку на розбірливість мовлення в середній частині приміщень дозволяє обґрунтувати необхідність прийняття відповідних дій для нейтралізації виявленого ефекту.

Практичне значення отриманих результатів полягає у підвищенні достовірності результатів артикуляційних випробувань та у підвищенні ступеня автоматизації таких випробувань, призначених для суб'єктивного оцінювання розбірливості українського мовлення в приміщеннях. Встановлення принципової можливості використання мір якості мовленнєвого сигналу замість мір розбірливості мовлення дозволяє спростити процедуру оцінювання розбірливості мови. Розробка програмно-апаратної системи для артикуляційних випробувань каналів передачі української мови в умовах дії шумової завади та реверберації дозволяє автоматизувати такі випробування, підвищити їх достовірність та суттєво знизити ресурсомісткість випробувань. Вдосконалення алгоритмів оцінювання часу реверберації прямим та сліпим методами дозволяє розширити границі застосування зазначених алгоритмів та підвищити ступінь автоматизації акустичних вимірювань.

Результати, представлені у дисертації, можуть бути використані для експертизи та корекції акустичних характеристик приміщень шляхом розрахунку та вимірювання якості та розбірливості мовленнєвих сигналів у приміщеннях, що проектуються, будуються або піддаються реновації. Володіння інформацією щодо характеру та ступеня впливу шуму, ранніх відбиттів звуку та пізньої реверберації на якість та розбірливість мовлення в різних точках приміщень дозволить підвищити точність акустичної експертизи. Автоматизація вимірювань акустичних характеристик приміщення, в тому числі якості та розбірливості мовлення, дозволить підвищити точність вимірювань, знизивши при цьому витрати матеріальних, часових, фінансових та людських ресурсів.

Зв'язок роботи з науковими програмами, планами, темами.

Викладені у дисертації нові теоретичні та практичні результати досліджень знайшли застосування у освітньому процесі кафедри акустичних та мультимедійних електронних систем за спеціальністю 171 Електроніка, освітньою програмою “Акустичні електронні системи та технології обробки акустичної інформації”, Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”.

Апробація результатів дисертації. Матеріали дисертаційних досліджень обговорювалися на міжнародних конференціях:

1. IEEE International Scientific-Practical Conference "Problems of Infocommunications. Science and Technology" (PIC S&T`2018), Україна, м. Харків, 9-12 жовтня 2018
2. IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO), Україна, м. Київ, 22-24 квітня 2020
3. IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), Україна, м. Київ, 20-23 жовтня 2020

Публікації. За результатами досліджень опубліковано 9 наукових праць, 5 статей у наукових фахових виданнях, (з них 1 стаття у періодичних наукових виданнях інших держав, які входять до ОЕСР та/або Європейського Союзу, фахових виданнях України категорії «А» або закордонних виданнях, що входять до WoS або Scopus), 1 патент України на корисну модель, 3 тези доповідей в збірниках матеріалів конференцій.

Структура та обсяг дисертації.

Робота складається зі вступу, чотирьох розділів, списку використаних джерел із 134 найменувань та 5 додатків. Робота містить 63 рисунки та 4 таблиці. Загальний обсяг дисертаційної роботи складає 157 сторінок.

1. АКУСТИЧНІ ХАРАКТЕРИСТИКИ ПРИМІЩЕННЯ, МОДЕЛЮВАННЯ ІМПУЛЬСНОЇ ХАРАКТЕРИСТИКИ ПРИМІЩЕННЯ

1.1. Приміщення як канал мовленнєвої комунікації

Результати численних досліджень свідчать, що акустичні характеристики приміщень істотно впливають на якість сприйняття інформації слуховою системою людини [10], [11], [12], [13], [14], [15]. Це означає, що архітектори і дизайнери повинні керуватися вимогами акустичного паспорта при проектуванні нових навчальних закладів, оскільки засвоєння звукової інформації при навчанні особливо важливо. З іншого боку, це означає, що адміністрація навчальних закладів повинна приділяти значну увагу акустичній експертизі та паспортизації експлуатованих навчальних приміщень.

У період з 1999 по 2009 рр. канадськими акустиками Дж. Бредлі, Х. Сато і В. Янгом [10], [11], [12] отримані надзвичайно важливі результати у вигляді кількісних оцінок ступеня впливу шуму і реверберації на розбірливість мови. При цьому експериментально показано, що шумова перешкода значно небезпечніше реверберації. Причин тому кілька. По-перше, рівень шуму може бути вельми високим не тільки і не стільки через близькість до слухачів технічних пристроїв, скільки через близькість основних джерел шуму, якими є поруч сидять учні. Ситуація ускладнюється тим, що мовленнєвий шум володіє високими маскуючими властивостями через подібності спектрів шуму і сигналу. Разом з тим, час реверберації в навчальних приміщеннях рідко перевищує 0,9 с. Це означає, що дифузна частина звукових відображень, іменована «пізньою реверберацією», має порівняно низьку інтенсивність і тому має слабкий маскуючий ефект.

Іншим важливим результатом, представленим в роботах [10], [11], [12], є кількісна оцінка залежності розбірливості мови від віку слухачів. Наявність такої оцінки дозволило зробити висновок, що для досягнення словесної розбірливості 95% в класних приміщеннях для школярів 6 років відношення сигнал-шум має бути не нижче 16 дБ, тоді як для школярів 11 років такий поріг

може бути знижений до 9 дБ, а для студентів даний поріг знижується до 1 дБ.

Результати дослідження великих лекційних аудиторій, оснащених системами звукопідсилення і обладнаних індукційними петлями на допомогу людям з ураженням слуху, представлені в роботі швейцарського акустика К. Еггеншвілера [13]. При цьому зазначено, що рівень шуму в навчальних приміщеннях повинен бути в межах 30-40 дБА, а рівень мовленнєвого сигналу повинен складати 65-75 дБ. Час реверберації для приміщень об'ємом до 1000 м³ не повинна перевищувати 0,8 с. При оцінюванні розбірливості мови нормальними слід вважати значення індексу STI > 0,56 і значення коефіцієнта втрати приголосних ALcons > 8%. Відзначено також, що при використанні систем звукопідсилення можливе виникнення ряду проблем. Зокрема, таких як невірна установка і орієнтація гучномовців. Крім того, частотні характеристики гучномовців часто оптимізують під музику, але не під мова. Мікрофонні входи можуть не забезпечувати відсутність перевантаження, а користуватися ними надто складно. Через погану акустику приміщень доводиться використовувати системи придушення зворотного акустичного зв'язку. Системи допомоги людям із вадами слуху (індукційні петлі, системи з частотною модуляцією або інфрачервоними датчиками) або відсутні, або працюють погано.

У роботі К. Марино [14] також оцінювалися характеристики великого лекційного залу. Оцінювання проводилося як шляхом комп'ютерного моделювання (програма ODEON), так і шляхом вимірювань акустичних параметрів, таких як час реверберації T₆₀, EDT, індекс розбірливості мови STI і коефіцієнт ясності C₈₀. Оцінка якості приміщення показала, що час реверберації нижче, ніж належить для приміщень таких розмірів. Отримані результати пояснювались пірамідальною формою стелі і підвищеним поглинанням поверхонь. В результаті були зроблені висновки про необхідність використання системи звукопідсилення для підвищення рівня прямого звуку.

Результати оцінки акустики трьох лекційних приміщень в Лундському університеті (Швеція, університет заснований в 1666 р) наведені в роботі

Р. Юсефа [15]. При цьому оцінювалися як старі, так і недавно побудовані приміщення. Як проблеми відзначено все більш широке використання технічних аудіовізуальних засобів, які встановлювалося в приміщеннях, не розрахованих для цього. Наслідком таких дій може бути погіршення засвоєння інформації студентами. При експериментальних дослідженнях оцінювалися і порівнювалися з рекомендованими значеннями наступні параметри: час реверберації T_{20} , сила звуку G , коефіцієнт ясності C_{50} і індекс розбірливості мови STI . В результаті експериментів виявилось, що всі три приміщення дуже посередні в плані чіткості мови, оскільки $STI = 0,45-0,50$.

Очевидним недоліком робіт [13], [14], [15] є переважаюча увага до оцінювання ревербераційних характеристик приміщень, тоді як в роботах Д. Бредлі [10], [11], [12] ясно вказано на більш важливу роль шумової перешкоди. Частково зазначений недолік компенсується наявністю результатів вимірювань індексу розбірливості мови STI . Крім того, в роботах [10], [11], [12], [13], [14], [15] відсутні чіткі рекомендації щодо вибору оцінюваних параметрів для акустичної паспортизації приміщень.

1.2. Вплив шуму, пізніх і ранніх відбиттів на розбірливість мови

Суб'єктивна оцінка розбірливості мови, спотвореної шумом та реверберацією, важлива для поліпшення моделей слухової системи людини [16], [17], калібрування систем оцінки об'єктивної розбірливості мови [18], розробки та сертифікації пристроїв та каналів зв'язку, аудиторій та конференц-залів, слухових апаратів та кохлеарних імплантатів [19], [10], [20].

Кількісні оцінки ступеня впливу шуму та реверберації на розбірливість мови в аудиторіях представлені в роботах [10], [11], [21], [12]. Експериментально було показано, що шумові завади набагато небезпечніші, ніж реверберація, через близькість джерел шуму (студенти, що розмовляють, сидячі поруч) і через подібність спектра перешкод та мови. Час реверберації в

аудиторіях рідко перевищує 0,9 с, тому пізня реверберація мало впливає на розбірливість мови через відносно низьку інтенсивність перешкод такого типу.

У роботі [22] було виявлено, що ранні відбиття можуть збільшити SNR до 6-9 дБ, оскільки енергія раннього відбиття може бути на 9 дБ більшою, ніж енергія прямого звуку. Для конкретних ситуацій прослуховування, наприклад коли голова оратора відвернута від слухача, ранні відбиття можуть бути відносно важливішими за прямий звук. У той же час, можливість збільшення SNR до 6-9 дБ була поставлена під сумнів, а результати розширених досліджень були представлені в [23]. У роботі [23] було показано, що збільшення енергії ранніх відбиттів призводить до поліпшення розбірливості мови, але це покращення менше, ніж у разі збільшення прямої звукової енергії. Таким чином, лише частина енергії ранніх відбиттів корисна для розбірливості мови. Для дизайнерів кімнат це означає, що посилення ранніх відбиттів є розумним, але підвищення рівня голосу оратора є більш вигідним. Можливо, цей результат можна пояснити тим, що не тільки енергія, а й фаза ранніх відбиттів відіграє важливу роль.

Що стосується дії бінаурального прослуховування, то в [23] було показано, що бінауральне прослуховування може збільшити SNR до 2-3 дБ. Водночас можна припустити, що існує небезпека послаблення позитивного ефекту ранніх відбиттів при сукупній дії бінаурального прослуховування та пізньої реверберації [24].

У [23] також зазначалося, що різниця в результатах [22] та [23] може бути значною мірою обумовлена різницею експериментальних налаштувань. По-перше, для визначення розбірливості мови використовувалися різні стимули. По-друге, моделювання звукових полів не були ідентичними: у [22] було використано лише 7 ранніх відбиттів, тоді як 20 ранніх відбиттів було використано в [23]. По-третє, пізня реверберація розглядалася в [22], тоді як відбиття, що надходять пізніше 55 мс після прямого звуку, були відкинуті в [23]. Існували також різні моделі фонового шуму. Навколишній шум із формою

спектра, що відповідає контуру NC 40, був використаний у [22], а дифузний стаціонарний мовленнєвий шум – у [23].

Ідея використання співвідношення енергії ранніх до пізніх відбиттів, розроблена в [25], [26], дозволила врахувати негативні наслідки як фонового шуму, так і реверберації. Ця концепція, як було показано в роботах [27], [28], може бути використана для оцінки якості мовленнєвих сигналів та для оцінки розбірливості мови в широкому діапазоні приміщень.

Було показано, що шум набагато небезпечніший за реверберацію в аудиторії [11], [10].

Результати суб'єктивної оцінки шумової розбірливості українського мовлення були представлені в [29] на відстані 0,6–0,8 м між колонками комп'ютера та слухачами, коли реверберація практично не впливає на розбірливість мови. У [29] було запропоновано спробувати тестування на інших відстанях, де дія реверберації більш помітна.

На жаль, у згаданих документах бракує інформації про те, наскільки сильно вплив ранніх відбиттів впливає на розбірливість мови.

1.3. Імпульсна характеристика приміщення та її моделі

Імпульсна характеристика (ІХ) приміщення залежить від геометрії та акустичного оформлення кімнати. У [30] акцентують увагу на тому, що для одного приміщення існує безліч ІХ, оскільки ІХ залежить від двох параметрів: положення джерела і положення приймача. В [30] уточняється термін ІХ приміщення як «імпульсна характеристика передавального шляху між джерелом і приймачем».

В тій же роботі [30] приміщення характеризують як лінійну систему. Тому його імпульсна характеристика дає повний опис змін звукового сигналу на шляху між двома точками кімнати. Якщо шлях передачі звуку розглядається як лінійна система, співвідношення між спектральними формами вхідного та вихідного сигналів таке:

$$Y(j\omega) = H(j\omega)X(j\omega),$$

де $X(j\omega)$ – спектральна форма вхідного сигналу, $Y(j\omega)$ – спектральна форма вихідного сигналу, $H(j\omega)$ – передавальна функція лінійної системи.

Використавши зворотне перетворення Фур'є, отримаємо запис для імпульсної характеристики системи:

$$h(\tau) = \mathcal{F}^{-1}[H(j\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega)e^{j\omega\tau} d\omega.$$

Імпульсну характеристику приміщення як функцію таких параметрів як розташування джерела та розташування приймача, також розглянуто в роботі [31], де надано визначення поняття точкової імпульсної характеристики приміщення.

Точкова ІХ описує часово-просторові властивості приміщення наступним чином:

$$h_{\theta\phi}(t) = \left[\sum_{p=0}^P h_{(p\theta\phi)}(t) \right],$$

де p – номер відбиття звуку, θ – азимут, а ϕ – нахил траєкторії відбиття.

В роботі [32] для моделювання реверберації використовують модель джерела зображення (Image Source Model, ISM). ISM моделює реверберацію на основі припущення, що звук у приміщенні поширюється як промінь. Акустичний промінь прямує безпосередньо від джерела до приймача, але також опосередковано, дзеркально відбиваючись від поверхонь. Потім його можна екстраполювати як звук, що доходить до приймача з двох джерел, другим джерелом є джерело зображення за дзеркалом. Положення джерела зображення обчислюється з положення джерела і положення та кута відбивача. Пряма відстань між джерелом зображення та приймачем містить інформацію, необхідну для моделювання фактичного відбитого звукового шляху. Кожна відбиваюча поверхня сама по собі виробляє джерело зображення, а джерело зображення другого порядку створюється шляхом об'єднаного відбиття двох

поверхонь, що дає нам чотири ефективні джерела замість одного. Це можна поширити на бажану кількість джерел зображення, чим більша кількість джерел зображення, тим більший ступінь точності моделювання ІХ.

Існує аналогічна модель ІХ приміщення, яка називається моделлю розповсюдження променів [33]. Основна відмінність між моделлю джерела зображення і променевою моделлю полягає в методиці розрахунку шляхів відбиттів [24].

Модальна модель ІХ була розглянута в роботах [34], [35]. Ця модель побудована на основі власних резонансних частот кімнати, які називаються нормальними модами. Коли джерело звуку вмикається у приміщенні, воно збуджує одну або кілька нормальних мод кімнати. Коли джерело звуку вимкнено, моди продовжують резонувати своєю накопиченою енергією, кожна з яких розпадається зі швидкістю, що визначається константою спаду моди, δ .

В [36] за основу прийнято модель ІХ, показану на рис. 1.1. Можна бачити, що обвідна ІХ є експонентою, за нахилом якої в логарифмічному масштабі можна визначити час реверберації в приміщенні. Маючи ІХ приміщення, можна також оцінити структуру ранніх відбиттів звуку, співвідношення енергій початкової та хвостової частин (так званий «індекс чіткості» звуку). Що стосується часу реверберації, розрізняють кілька видів його оцінки. Наприклад, оцінка EDT (Early Decay Time) характеризує суб'єктивне сприйняття якості звуку в приміщенні слухачами, тоді як інші оцінки (T_{10} , T_{20} , T_{30}) характеризують приміщення як таке.

Для кожної точки приміщення ІХ описує послідовність приходу імпульсів, що відповідають прямому звуку та окремим відбиттям від поверхонь приміщення. Розподіл цих імпульсів в часі визначається геометрією приміщення і координатами точок передачі і приймання звуку.

Структура ранніх звукових відбиттів, що визнається їх рівнями і часом запізнення відбитого від огороження звуку по відношенню до прямого звуку, має значний вплив на якість сприйняття в залі як мовленнєвих, так і музичних

програм. Ця характеристика, будучи локальною, тісно пов'язана з усіма архітектурними параметрами приміщення: його розмірами, формою, обрисами і оздобленням внутрішніх поверхонь.



Рис. 1.1 – Спрощена модель імпульсної характеристики приміщення [36].

На рис. 1.1 добре видно, що перші після випромінення прямого звуку відбиття є дискретними, вони чітко розділені між собою. Згідно ефекту Хааса [37], ранні відбиття сприймаються як одне ціле, якщо вони надходять до слухача не більше, ніж за 50 мс. Тому їх прийнято вважати корисними, оскільки ранні відбиття підсилюють прямий звук, сприймаючись з ним як одне ціле.

Можна побачити, що імпульси, які надходять в точку прийому, з часом стають дифузними, зливаються між собою і формують дифузний ревербераційний хвіст.

Співвідношення між енергіями ранньої та пізньої складових реверберації описують індекс чіткості (ясності) мови C_{50} (для мовленнєвих сигналів) та індекс прозорості звуку C_{80} (для музичних програм):

$$C_{50} = 10 \lg \frac{E_{50}}{L_{50}}, \quad C_{80} = 10 \lg \frac{E_{80}}{L_{80}},$$

де $E_{80} = \int_0^{80} h^2(t) dt$, $L_{80} = \int_{80}^{\infty} h^2(t) dt$, $h(t)$ – імпульсна характеристика приміщення. За результатами оцінювання C_{50} можна обчислити коефіцієнт складової розбірливості [38].

$$k_c = (1 + 10^{-0,1 \cdot C_{50}})^{-0,2}.$$

Мірою якості мовленнєвого сигналу також може служити коефіцієнт низького тону

$$k_{HT} = \frac{T_{125} + T_{250}}{T_{500} + T_{1000}},$$

де в чисельнику і знаменнику вказані значення часу реверберації на частотах 125, 250, 500 і 1000 Гц.

1.4. Реверберація як різновид спотворення мовленнєвого сигналу

Інерційність слухової системи людини призводить до інтеграції ранньої реверберації і прямого звуку, внаслідок чого збільшується сила прямого звуку, який сприймається слухачем. Про цю властивість у 1935 році Ф. Айгнер і М. Страт повідомили у своїй статті [39]. Вони вперше запропонували співвідношення акустичної енергії для вимірювання кількісної оцінки впливу фонового шуму і акустичних параметрів приміщення на розбірливість мови. Це співвідношення можна записати таким чином:

$$Q = \frac{E_{Пр} + E_P}{E_{П} + E_{Ш}}, \quad (1.1)$$

де $E_{Пр}$ – енергія прямого звуку, E_P – енергія ранніх відбиттів (таких, що приходять до слухача не пізніше 60 мс від моменту випромінення), $E_{П}$ – енергія пізніх відбиттів, які приходять після 60 мс, $E_{Ш}$ – енергія шуму.

Айгнер і Страт розвинули цю ідею, сказавши, що $Q = 1$ – нижній поріг для отримання задовільного звукового враження. Вираз (1.1) є основою більшості показників розбірливості мови, які були згодом розроблені. Важливо відзначити, що звукове враження може бути покращене за рахунок зменшення як $E_{П}$, так і $E_{Ш}$.

Приголосні грають більшу роль в розбірливості мови, ніж голосні. В 1971 В. Пейтц [40] запропонував такий параметр як втрата артикуляції приголосних, яка чисельно виражає зниження сприйняття приголосних внаслідок

реверберації і залежить від відстані джерело-приймач. Можна знизити втрату артикуляції, а отже, підвищити розбірливість мови зниженням відстані джерело-приймач або часу реверберації.

В 1980 Д. Берклі [41] дійшов висновку, що реверберація виражається двома компонентами: забарвленням та луною. Він показав, що спектральне відхилення σ добре корелюється з суб'єктивним сприйняттям забарвлення, а луна напряму відповідає часу реверберації T_{60} .

В 1982 Дж. Аллен [42] опублікував формулу, яка дає змогу передбачити суб'єктивне сприйняття реверберованої мови:

$$P = P_{\max} - \sigma T_{60}, \quad (1.2)$$

де P – суб'єктивне сприйняття в деяких одиницях вимірювання, P_{\max} – максимально можливе значення цього сприйняття. Згідно (1.2), понижаючи як спектральне відхилення, так і час реверберації, можна підвищити якість реверберованої мови.

А. Болт і Р. Макдональд [43], а також А. Набелек [44] запропонували два фактори погіршення реверберованої мови: самомаскування та перекриваюче маскування. Самомаскування відповідає зміні в часі та по частоті однієї фонемі, що порушує фонетичну інформацію, яка є необхідною для впізнавання звуку. Перекриваюче маскування проявляється тоді, коли попередня фонема та її відбиття маскують наступну фонему.

Вплив реверберації при монауральному сприйнятті набагато більший, ніж при бінауральному [24]. Якщо людина чує двома вухами, пізня реверберація сприймається як некорельований шум, і мозок її відфільтровує, а мова є більш зрозумілою, ніж при прослуховуванні її одним вухом в тому ж приміщенні.

В [45] дуже добре показаний характер пізньої реверберації як завади.

Модель мовленнєвого сигналу $y(t)$, який спотворено реверберацією, можна представити у вигляді згортки чистого мовленнєвого сигналу $x(t)$ з ІХ $h(t)$ приміщення:

$$y(t) = \int_0^{\infty} h(v)x(t-v)dv = x(t) \otimes h(t), \quad (1.3)$$

де \otimes – символ згортки.

Використовуючи модель Полека [46] для ІХ приміщення

$$h(t) = \xi(t)e^{-\delta t}, \quad (1.4)$$

де $\xi(t)$ – стаціонарний білий шум, $\delta = 3 \ln 10/T_{60}$ – показник швидкості загасання рівня звуку в приміщенні, а також розділяючи в ІХ $h(t)$ (рис. 1.1) області, що відповідають раннім та пізнім відбиттям

$$h_i(t) = \begin{cases} h(t), & 0 \leq t \leq T_l; \\ 0, & \text{дп.т} \end{cases} \quad (1.5)$$

$$h_l(t) = \begin{cases} h(t+T_l), & t \geq 0; \\ 0, & \text{дп.т} \end{cases}, \quad (1.6)$$

спотворений реверберацією сигнал можна представити у вигляді:

$$y(t) = h_i(t) \otimes x(t) + r(t), \quad (1.7)$$

де $r(t) = h_l(t) \otimes x(t - T_l)$ – компонент, спричинений дією пізньої реверберації; T_l – момент початку пізньої реверберації. Вважаючи, що доданки співвідношення (1.7) статистично незалежні, приходимо до трактування пізньої реверберації як різновиду адитивного шуму [45]. Але, на жаль, нестаціонарний характер цього шуму робить неефективним безпосереднє використання традиційних методів придушення шуму, які розраховані на стаціонарний або повільний нестаціонарний [45].

Проілюструємо дію реверберації на мовленнєвий сигнал за допомогою графіків і спектрограм чистого і реверберованого сигналів (рис. 1.2). На рис. 1.2 показана спектрограма і графік чистого тестового мовленнєвого сигналу. На спектрограмі (рис. 1.2) можна чітко побачити мовленнєві форманти, які також називають максимумами спектру. Крім того, з рисунку видно, що фонемі добре розділені у часі.

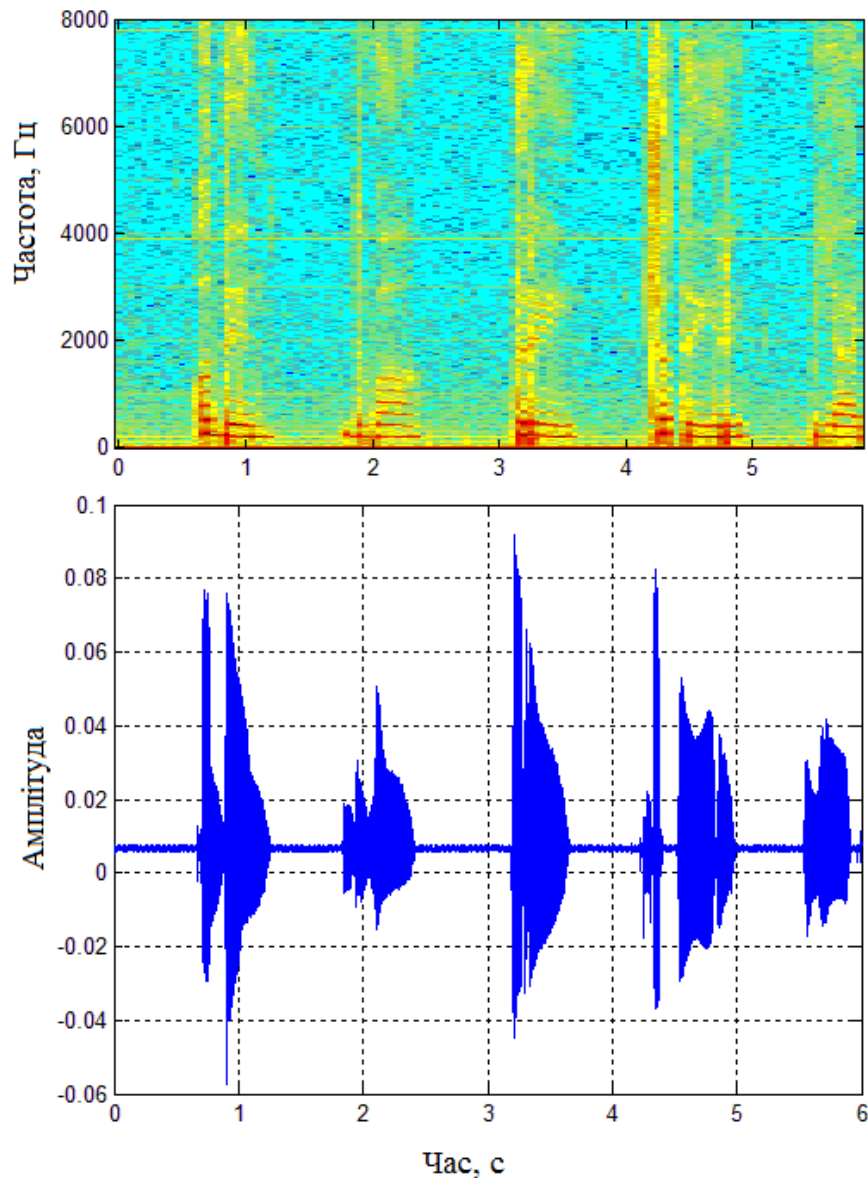


Рис. 1.2 – Спектрограма (зверху) і графік (знизу) чистого мовленнєвого сигналу.

Чистий мовленнєвий сигнал був згорнутий з імпульсною характеристикою (ІХ) кімнати, в результаті чого був отриманий реверберований мовленнєвий сигнал (рис. 1.3). На рисунку 1.3 дуже добре видно спотворення сигналу внаслідок його проходження через акустичний тракт (АТ) (згортка чистого сигналу з ІХ імітує проходження сигналу через АТ [47]). Спектрограма (рис. 1.3) ілюструє розмивання мовленнєвих формант як за часом, так і за частотою. І на спектрограмі, і на графіку (рис. 1.3) дуже добре видно розмивання фонем у часі, внаслідок чого паузи між словами заповнені

реверберацією, а послідовні фонемі перекриваються. Згорнувши чистий сигнал з IX кімнати, ми начебто пропустили мову через НЧ-фільтр (якщо зіставити спектрограми рис. 1.2 і рис. 1.3 можна побачити, що ревербована мова втратила значну частину високочастотної енергії). Отже, реверберація не тільки розмиває фонемі, але і змінює тембр сигналу. Внаслідок цих спотворень не тільки добре чути різницю між чистим і ревербованим сигналом, але й значно погіршується розбірливість мовлення. Крім того, якщо ставиться завдання автоматичного розпізнавання мовлення, помітно знижується точність (процент правильно сприйнятих слів) такого розпізнавання.

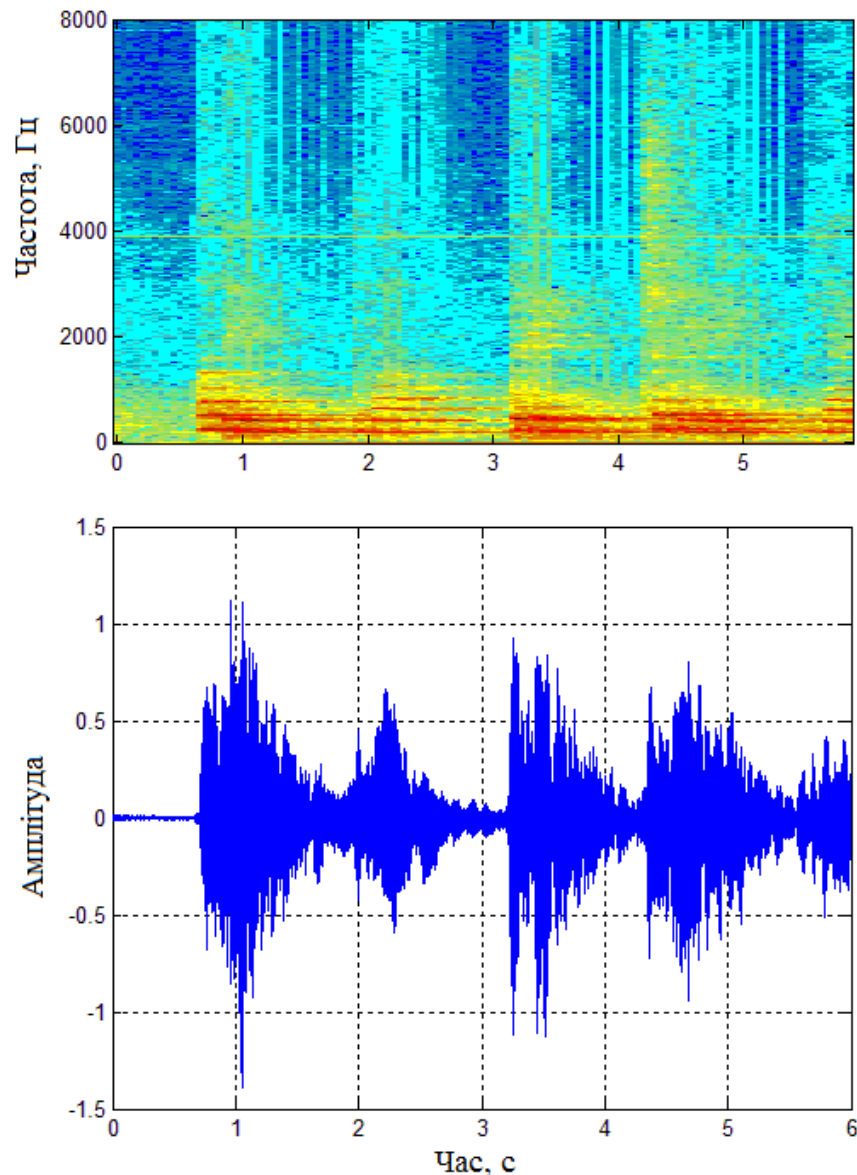


Рис. 1.3 – Спектрограма (зверху) і графік (знизу) ревербованого мовленнєвого сигналу.

Було б зручно вважати, що реверберація виключно зменшує розбірливість, але це припущення загалом є не зовсім вірним [48]. Ранні відбиття часто бажані, оскільки вони збільшують амплітуду сигналу, який сприймає слухач. Це збільшення амплітуди може збільшити розбірливість мови, якщо рівень мовленнєвого сигналу підвищується відносно рівня навколишнього шуму [49]. Ранні відбиття лише тоді не підвищують розбірливість, коли рівень звукового тиску чистого сигналу дорівнює рівню звукового тиску реверберованого сигналу [24].

1.5. Використання спектра пізньої реверберації для компенсації впливу пізньої реверберації в частотній області

Методи корекції мови, що спотворена шумом, на даний момент досить добре розвинуті і широко використовуються. Огляд таких методів представлений в [50]. Однак для практичної корекції мови, що спотворена реверберацією, такі методи стали використовуватись пізніше. Так, в 2001 році [51] було вперше запропоновано використовувати метод віднімання спектрів для придушення реверберації. Він є дуже зручним в силу мінімальних вимог до апріорної інформації про сигнал та заваду, але має той недолік, що призводить до появи артефакту, що має назву «музичні тони» та представляє собою вузькосмугові флуктуації в залишковому шумі. В дисертації [24] можна знайти два прийоми, які дозволяють значною мірою придушити цей небажаний ефект, але повністю позбавитись від нього не вдається. Незважаючи на це, метод спектрального віднімання є прийнятним для ряду ситуацій.

Іншим напрямком придушення шумової завади є так звані «статистичні» методи, які були викладені в [52], [53] і розвинуті в [24], [54]. Ці методи базуються на статистичних моделях спектрів мовленнєвих сигналів та полягають в мінімізації помилки відновлення (різниці між чистим та відновленим сигналами). Статистична модель сприйняття мови та значимі заходи спотворення, які є найбільш придатними для спектральної корекції, досі

не з'ясовані. Цим пояснюється значна різноманітність статистичних моделей і заходів спотворення, що використовуються, а також конкретних реалізацій алгоритмів спектральної корекції. Метод оцінювання OM-LSA (Optimally-Modified Log Spectral Amplitude, оптимально модифікований логарифм амплітудного спектру) є більш ефективним в порівнянні з методом спектрального віднімання.

В дисертації [24] розглянуто метод спектрального віднімання та статистичний метод, які скориговані таким чином, що їх можна використовувати як для придушення реверберації, так і для придушення фонового шуму.

У методах третього напрямку – спектральної корекції на основі прихованих марківських процесів (НМР) – робиться спроба обійти припущення про конкретні розподіли мови і шумових процесів. Щільності ймовірностей чистої мови та шуму оцінюються в процесі навчання, а потім спільно використовуються при відновленні спотвореної мови для заданої заходи спотворення. Істотні недоліки даного напрямку:

1) відновлення мови незадовільне, якщо модель шуму при розпізнаванні відрізняється від такої при навчанні;

2) істотна складність алгоритмів і високі вимоги до обчислювальних ресурсів.

Методи четвертого напрямку – це методи підпросторів, сутність яких полягає в поділі простору сигналу за наявності фонового шуму на підпростори «сигнал-плюс-шум» і «шум», з подальшим видаленням шумового підпростору і оцінюванням сигналу по підпростору, який залишився.

Процедуру придушення пізньої реверберації можна реалізувати практично тими ж алгоритмічними і програмними засобами, які використовуються для придушення шумової завади. Відмінність лише в тому, що замість спектра шумової завади достатньо оцінити спектр пізньої реверберації.

Для відстаней між джерелом звуку та мікрофоном, які більше критичної дистанції (коли рівень прямого звуку стає рівним рівню відбитого звуку), спектр потужності пізньої реверберації $\lambda_r(l, k)$ може бути обрахований через спектр потужності $\lambda_y(l, k)$, сигналу $y(t)$, що спостерігається [51]:

$$\lambda_r(l, k) = e^{-2\delta(k)T_l} \cdot \lambda_y(l - N_l, k), \quad (1.8)$$

де $N_l = T_l F_s / R$; R – зсув фреймів, що виражений у відліках; $\delta(k) = 2 \ln 10 / T_{60}(k)$; $T_{60}(k)$ – час реверберації.

Сенс співвідношення (1.8) досить простий: поточні звуки мови маскуються «ревербераційний хвостом» від попередніх звуків.

Для підвищення точності вимірювання спектру $\lambda_y(l, k)$ в роботах [24], [51] запропоновано робити усереднення:

$$\hat{\lambda}_y(l, k) = \eta_y(k) \hat{\lambda}_y(l - 1, k) + (1 - \eta_y(k)) |Y(l, k)|^2, \quad (1.9)$$

де $Y(l, k)$ – дискретне перетворення Фур'є l -го сегменту сигналу $y(t)$;

$$\eta_y(k) = \begin{cases} \eta_y^d(k), & |Y(l, k)|^2 \leq \hat{\lambda}_y(l - 1, k); \\ \eta_y^a(k), & \text{в решті випадків.} \end{cases} \quad (1.10)$$

Верхнє значення параметру $\eta_y^d(k)$ ($0 \leq \eta_y^d(k) < 1$) при цьому обмежують величиною

$$\eta_y^d(k) = \frac{1}{1 + 2\delta(k)R/F_s}, \quad (1.11)$$

а значення параметру $\eta_y^a(k)$ обирають з умови $0 \leq \eta_y^a(k) < \eta_y^d(k)$.

На жаль, в роботах [46], [51] відсутні вказівки щодо раціонального вибору конкретних значень параметрів T_l та $\eta_y^a(k)$.

1.6. Прямі методи вимірювання часу реверберації

В 1964 році Шредер запропонував новий метод вимірювання часу реверберації (ЧР) [55] шляхом випромінення білого шуму до моменту досягнення стаціонарності, після чого джерело вимикалось, а самописець

будував графік квадрата тиску. Оскільки випромінюваний сигнал має випадковий характер, записана реалізація також має випадковий характер. Тому ЧР визначається за нахилом обвідної дисперсії ревербераційного процесу як функції часу.

Оскільки при вимірюванні ЧР білий шум зазвичай фільтрують октавним або третьоктавним фільтром для отримання шуму з необхідною шириною спектру [56], і в початковий момент реєстрації сигналу джерело шуму вимикається, прийнятий сигнал $s(t)$ можна описати виразом:

$$s(t) = \int_{-\infty}^0 n(\tau)h(t-\tau)d\tau, \quad (1.13)$$

де $h(t)$ – ІХ системи. Верхня межа інтегрування – момент часу, в який вимикають джерело шуму. Нижня межа інтегрування каже про те, що шум $n(t)$ діяв настільки довго, що в момент часу $t = 0$ поле в приміщенні можна вважати стаціонарним випадковим процесом.

Тоді дисперсія процесу:

$$D_s(t) = N \int_{-\infty}^0 h^2(t-\tau)d\tau = N \int_t^{\infty} h^2(x)dx \quad (1.14)$$

Згідно (1.14), дисперсія ревербераційного процесу – визначений інтеграл квадрату ІХ системи, помножений на потужність шуму.

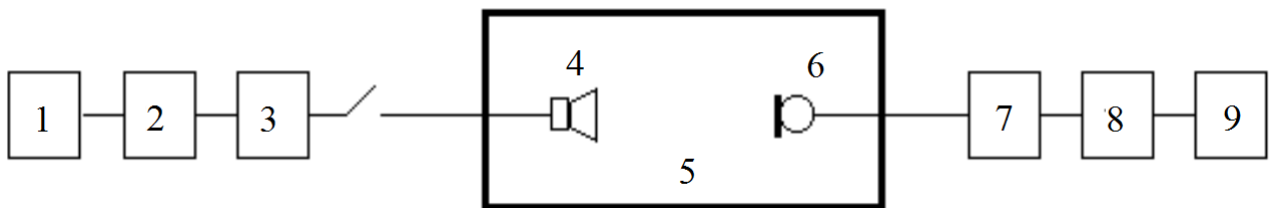


Рис. 1.6 – Схема вимірювання ЧР методом Шредера, де 1 – джерело шуму, 2, 8 – смугові фільтри, 3, 7 – підсилювачі, 4 – гучномовець, 5 – кімната, в якій проводять вимірювання, 6 – вимірювальний мікрофон, 9 – блок обробки даних.

Згідно [56] гучномовець має випромінювати шум, фільтрований третьоктавними фільтрами. В якості звукового сигналу допускається використання нефільтрованого шуму та імпульсу вибухового характеру (хлопок, вибух повітряної кульки або хлопавки, постріл зі стартового пістолета). Схема проведення вимірювань ЧР за [56] наведена на рис. 1.6.

Міжнародний стандарт [57] допускає використання лінійно-частотно-модульованого (ЛЧМ) імпульсу (рис. 1.7) та так званої «послідовності максимальної довжини» (М-послідовності або MLS-сигналу, рис. 1.8) в якості вимірювального сигналу. Використання таких сигналів відносять до методу інтегрування імпульсного відгуку, за якого вимірювальний сигнал не вимагає попередньої фільтрації, тому блок 2 зі схеми, наведеної на рис.1.6 можна виключити.

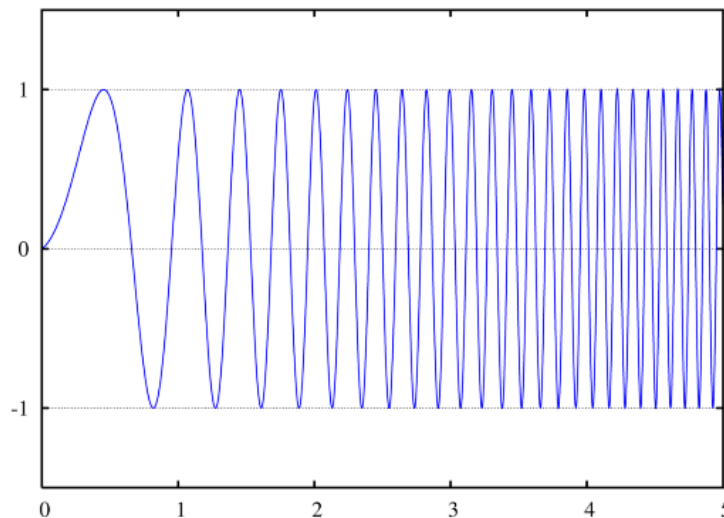


Рис. 1.7 – Зміна частоти ЛЧМ-сигналу за період.

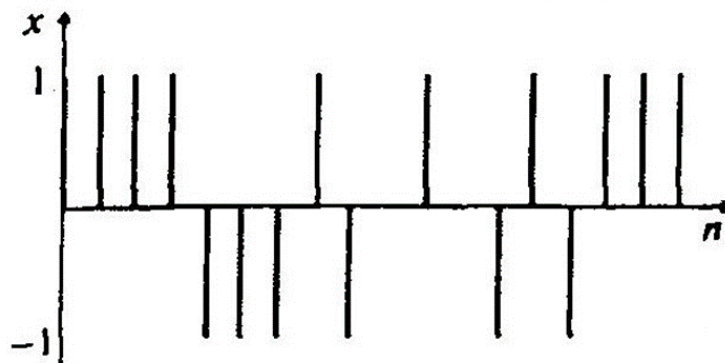


Рис. 1.8 – Приклад реалізації М-послідовності.

При використанні ЛЧМ-сигналу тривалість впливу має бути набагато більшою за тривалість відгуку системи. Таким чином, якщо час реверберації дорівнює одній секунді, частота сигналу має змінюватись повільно відносно однієї секунди [58].

1.7. Сліпі методи вимірювання часу реверберації

Зазвичай час реверберації вимірюють напряму, використовуючи ІХ приміщення $h(t)$, як показано в попередньому підрозділі роботи. Якщо ІХ невідома і немає змоги її записати, використовують так звані «сліпі» методи вимірювання ЧР.

Першим сліпим підходом до визначення ЧР став такий, при якому передбачається, що частотна характеристика фільтра, зворотного до ІХ приміщення, є мінімально-фазовою [59], [60], але на практиці зазвичай це твердження виявляється хибним. В 1999 році Дж. Наннаріелло та Ф. Фріке [61] а в 2001 році Т. Кокс, Ф. Лі та П. Дарлінгтон [62] запропонували використовувати нейронні мережі для отримання інформації про час реверберації із запису мовленнєвого сигналу. Тоді ж К. Лебарт, Дж. Буше та П. Денбі запропонували виділяти паузи у записі мови і визначати час реверберації під час цих пауз [51]. Метод максимальної правдоподібності (МП) вперше був запропонований Р. Ратнамом, Д. Джонсом та Б. Уілером в 2003 році [63]. Згідно [64], метод МП є достатньо завадостійким, потребує менше часу на виконання алгоритму та не поступається в точності вимірювань методам, описаним в [65] та [66].

Метод МП базується на параметризації ревербераційного процесу [63]. Тому реверберацію представляють у вигляді добутку широкосмугового гаусівського випадкового процесу $x(n)$, $n \geq 0$ з параметрами $[0, \sigma]$ і детермінованої функції $a(n) > 0$:

$$y(n) = x(n)a(n). \quad (1.15)$$

Тоді середнє значення процесу $y(n)$:

$$\langle y(n) \rangle = \langle x(n) \rangle a(n) = 0, \quad (1.16)$$

а дисперсія:

$$D[y(n)] = D[x(n)]a^2(n) = \sigma^2 \cdot a^2(n). \quad (1.17)$$

Одномірна щільність імовірності процесу $y(n)$ в n -й момент часу:

$$P[y(n)] = \frac{1}{\sqrt{2\pi\sigma \cdot a(n)}} \exp\left(-\frac{y^2(n)}{2\sigma^2 a^2(n)}\right). \quad (1.19)$$

Якщо \mathbf{y} – N -вимірний вектор, тоді оскільки відліки процесу $x(n)$ статистично незалежні, багатовимірна щільність імовірності цього вектора, яка є функцією правдоподібності, записується у вигляді добутку N одновимірних гаусівських розподілів:

$$L(\mathbf{y}; \mathbf{a}, \sigma) = \frac{1}{a(0) \cdots a(N-1)} \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left(-\frac{\sum_{n=0}^{N-1} (y(n)/a(n))^2}{2\sigma^2}\right), \quad (1.20)$$

де \mathbf{a} і σ – $N + 1$ невідомих параметрів, які мають бути оцінені зі спостереження \mathbf{y} .

Нехай параметр загасання τ описує затухання обвідної звукового сигналу в процесі вільного загасання. Тоді послідовність $a(n)$ записується у вигляді виразу:

$$a(n) = \exp(-n/\tau). \quad (1.21)$$

Введемо позначення $a = \exp(-1/\tau)$, тоді (1.21) записується у вигляді:

$$a(n) = a^n. \quad (1.22)$$

Підставивши (1.22) в (1.20), отримаємо:

$$L(\mathbf{y}; a, \sigma) = \left(\frac{1}{2\pi a^{(N-1)} \sigma^2} \right)^{N/2} \exp \left(- \frac{\sum_{n=0}^{N-1} a^{-2n} y(n)^2}{2\sigma^2} \right). \quad (1.23)$$

Для фіксованого вікна спостереження N і послідовності спостережень $y(n)$, параметрами функції правдоподібності є параметр загасання a і потужність розсіювання σ . Вони можуть бути оцінені шляхом використання максимально правдоподібного наближення [67]. Для цього спочатку необхідно прологарифмувати (1.23):

$$\ln L(\mathbf{y}; a, \sigma) = - \frac{N(N-1)}{2} \ln(a) - \frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} a^{-2n} y(n)^2. \quad (1.24)$$

Оптимальні оцінки параметрів a і σ – це такі значення цих параметрів, які відповідають максимуму функції (1.24).

Продиференціюємо логарифм функції правдоподібності (1.24) по a для знаходження максимуму $\ln(L)$ і отримаємо функцію s_a :

$$s_a(a; \mathbf{y}, \sigma) = \frac{\partial \ln L(\mathbf{y}; a, \sigma)}{\partial a} = - \frac{N(N-1)}{2a} + \frac{1}{a\sigma^2} \sum_{n=0}^{N-1} n a^{-2n} y(n)^2. \quad (1.25)$$

Для того, щоб знайти екстремум логарифма функції правдоподібності, покладемо $s_a = 0$:

$$- \frac{N(N-1)}{2a} + \frac{1}{a\sigma^2} \sum_{n=0}^{N-1} n a^{-2n} y(n)^2 = 0. \quad (1.26)$$

Вираз (1.26) можна задовольнити деяким значенням a^* . Логарифм функції правдоподібності буде максимальним тоді, коли друга його похідна в точці a^* буде від'ємною. Згідно [63], можна показати, що ця умова виконується, однак згідно [68], неможливо для всіх можливих значень a^* аналітично довести виконання цієї умови, тому її необхідно перевіряти.

Аналогічно можна оцінити потужність розсіювання ревербераційного хвоста. Диференціюючи (1.24) по σ отримаємо:

$$s_{\sigma}(\sigma; \mathbf{y}, a) = \frac{\partial \ln L(\mathbf{y}; a, \sigma)}{\partial \sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma^3} \sum_{n=0}^{N-1} a^{-2n} y(n)^2. \quad (1.27)$$

Вираз (1.27) досягає екстремуму, коли

$$\sigma^2 = \frac{1}{N} \sum_{n=0}^{N-1} a^{-2n} y(n)^2. \quad (1.28)$$

В роботі [69] запропоновано і зіставлено два алгоритми оцінювання параметрів a і σ : онлайн алгоритм та блочний алгоритм.

В роботі [70], яка заснована на роботі [68] представлений покращений алгоритм сліпої оцінки ЧР. Основна ідея покращення полягає в тому, щоб розглядати лише ті ділянки сигналу, на яких енергія загасає, оскільки саме вони містять інформацію про ЧР. Тоді інші ділянки мовленнєвого сигналу не аналізуються. Тим самим зменшується обсяг обчислень, підвищується точність вимірювань та стійкість алгоритму до дії завади.

За [70] мовленнєвий сигнал поділяють на фрейми:

$$x_f(\lambda, m) = x(\lambda M_{\Delta} + m), \quad m = 0, 1, \dots, M - 1, \quad (1.29)$$

де λ – номер фрейму, M_{Δ} – зсув фрейму, M – кількість відліків в фреймі, m – номер відліку всередині фрейму.

Фрейм в свою чергу розбивається на підфрейми:

$$y(\lambda, l, k) = x_f(\lambda, lP + k), \quad k = 0, 1, \dots, P - 1, \quad l = 0, 1, \dots, L - 1, \quad (1.30)$$

де l – номер підфрейму, який складається з P відліків, k – номер відліку в підфреймі.

Після цього необхідно перевірити чи спадає енергія в підфреймах всередині фрейму за такими правилами:

$$\begin{aligned} \sum_{k=0}^{P-1} y^2(\lambda, l, k) &> w_{l+1}^{(\text{var})} \cdot \sum_{k=0}^{P-1} y^2(\lambda, l+1, k), \\ \max_k \{y(\lambda, l, k)\} &> w_{l+1}^{(\text{max})} \cdot \max_k \{y(\lambda, l+1, k)\}, \\ \min_k \{y(\lambda, l, k)\} &> w_{l+1}^{(\text{min})} \cdot \min_k \{y(\lambda, l+1, k)\}, \end{aligned} \quad (1.31)$$

де номер підфрейму $l = 0, 1, \dots, L - 2$, а ваговий фактор $0 \leq w_l \leq 1$. Якщо хоча б одна з умов (1.31) порушується, необхідно перевірити, чи досяг лічильник l свого мінімального значення $1 < l_{min} < L - 2$. Якщо жоден підфрейм всередині фрейму не задовольняє (1.31), аналізується наступний фрейм. В іншому випадку, якщо для деяких підфреймів виконується (1.31), вони визначаються як сегмент, на якому звук загасає. Після цього на сегменті оцінюють ЧР методом МП.

Гістограма оновлюється з отриманням кожної наступної оцінки. Кількість оцінок в гістограмі становить K_f . ЧР визначається за максимальним значенням гістограми $\hat{T}_{60}^{(1)}$, а не за першим її значенням, як у роботах [63], [68], [69]. Дисперсія оцінки знижується за допомогою процедури згладжування:

$$\hat{T}_{60}(\lambda) = \beta(\lambda) \cdot \hat{T}_{60}(\lambda - 1) + (1 - \beta(\lambda)) \cdot \hat{T}_{60}^{(1)}(\lambda), \quad 0 < \beta(\lambda) < 1. \quad (1.32)$$

В [70] представлено порівняння двох алгоритмів оцінки ЧР: старого [68] і нового [70]. Новий алгоритм дуже добре відслідковує зміну ЧР в часі. Також показано, що новий метод дає вірні значення оцінки ЧР навіть за наявності шумової завади при відношенні сигнал/завада 25 дБ, хоча і з деякою похибкою.

Висновки до розділу 1

1. Імпульсна характеристика приміщення є набором всіх точкових імпульсних характеристик, що існують для всіх шляхів джерело-приймач. Тому знання ІХ приміщення дозволяє оцінити розбірливість мовлення в різних точках приміщення.

2. Перелік моделей ІХ приміщення є значним. В даній дисертаційній роботі для подальшого використання обрано комбіновану модель, згідно якої ІХ поділяється на дві частини: ранні дискретні відбиття та дифузний ревербераційний хвіст (пізня реверберація).

3. Дія ранньої реверберації на якість та розбірливість мовленнєвого сигналу досліджена не у повній мірі. Зазвичай ранні відбиття розглядають як фактор, що сприяє підвищенню розбірливості мовлення. Між тим, можна припустити, що ранні відбиття можуть грати негативну роль, погіршуючи розбірливість мовлення в окремих точках приміщення.

4. Час реверберації є важливою акустичною характеристикою приміщення, й на сьогодні класичним прямим методом вимірювання ЧР є метод Шредера. Проте, існуючим алгоритмам вимірювання ЧР прямим методом притаманні суттєві недоліки, спричинені наявністю фонові шумової завади.

5. Сліпі методи вимірювання ЧР є зручними за відсутності оцінок ІХ приміщень. Однак існуючим алгоритмам вимірювання ЧР сліпим методом притаманні обмеження, які знижують ефективність акустичних вимірювань параметрів приміщень.

6. В розглянутих роботах, де аналізується вплив приміщення на мовленнєвий сигнал, бракує аналізу зв'язку об'єктивних показників якості мовлення із суб'єктивною оцінкою розбірливості. Окрім цього, традиційна методологія проведення артикуляційних випробувань є ресурсомісткою й недостатньо автоматизованою.

2. ПРЯМІ МЕТОДИ ВИМІРЮВАННЯ АКУСТИЧНИХ ХАРАКТЕРИСТИК

2.1. Вимірювання акустичних характеристик приміщення за умови наявності фонові завади

Оцінки EDT, T20 і T30 обчислювалися методом Шредера, по обвідній $R_\phi(t)$ функції $h_\phi(t)$, отриманої з імпульсної характеристики $h(t)$ шляхом фільтрації її набором октавних фільтрів [71]:

$$R_\phi(t) = \int_{-\infty}^{-t} h_\phi^2(-\tau) d\tau.$$

Схема розрахунку оцінок EDT, T10, T20 і T30 представлена на рис. 2.1.

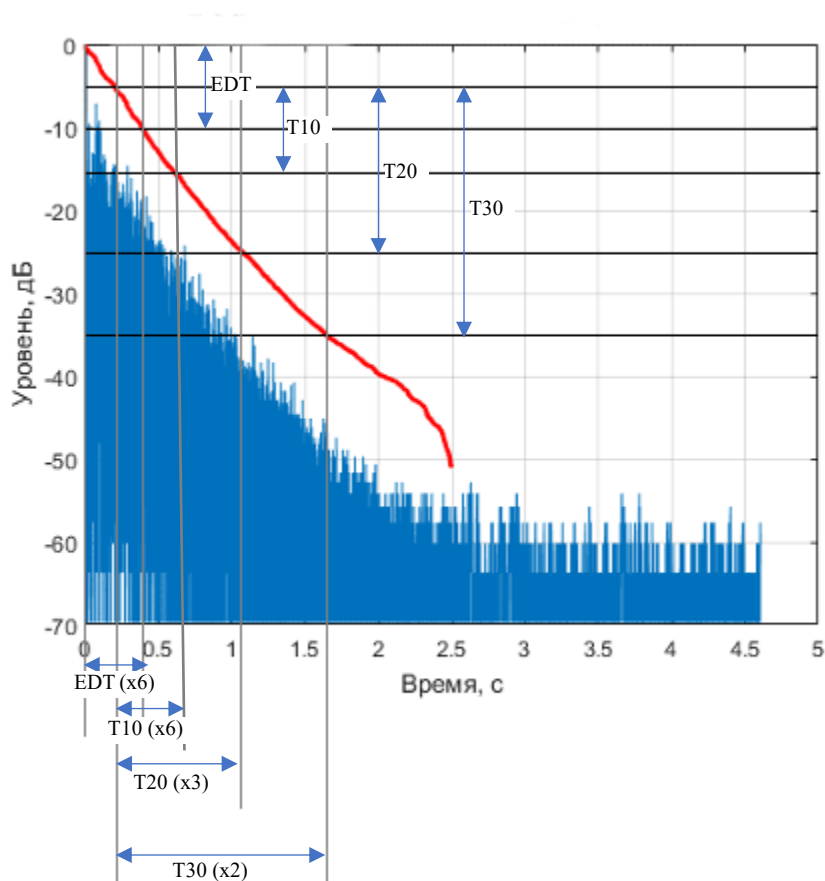


Рис. 2.1 – Схема розрахунку ЧР по обвідній ІХ.

При вимірюванні ЧР прямим методом із використанням формули Шредера (1.14) фоновий шум призводить до утворення «полиці», яка тягнеться за обвідною ревербераційного процесу, і невірному визначення часу реверберації (рис. 2.2).

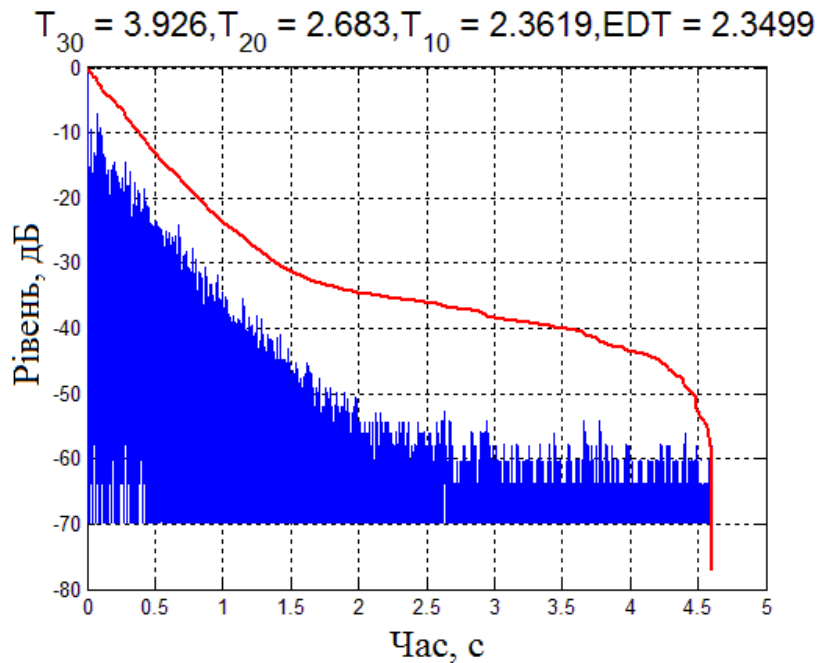


Рис. 2.2 – Результат безпосереднього використання методу Шредера для знаходження ЧР з приміщенні з фоновим шумом.

Шумовий хвіст можна обрізати вручну в момент часу, коли рівень ревербераційного процесу стає порівняним з рівнем фонового шуму. Це дуже простий вихід з положення, однак він вимагає проміжного втручання людини в процедуру обчислень.

Спосіб боротьби з таким шумовим хвостом залежить від того, який тестовий сигнал було випромінено.

Найкращим буде рішення, при якому в якості тестового сигналу випромінюють лінійно-частотно-модульований сигнал або послідовність максимальної довжини (М-послідовність, MLS-сигнал). У першому випадку може бути застосована методика, яку називають спектрометрією із затримкою у часі (Time Delay Spectrometry – TDS) [72]. У другому випадку імпульсна характеристика приміщення оцінюється за допомогою взаємно кореляційної функції між вихідним MLS-сигналом і сигналом в точці прийому [73]. Пік-фактор М-послідовності дорівнює одиниці, що дає високе значення відношення сигнал-завада при вимірюваннях в зашумленому приміщенні.

Для вирішення проблеми шумності записаного відгуку при використанні в якості тестового сигналу короткого імпульсу випадкового характеру (сплеск руками, вибух повітряної кульки або хлопавки, постріл стартового пістолета) дисертантом запропоновано використати умови знаходження спаду сигналу:

$$\begin{aligned} \overline{y^2}(l, k) &> w_{l+1}^{(avg)} \cdot \overline{y^2}(l + 1, k), \\ \max_k \{y^2(l, k)\} &> w_{l+1}^{(max)} \cdot \max_k \{y^2(l + 1, k)\}, \\ \min_k \{y^2(l, k)\} &> w_{l+1}^{(min)} \cdot \min_k \{y^2(l + 1, k)\}, \end{aligned} \quad (2.1)$$

і інтегрувати сигнал виключно на спадаючому відрізку. Тоді можна нівелювати вплив шумового «хвоста» [74]. У співвідношеннях (2.2) використано такі позначення y – сигнал, який ми спостерігаємо (позначений синім на рис. 2.2, у нашому випадку, фактично, це ІХ приміщення), l – номер фрейму, k – номер відліку у фреймі, w – ваговий фактор.

Опишемо детальніше запропоновану процедуру. Згідно формул (2.1) сигнал розбивається на фрейми, знаходиться мінімальне, середнє та максимальне значення потужності сигналу на сусідніх фреймах. Якщо (2.1) виконується, приймається рішення про те, що на проаналізованих фреймах сигнал загасає. Фоновий шум призводить не тільки до утворення «полиці» на записі ІХ приміщення, але і до сповільнення загасання сигналу при низьких значеннях SNR (на рис. 2.1 видно, що на часовому відрізку від 2 до 3 с від початку запису загасання сигналу уповільнюється), що також знижує точність вимірювання ЧР. Для точного визначення моменту, коли швидкість загасання сигналу зменшується, автор пропонує робити віконний аналіз сигналу з перекриттям щонайменше 50%. Аналіз продовжується до першого фрейму, на якому (2.1) не виконуються.

Аналізуючи виключно середнє значення квадрату сигналу (верхній рядок виразів (2.1)), можна хибно визначити «корисний» відрізок записаного сигналу через стохастичний характер пізніх відбиттів. Так, при проведенні експериментів з використанням виключно першого рядка виразів (2.1), були

отримані хибні значення оцінки ЧР, оскільки фактично сигнал продовжував повільно загасати навіть на відрізку, що відповідає «полиці». Точу були введені контрольні умови (другий і третій рядки виразів (2.1)), за яких мають зменшуватись і середнє, і максимальне, і мінімальне значення квадрату записаного відгуку. Вагові коефіцієнти w також були введені в умови (2.1) для корекції неточностей при визначенні корисного відрізка записаного сигналу.

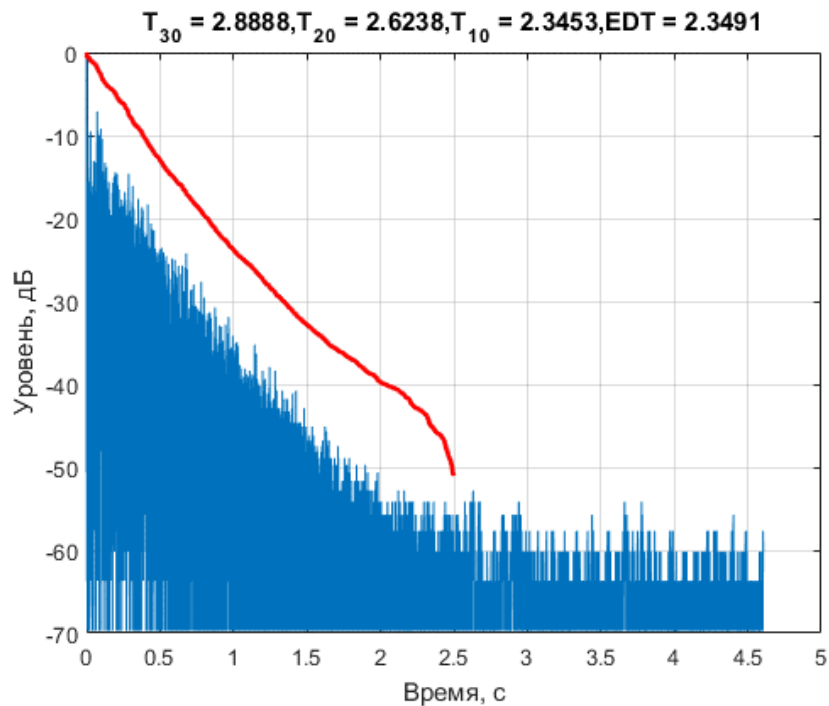


Рис. 2.3 – Результат «розумного» автоматичного вимірювання ЧР методом Шредера.

Алгоритм для автоматичного, вимірювання часу реверберації методом Шредера за умови наявності фоновий завади був змодельований у Matlab. Результат роботи програми наведений на рис. 2.3.

Як видно на рис. 2.3, програма досить точно розпізнала момент часу, у який рівень затухаючого сигналу стає рівним рівню фоновий завади, в результаті чого було отримане більш точне значення часу реверберації. Особливо добре це видно для T_{30} : різниця між часом реверберації, що був виміряний за допомогою безпосереднього застосування методу Шредера (рис. 2.2), і таким, що був виміряний за вдосконаленим алгоритмом (рис. 2.3), склала більше 1 с.

З точки зору завадостійкості бажано використовувати методики вимірювання ІХ, в яких в якості тестового сигналу випромінюється ЛЧМ-сигнал або М-послідовність.

2.2. Вимірювання бінауральних імпульсних характеристик приміщення

Акустичні характеристики приміщень, виміряні одноканальними системами, неповно відображають якість сприйняття акустичної інформації людиною у приміщенні. Бінауральне сприйняття суттєво впливає на здатність людини до сприйняття мовлення за умови наявності завади. Для врахування явища бінаурального слуху при записі відгуку приміщення на звуковий стимул для подальшого обчислення ІХ приміщення використовують штучну голову (ШГ).

Окрім цього, ШГ з моделлю зовнішнього вуха використовують при розв'язанні таких задач як запис звуку в концертній залі, контроль якості навушників, контроль якості та розбірливості мови в лініях зв'язку [16], [75], [76].

За допомогою ШГ були записані бінауральні ІХ кількох приміщень різного призначення, з цих записів були сформовані бази даних [77], [78], [25]. За допомогою таких записів можна змодельовати звук у різних точках приміщення (ефект «акустичне дзеркало» [79]), оцінити якість та розбірливість мовленнєвої інформації, дослідити алгоритми дереверберації [80], [22], [23], [21], [2], [6].

На жаль, використання бінауральної системи ускладнює процедуру зняття ІХ приміщення, тому згадана база бінауральних ІХ [77], [75] є досить обмеженою і містить записи лише для однієї аудиторії.

Спрощену структурну схему процедури оцінювання ІХ приміщення за допомогою ШГ представлено на рис. 2.4.

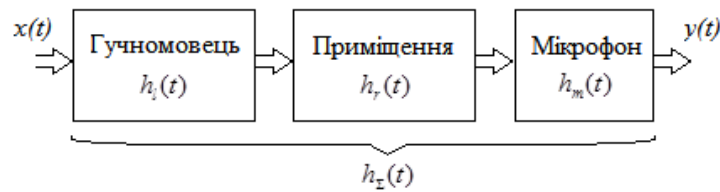


Рис. 2.4 – Структурна схема процедури оцінювання ІХ приміщення

Тестовий звуковий сигнал $x(t)$ випромінюють за допомогою гучномовця, розміщеного в точці приміщення, де зазвичай знаходиться диктор. Відгук приміщення $y(t)$ на стимул $x(t)$ сприймають парою мікрофонів, притиснутих до вушних раковин штучної голови, розміщеної в одній із точок приміщення, де зазвичай знаходяться слухачі.

Оскільки амплітудно-частотні характеристики гучномовця та мікрофона не є ідеально рівномірними в смузі частот тестового сигналу, замість ІХ приміщення $h_r(t)$ насправді буде оцінено згортку

$$h_{\Sigma}(t) = h_l(t) \otimes h_r(t) \otimes h_m(t) = h_r(t) \otimes h_{lm}(t), \quad (2.2)$$

де \otimes – символ згортки; $h_l(t)$ – ІХ гучномовця; $h_m(t)$ – ІХ одного із мікрофонів; $h_{lm}(t) = h_l(t) \otimes h_m(t)$ – ІХ системи «гучномовець-мікрофон» (ГМ).

Спосіб обчислення ІХ приміщення $h_r(t)$ є досить очевидним. Дійсно, оскільки в частотній області співвідношенню (2.2) відповідає добуток відповідних частотних характеристик

$$H_{\Sigma}(f) = H_l(f) \cdot H_r(f) \cdot H_m(f) = H_r(f) \cdot H_{lm}(f), \quad (2.3)$$

де $H(f) = F\{h(t)\}$, F – символ прямого перетворення Фур'є, із (2.3) отримуємо

$$h_r(t) = F^{-1}\{H_r(f)\} = F^{-1}\left\{\frac{|H_{\Sigma}(f)|}{|H_{lm}(f)|} \cdot \exp[\theta_{\Sigma}(f) - \theta_{lm}(f)]\right\}, \quad (2.4)$$

де $\theta_{\Sigma}(f)$ та $\theta_{lm}(f)$ – фазочастотні характеристики (ФЧХ) функцій $H_{\Sigma}(f)$ та $H_{lm}(f)$, відповідно.

Проблемним місцем для реалізації обчислень за виразом (2.4) є операція ділення, оскільки модуль $|H_{lm}(f)|$ (амплітудно частотна характеристика (АЧХ) підсистеми ГМ) функції $H_{lm}(f)$ може містити малі числові значення, що

приведе до переповнення розрядної сітки обчислювальної системи та аварійної зупинки комп'ютера. Разом із тим, зрозуміло, що якщо забезпечити належний контроль за властивостями функції $|H_{lm}(f)|$, рівняння (2.4) може використовуватися для обчислень.

Що стосується ФЧХ $\theta_{\Sigma}(f)$ та $\theta_{lm}(f)$, тут є два варіанти: 1) враховувати інформацію про обидві ФЧХ; 2) враховувати інформацію лише про $\theta_{\Sigma}(f)$, припускаючи, що $\theta_{lm}(f) = 0$.

В даній дисертації і опублікованій роботі [9] запропоновано оцінювати ІХ приміщення $h_r(t)$ із використанням виразу

$$h_r(t) = F^{-1} \{H_r(f)\} = F^{-1} \left\{ \frac{H_{\Sigma}(f)}{|H_{lm}(f)|} \cdot M_R(f) \right\}, \quad (2.5)$$

де $H_{\Sigma}(f) = F\{h_{\Sigma}(t)\}$ – частота характеристика системи рис. 2.3; F – символ прямого перетворення Фур'є; $|\cdot|$ – символ модуля; $M_R(f)$ – множник регуляризації [81], призначений для зменшення дисперсії оцінки ІХ. В [9] запропоновано в якості такого множника використовувати вікно Ханна:

$$M_r(f) = 0,5 \left[1 - \cos(\pi(f - F_s/2)/\Delta F) \right], \quad f = 0, \dots, F_s,$$

де ΔF – параметр регуляризації; F_s – частота дискретизації (для $F_s=44,1$ кГц в [9] рекомендовано прийняти $\Delta F = 18$ кГц).

Організація обчислень АЧХ підсистеми «гучномовець-мікрофон».

В роботі [80] вимірювання ІХ приміщення проводилися за допомогою професійної аудіоапаратури, щоб отримати високоякісні та низькошумні результати. В якості джерела звуку було використано акустичний двоканальний активний студійний монітор Genelec 8130, що має цифровий інтерфейс AES/EBU. Для запису звуку були використані два всенаправлених конденсаторних вимірювальних мікрофони Beyerdynamic MM1 та мікрофонний підсилювач RME Octamic II у поєднанні з аудіоінтерфейсом RME Multiface II. Всі вимірювання проводилися з частотою дискретизації 48 кГц і 24-бітовою точністю. Для бінауральних записів використано штучну голову HMS2 фірми

«HEAD acoustics». Два мікрофони розташовувалися впритул до вушної раковини на відстані 1 см від вушного каналу.

Базовим елементом при побудові тестового сигналу слугував MLS-сигнал із ступенем 16 (тобто сигнал містив 2^{16} вибірок), що відповідає довжині сигналу 1,36 с при частоті дискретизації 48 кГц. Тестовий сигнал створювався шляхом повторювання цього базового елемента 17 разів. Імпульсну характеристику приміщення обчислювали, усереднюючи 16 останніх сплесків оцінки взаємнокореляційної функції (ВКФ) між сигналом з виходу мікрофона та тестовим сигналом, що дозволило підвищити відношення сигнал-шум на 12 дБ.

Автор даної дисертації використав такі ж тестові сигнали, як і в [77], з тією лише відмінністю, що частота дискретизації становила 44,1 кГц. Якщо вважати ланцюжок «гучномовець-приміщення-мікрофон» лінійною системою із ІХ $h_{\Sigma}(t)$, тоді, як відомо [78], така ІХ може бути обчисленою за виразом

$$K_{XY}(\tau) = \langle x(t)y(t+\tau) \rangle = \int_{-\infty}^{\infty} h_{\Sigma}(v) \langle x(t)x(t+\tau-v) \rangle dv = \int_{-\infty}^{\infty} h_{\Sigma}(v) K_X(\tau-v) dv, \quad (2.6)$$

де $\langle \cdot \rangle$ – символ математичного очікування; $K_X(\tau)$ – автокореляційна функція тестового сигналу $x(t)$; $K_{XY}(\tau)$ – ВКФ тестового сигналу $x(t)$ та сигналу $y(t)$ на виході мікрофона. Дійсно, за умови

$$K_X(\tau) = D_X \cdot \delta(\tau),$$

де $\delta(\tau)$ – дельта-функція Дірака, D_X – дисперсія тестового сигналу $x(t)$, із (2.3) отримуємо

$$K_{XY}(\tau) = D_X \cdot h_{\Sigma}(\tau). \quad (2.7)$$

Фур'є-перетворення від (2.7) приводить до виразу

$$S_{XY}(f) = H_{\Sigma}(f) \frac{N_0}{2}, \quad (2.8)$$

де $S_{XY}(f)$ – взаємний спектр сигналів $x(t)$ та $y(t)$; N_0 – спектральна щільність білого шуму $x(t)$.

Якщо вимірювання виконувати в заглушеній кімнаті, на малій відстані між мікрофоном та гучномовцем, тоді, згідно (2.7) та (2.8), можна отримати інформацію про ІХ та частотну характеристику підсистеми ГМ:

$$K_{xya}(\tau) = D_X \cdot h_{lm}(\tau), \quad S_{xya}(f) = H_{lm}(f) \frac{N_0}{2}, \quad (2.9)$$

оскільки форма вихідного сигналу $y(t)$ буде визначатися властивостями підсистеми ГМ.

Оскільки при практичних вимірюваннях запис та обробка сигналів виконуються у цифровій формі, при використанні виразів (2.9) достатньо забезпечити, щоб спектр потужності тестового сигналу $x(t)$ був рівномірним у смузі частот $F_s/2$, де F_s – частота дискретизації. Для обраного тестового mls-сигналу можна вважати, що ця вимога виконується із достатньою для практичної діяльності точністю, оскільки ширина центрального сплеску є мінімально можливою та становить $1/F_s$, а рівень бокових пелюсток є близьким до -47 дБ (рис. 2.5).

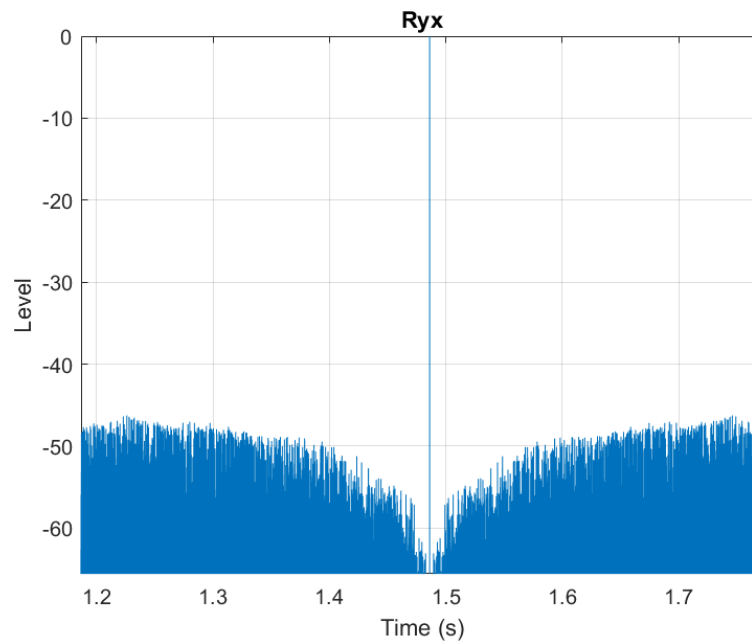


Рис. 2.5 – Автокореляційна функція MLS-сигналу

Приклади форми сигналів $x(t)$ та $y(t)$ на вході та виході, відповідно, підсистеми ГМ наведено на рис. 2.6, а відповідних спектрів потужності – на

рис. 2.7. Як бачимо із рис. 2.6, вибірки сигналу $y(t)$, на відміну від вибірок сигналу $x(t)$, мають різну висоту, що свідчить про спотворення вхідного сигналу $x(t)$ підсистемою ГМ. Наведені на рис. 2.7 оцінки спектрів потужності більш інформативні, оскільки дозволяють визначити, на яких частотах відбувається найбільш сильне спотворення вхідного сигналу. Дійсно, якщо нерівномірність спектру вхідного сигналу (рис. 2.7а) не перевищує $\pm 0,5$ дБ, то нерівномірність спектру вихідного сигналу в смузі частот від 300 Гц до 10 кГц становить ± 3 дБ, є близькою до ± 6 дБ в смузі частот від 10 кГц до 16 кГц, а в смузі від 50 Гц до 300 Гц спостерігається провал глибиною -15 дБ.

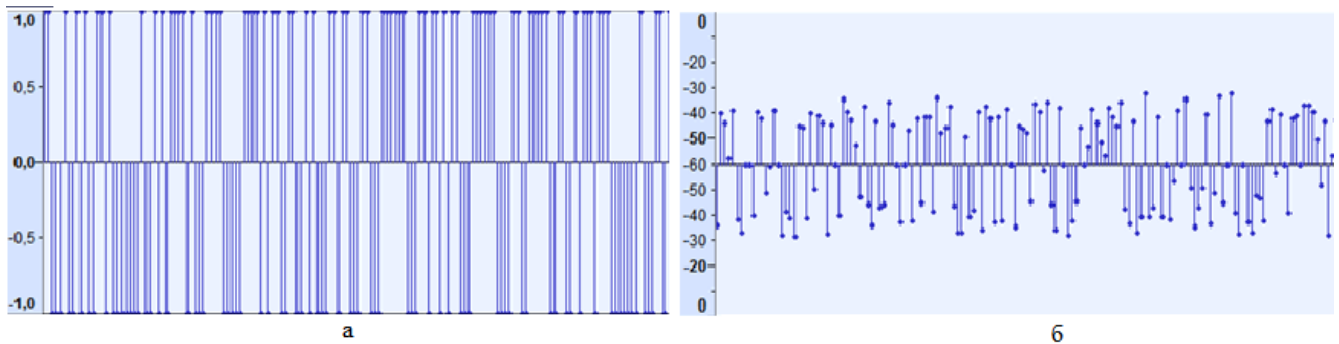


Рис. 2.6 – Сигнали $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ

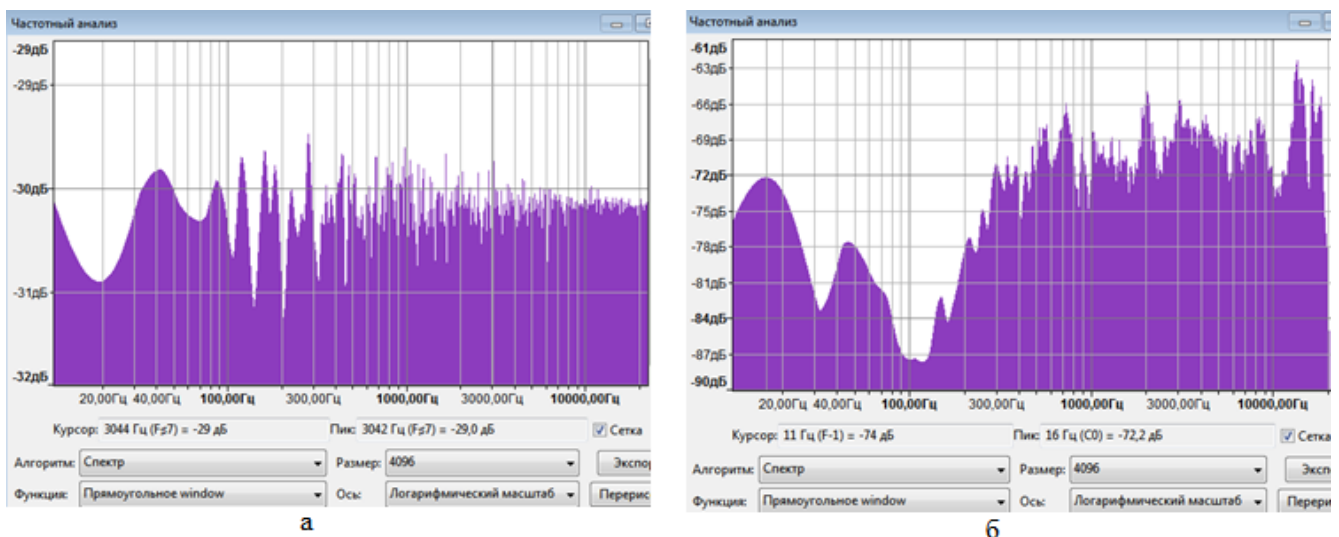


Рис. 2.7 – Спектри потужності сигналів $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ

Згідно із (2.9), форму АЧХ підсистеми можна також визначити шляхом взаємно-кореляційного аналізу сигналів $x(t)$ та $y(t)$. На рис. 2.8 наведено форму оцінки ВКФ.

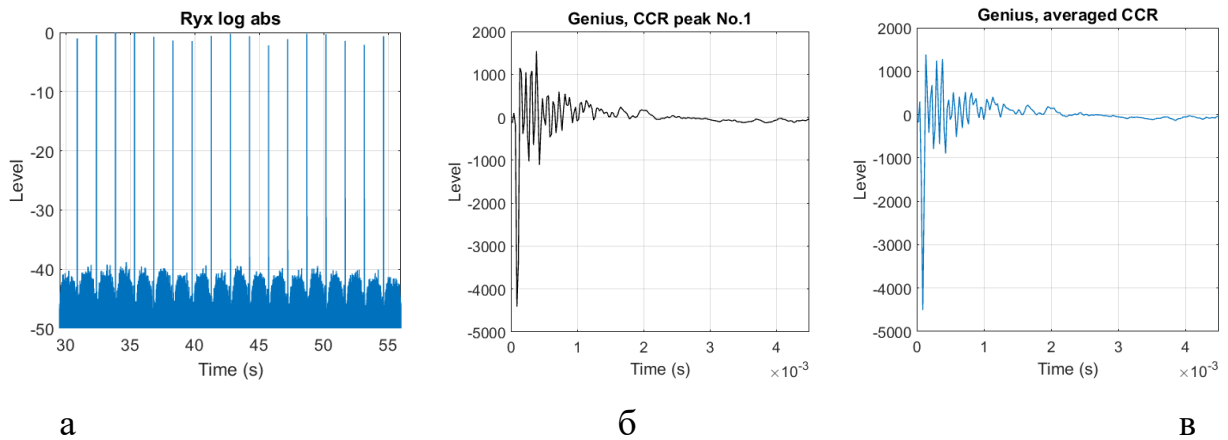


Рис. 2.8 – ВКФ сигналів $x(t)$ та $y(t)$ (а), форма першого (б) та усередненого (в) сплесків ВКФ

Модуль Фур'є-перетворення усередненого сплеску ВКФ, згідно (2.9), є оцінкою АЧХ підсистеми ГМ.

Результати досліджень.

На рис. 2.9 показано оцінки АЧХ $|H_{lm}(f)|$ підсистеми ГМ для 3-х видів гучномовців:

1. малогабаритний MP3-плеєр Optima МК-4 (вихідна потужність 5 Вт, частотний діапазон 150-18000 Гц, відношення сигнал-шум 86 дБ);
2. активна колонка Genius SP-HF 2.0 500 (вихідна потужність 14 Вт, частотний діапазон 65-20000 Гц, відношення сигнал-шум 73 дБ);
3. активні акустичні монітори Yamaha MSP5 (вихідна потужність 67 Вт, частотний діапазон 50-40000 Гц, відношення сигнал-шум 102 дБ).

На кожному із рисунків наведено графіки оцінок АЧХ $|H_{lm}(f)|$ чотирьох типів:

1. Туре 0 – оцінка спектра потужності Велча як усереднення спектрів потужності 17 сегментів сигналу $y(t)$, кожен сегмент має довжину 65536

вибірок та не піддається ваговому зважуванню:

$[P_y, f_{yy}] = \text{pwelch}(y, \text{rectwin}(65536), 0, 65536, F_s);$

2. Type 1 – згладжений за частотою, трикутним вікном шириною 50 Гц, модуль дискретного перетворення Фур'є (ДПФ) від результату усереднення 17 піків ВКФ;

3. Type 2 – згладжений за частотою, трикутним вікном шириною 50 Гц, результат усереднення (17 разів) модулів ДПФ від 17 сплесків ВКФ;

4. Type 3 – оцінка спектра потужності Велча із використанням вагового вікна Ханна, що забезпечує роздільну здатність 50 Гц: $[P_{yu}, f_{yu}] = \text{pwelch}(y, \text{hann}(\text{delw}), \text{round}(\text{delw}/2), 65536, F_s)$, $\text{delw} = (1,5F_s)/df$, $F_s = 44100$ Гц, $df = 50$ Гц.

Процедуру згладжування трикутним вікном, використану в оцінках 1-го та 2-го типів, було реалізовано шляхом цифрової фільтрації незгладженої оцінки спектру, для чого було застосовано рекурсивний фільтр першого порядку. Це двоетапна процедура, на першому етапі якої виконується згладжування прямокутним вікном зліва направо:

$$y_n = y_{n-1} + \frac{1}{N}(x_n - x_{n-N}), \quad n = N+1, \dots, N_{fft}/2+1,$$

а на другому етапі виконується згладжування прямокутним вікном справа наліво:

$$y_n = y_{n+1} + \frac{1}{N}(x_n - x_{n+N}), \quad n = N_{fft}/2 - N + 1, \dots, 1,$$

де $N = \frac{N_{fft}}{F_s} \cdot \frac{df}{2}$, $N_{fft} = 65536$, $F_s = 44100$ Гц, $df = 50$ Гц.

Як бачимо, значення параметру N обчислювалося, виходячи із ширини прямокутного вікна 25 Гц. Це зумовлене тим, що подвійне згладжування прямокутним вікном еквівалентне одиночному згладжуванню трикутним вікном подвоєної ширини. Необхідність в застосуванні двох напрямків фільтрації зумовлена тим, що при цьому не треба дописувати якісь вибірки поза краями первинної оцінки спектру, що підлягає згладжуванню.

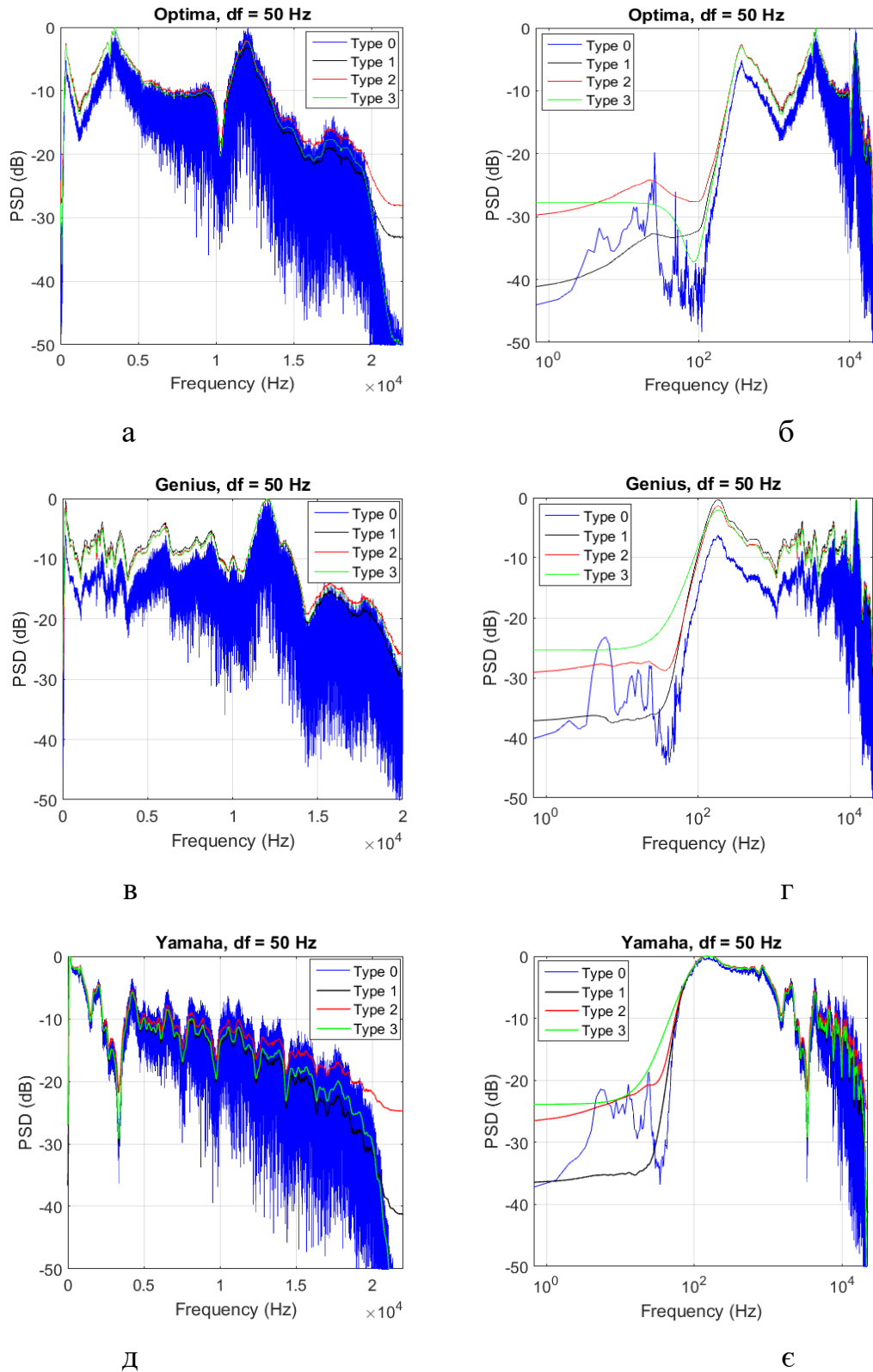


Рис. 2.9 – Оцінки спектру сигналу гучномовців (заглушене приміщення)

Порівнюючи між собою наведені на рис. 2.8 оцінки $|H_{lm}(f)|$, доходимо наступних висновків.

По-перше, виходячи із ступеня фактичної нерівномірності та фактичному частотному діапазону оцінки АЧХ $|H_{lm}(f)|$, слід віддати перевагу активній колонці Genius. Як це не дивно, але акустичний монітор Yamaha виявився неконкурентним, що можна пояснити його фізичним старінням.

По-друге, як видно із наведених графіків, незгладжену оцінку $|H_{lm}(f)|$ нульового типу недоцільно використовувати для подальшої корекції, з огляду на неприйнятно велику дисперсію в області середніх та, що особливо, високих частот.

По-третє, серед згладжених оцінок найбільш привабливою виглядає оцінка третього типу як така, що найкращим чином забезпечує зниження дисперсії оцінки без надмірного підвищення її зміщення. Зауважимо, що певна доля суб'єктивності при виборі типу оцінки є неминучою, оскільки істинна АЧХ підсистеми ГМ невідома.

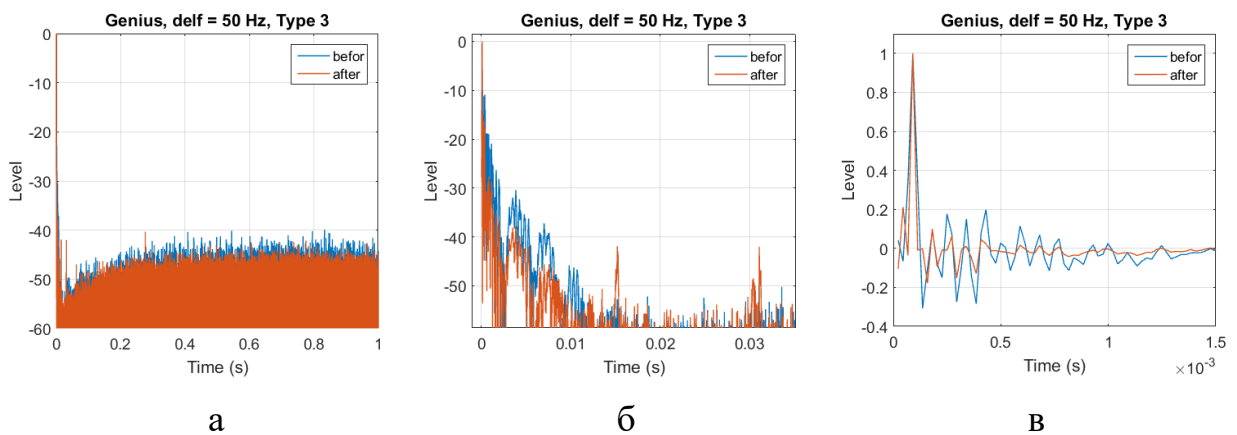


Рис. 2.10 – Скориговані оцінки ІХ для гучномовця Genius

Результати обчислень ІХ підсистеми ГМ згідно рівняння (2.4) із використанням згладженої оцінки $|H_{lm}(f)|$ для гучномовця Genius показані на рис. 2.10 для різних значень часового масштабу.

Результат коригування, як бачимо, має як позитивні, так і негативні якості. Негативною є поява окремих дискретних сплесків достатньо високого рівня -40 дБ (рис. 2.10а). Позитивні: 1) рівень бокових пелюсток знизився приблизно на

3-4 дБ (рис. 2.8а); 2) головний пік ІХ став вужчим в 2 рази, а рівень найближчих до нього сплесків став нижчим майже на 6 дБ (рис. 2.8б,в).

З огляду на гостроту зазначених вище сплесків доречно припустити, що вони спричинені збільшенням дисперсії оцінки АЧХ, отриманої із використанням виразу (2.14), на високих частотах, через ділення на малі значення згладженої оцінки $|H_{lm}(f)|$. Очевидним та простим способом пригнічення таких сплесків є множення відкоригованої оцінки АЧХ на вагове вікно. В часовій області це відповідає процедурі згладжування оцінки ІХ шляхом ковзного усереднення.

В рамках даної роботи розглянуто два види такого вагового вікна. В першому випадку це вікно описується виразом:

$$w_1(r) = \begin{cases} 1, & r = 1, \dots, \frac{N_{ff}}{2} - k, \quad r = \frac{N_{ff}}{2} + 2 + k, \dots, N_{ff}, \\ 0,5 \left[1 - \cos \frac{\pi}{k} \left(r - \frac{N_{ff}}{2} - 1 \right) \right], & r = \frac{N_{ff}}{2} + 1 - k, \dots, \frac{N_{ff}}{2} + 1 + k, \end{cases}$$

де k визначає інтервал, що відповідає половині періоду косинусоїди, що описує кінцеву частину вікна, яка відповідає за пригнічення високих частот:

$$k = (F_s/2 - f_1) / (F_s/N_{ff}).$$

Друге вагове вікно – це вікно Ханна

$$w_2(r) = 0,5 \left[1 - \cos \left(\pi(r - N_{ff}/2 - 1) / \Delta N \right) \right], \quad r = 1, \dots, N_{ff},$$

де ΔN – половина ширини вікна Ханна у вибірках.

Графіки вагових вікон для $f_1 = 19$ кГц та $f_2 = 18$ кГц ($f_2 = \Delta N \cdot F_s / N_{fft}$) наведено на рис. 2.11. Неважко помітити, що в окремому випадку $f_1 = 0$, $f_2 = F_s/2$ ці вікна співпадають: $w_1(r) = w_2(r)$.

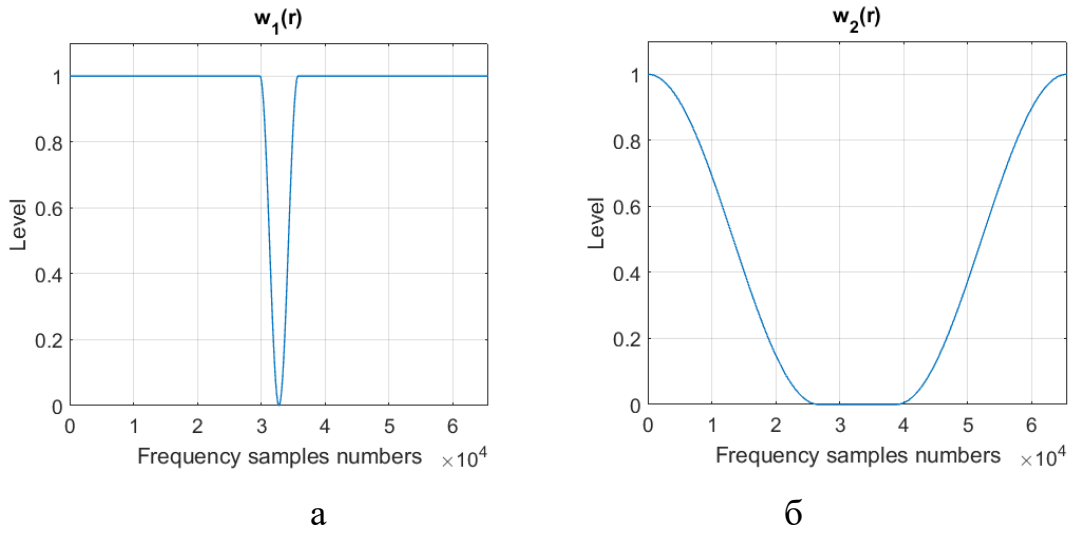


Рис. 2.11 – Вагове вікно $w_1(r)$ для $f_1=19$ кГц (а) та вікно $w_2(r)$ для $f_2=18$ кГц (б)

Результати застосування вікон $w_1(r)$ та $w_2(r)$ наведено на рис. 2.12 та 2.13.

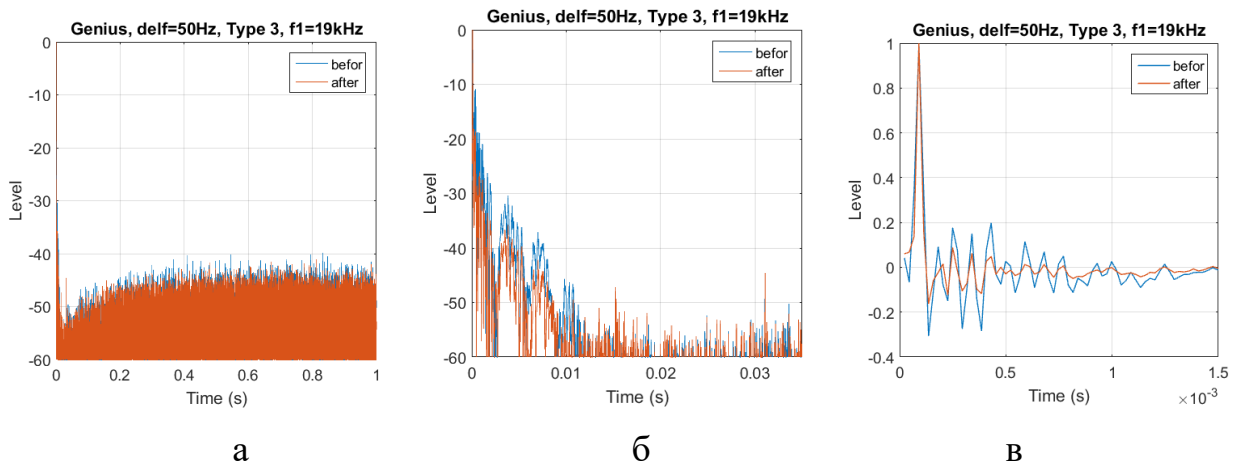


Рис. 2.11 – Скориговані ІХ для гучномовця Genius, із вікном $w_1(r)$ для $f_1=19$ кГц

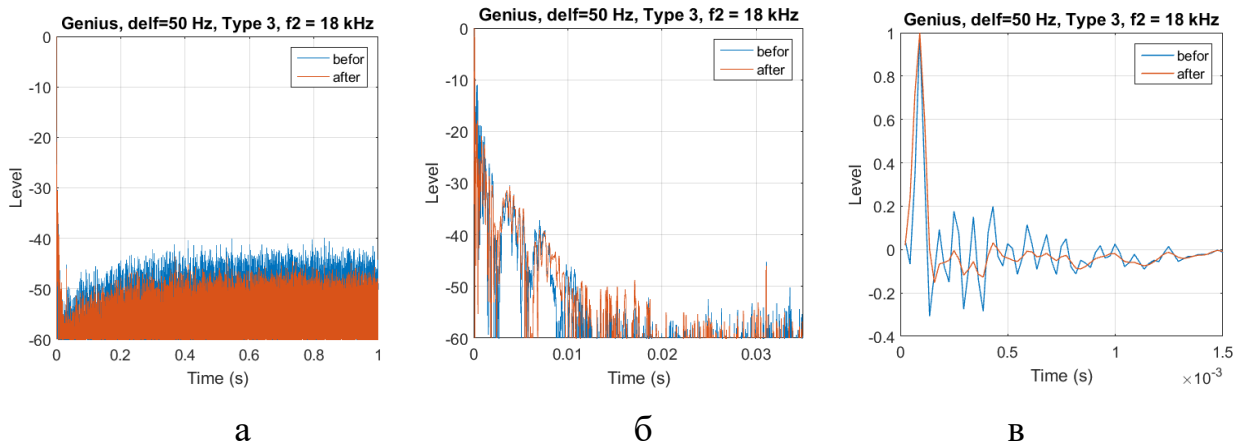


Рис. 2.12 – Скориговані ІХ для гучномовця Genius, із вікном $w_2(r)$ для $f_2=18$ кГц

Порівнюючи рис. 2.11 та 2.12, неважко зробити висновок, що вікно $w_2(r)$ із параметром $f_2=18$ кГц є більш перспективним, оскільки дозволяє зменшити рівень бокових пелюсток оцінки ІХ підсистеми ГМ на 4 дБ, тоді як вікно $w_1(r)$ із параметром $f_1=19$ кГц дозволяє зменшити рівень бокових пелюсток лише на 1 дБ (рис. 2.11а та 2.12а). Правда, в інтервалі 0-10 мс вікно $w_1(r)$ є кращим, оскільки забезпечує рівень бокових пелюсток на 6 дБ нижчий, ніж вікно $w_2(r)$. Проте ця перевага не є принциповою, оскільки у випадку вікна $w_2(r)$ рівень цих бокових пелюсток не перевищує -30 дБ, а ширина головного сплеску ІХ на рівні -30 дБ не перевищує 2 мс, що забезпечує достатню для інженерних застосувань точність оцінювання ІХ приміщення.

Зазначимо, що використання вагового вікна $w_2(r)$ можна розглядати як регуляризацію, із регуляризуючим множником $w_2(r)$ та параметром регуляризації f_2 , рішення нестійкої оберненої задачі [81].

Оскільки в даній роботі розглянуто варіант коригування частотної характеристики тракту ШГ, де враховується лише нерівномірність АЧХ підсистеми ГМ, в подальшому слід розглянути доцільність додаткового врахування оцінки фазової характеристики підсистеми ГМ.

Висновки до розділу 2

1. Вдосконалено алгоритм автоматичного оцінювання часу реверберації прямим методом, що полягає у автоматичному віконному з перекриттям аналізі ІХ приміщення. Працездатність та точність оцінювання часу реверберації за допомогою вдосконаленого алгоритму були експериментально перевірені при оцінюванні акустичних параметрів учбових приміщень університету.

2. Розроблено рекомендації і алгоритми коригування характеристик акустичної апаратури непрофесійного рівня якості при проведенні оцінювання якості та розбірливості мовлення із використанням штучної голови. Працездатність та точність розроблених алгоритмів були експериментально перевірені при дослідженні учбових приміщень.

3. Отримані результати роблять можливими виконання наступних досліджень:

- створення бази даних бінауральних імпульсних характеристик приміщень;
- дослідження алгоритмів дереверберації двоканальних записів мовленнєвих сигналів із використанням бази бінауральних ІХ;
- проведення досліджень розбірливості мовлення в різних точках приміщення із використанням відповідних записів бінауральних ІХ приміщення.

3. МОДЕЛЮВАННЯ АЛГОРИТМУ СЛПОВОГО ВИМІРЮВАННЯ ЧАСУ РЕВЕРБЕРАЦІЇ І ПРИДУШЕННЯ РЕВЕРБЕРАЦІЇ

3.1. Комп'ютерне моделювання алгоритму вимірювання часу реверберації методом максимальної правдоподібності

Алгоритм з вимірювання ЧР методом МП вже реалізовано, тому доцільно будувати експерименти і покращення на базі цього програмного коду. Цей пакет є у вільному доступі [82], його експлуатація здійснюється в програмному середовищі Matlab. Не дивлячись на те, що пакет був розроблений на основі [70], він не є тим самим програмним інструментом, який використовувався під час написання вищезгаданої роботи: алгоритм дещо спрощено, не будується друга гістограма для відслідковування зміни ЧР в часі, значення параметрів не ідентичні до тих, що використовувались в роботі [70]. Окрім цього, пакет має деякі обмеження:

- 1) ЧР вимірюється в межах $0,2 \dots 1,2$;
- 2) вважається, що відстань між джерелом і приймачем більше критичної;
- 3) пакет скриптів призначений для роботи з сигналами, які попередньо пройшли процедуру придушення шуму.

Схема алгоритму оцінки ЧР методом МП, який був реалізований в пакеті скриптів [82] представлена на рис. 3.1. Алгоритм більш детально описаний в роботі [69] і в першому розділі цієї роботи.

Детальне вивчення програмного коду [82] показало, що алгоритм не зовсім такий, як було описано у [70]. Так, ваговий фактор у виразі (2.1) рівний одиниці для кожної з умов. Було сказано, що: «Якщо жоден підфрейм всередині фрейму не задовольняє (2.1), аналізується наступний фрейм». Насправді ж, коли аналізуються початкові підфрейми у фреймі, і якщо кількість підфреймів, що задовольняють умові (2.1) менше за деяке мінімальне значення, аналіз підфреймів у фреймі далі не виконується. Тобто, навіть якщо для перших двох

фреймів не виконується (2.1), інші підфрейми не розглядаються, а фрейм зсувається на півфрейма.

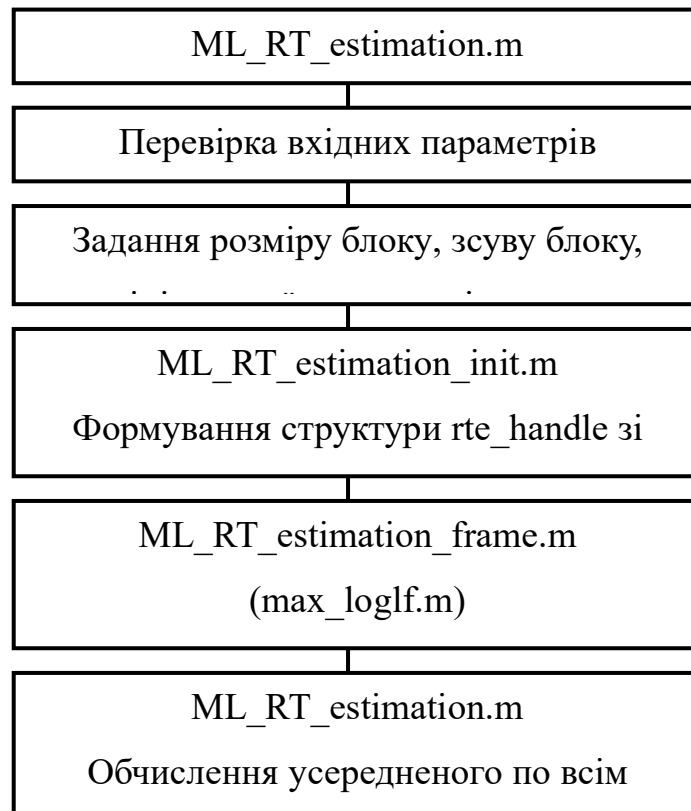


Рис. 3.1 – Схема алгоритму оцінювання ЧР методом МП.

В [6], [81] для дослідження похибки вимірювання ЧР всліпу використовувались синтезовані ІХ (спочатку формувалась їх огибаюча, яка потім заповнювалась білим гаусівським шумом), але у цій роботі були використані ІХ з бази даних ІХ [83], а також ІХ, що були виміряні під час дослідження [3]. Зразки мови були взяті з бази даних ТІМІТ [84], а саме записи 20 дикторів (10 жінок та 10 чоловіків), які промовляють різні речення. Для отримання реверберованого сигналу чистий сигнал згортався з ІХ відповідного приміщення. Розмір фрейму 320 мс, зсув фрейму 160 мс.

Час реверберації був виміряний двома способами: методом Шредера (таким чином були отримані еталонні значення ЧР) і методом МП по реверберованому мовленнєвому сигналу.

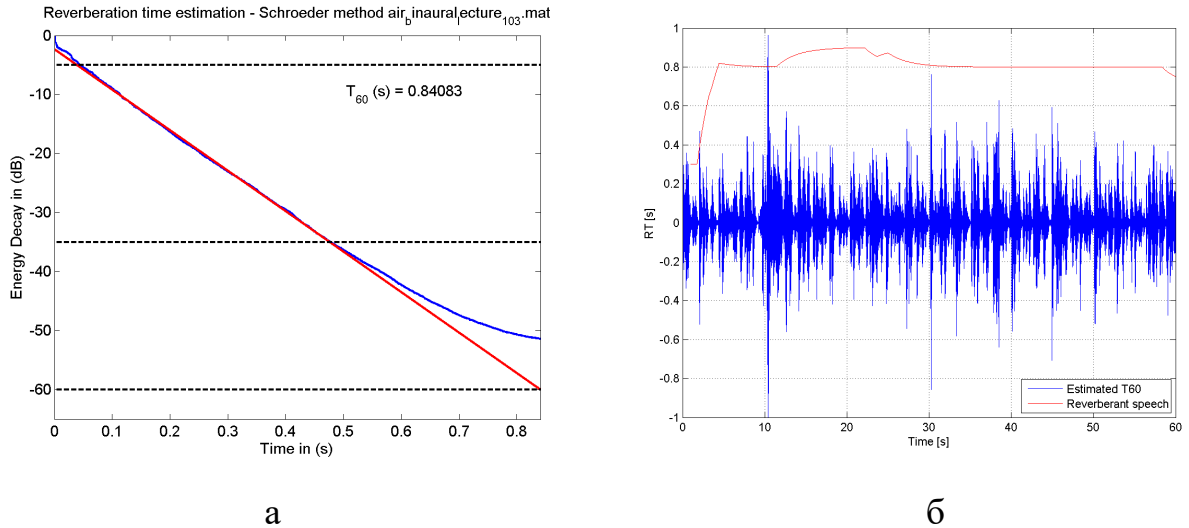


Рис. 3.2 – Приклад вимірювання ЧР (а) методом Шредера і (б) методом МП. ІХ лекційної кімнати.

З рис. 3.2. видно, що методом МП було досить точно виміряно ЧР відносно метода Шредера, однак така ситуація не для всіх ІХ.

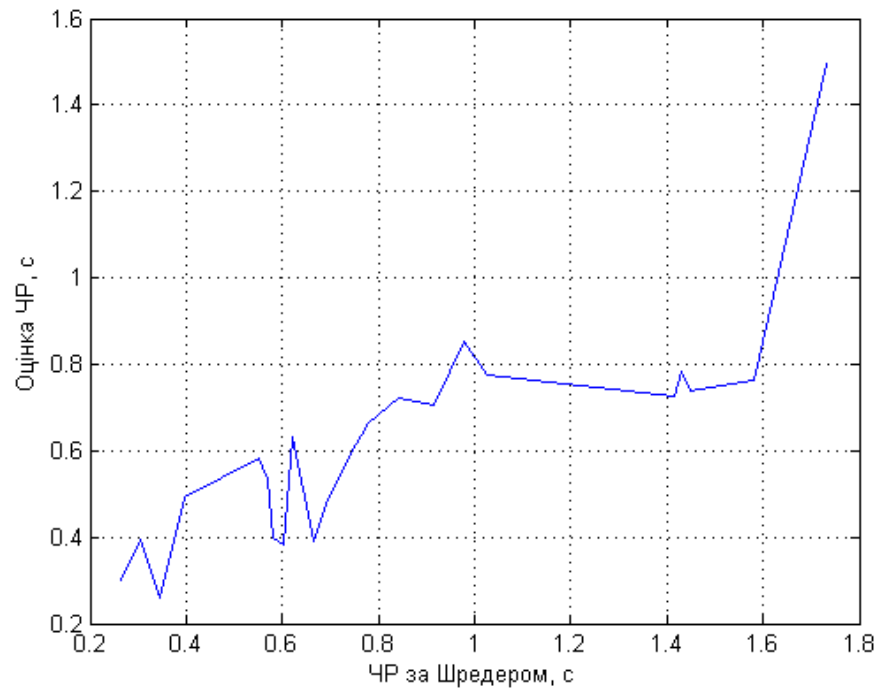


Рис. 3.3 – Графік залежності оцінки ЧР методом МП від дійсного значення ЧР (метод Шредера), довжина підфрейма стандартна (0,032 с).

В ідеальному випадку залежність оцінки ЧР методом МП від дійсного значення ЧР має бути прямою, однак, як видно з рис. 3.3. це твердження не справджується. Вірогідно це викликано тим, що алгоритм не досить точно обирає ділянки для вимірювання ЧР.

3.2. Оптимізація алгоритму вимірювання часу реверберації методом максимальної правдоподібності

Проведемо експеримент і побудуємо графіки відповідності ЧР виміряного методом МП дійсному ЧР. Графіки наведено у додатку Б. Можна побачити, що найкраща відповідність між ЧР виміряним методом МП та дійсним ЧР досягається при довжині підфрейму 0,07-0,085 с.

Побудуємо графік залежності абсолютної похибки вимірювання від часу реверберації для цих довжин підфреймів (додаток В). Для довжини підфрейма 0,07 с для ЧР 1,4 с похибка складає 0,12с, що є досить непоганим результатом, однак для ЧР від 0,2 до 0,55 с похибка становить від 0,1 до 1,5 с. Найменша похибка вимірювання у діапазоні ЧР від 0,2 до 1 с досягається при довжині фрейму 0,08 с.

У першому розділі було показано, що алгоритм, покладений в основі [82], можна вдосконалити, будуючи другу гістограму, буфер якої значно менший з буфер першої. Графіки залежності абсолютної похибки вимірювання від часу реверберації для довжин підфреймів 0,07 с, 0,075 с, 0,08 с, 0,085 с, показано у додатку В (рис. В.2).

Покращений алгоритм дає найкращі результати мінімізації максимуму похибки для довжини підфрейму 0,75 с. В діапазоні ЧР 1-1,8 с максимальна похибка складає 0,16 с, однак всередньому приблизно становить 0,07 с. Для довжини підфрейму 0,085 с в діапазоні ЧР 1-1,8 с максимальна похибка складає 0,2 с, що більше за максимальну похибку для довжини підфрейму 0,075 с. Однак всередньому в діапазоні ЧР 1-1,8 с похибка приблизно становить 0,077 с.

Можна помітити, що оригінальний алгоритм з одною гістограмою забезпечує найменшу похибку вимірювання ЧР в діапазоні 0,2-1 с, а покращений алгоритм з двома гістограмами – в діапазоні 1-1,8 с. Тож скомбінуємо ці два алгоритми. Якщо для фрейму оцінка ЧР становить менше 1 с, використовуємо алгоритм з однією гістограмою. Якщо ж для фрейму оцінка ЧР становить більше 1 с, використовуємо алгоритм з двома гістограмами.

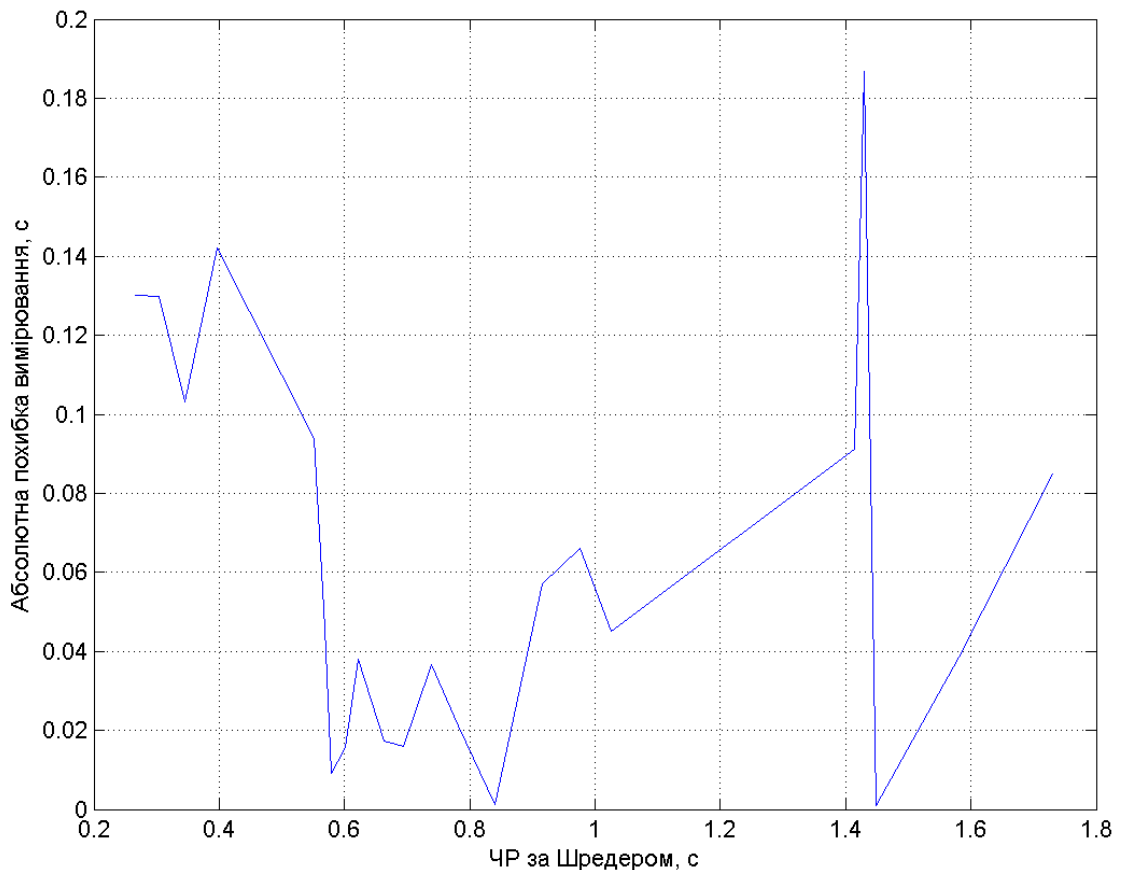


Рис. 3.4 – Графік залежності абсолютної похибки вимірювання від часу реверберації. Комбінований алгоритм.

Як видно з рис. 3.4, комбінований алгоритм вимірювання часу реверберації методом МП забезпечує похибку вимірювання на всьому діапазоні ЧР до 0,19 с, при цьому середня похибка вимірювання становить 0,0623 с, що є дуже хорошим результатом.

3.3. Оптимізація параметрів алгоритму придушення реверберації

Як було показано у розділі 1.5, для придушення реверберації методом спектрального віднімання, необхідно знати спектр реверберації. На жаль, в роботах [46], [51] відсутні вказівки щодо раціонального вибору конкретних значень параметрів T_l та $\eta_z^d(k)$. З метою оптимізації цих параметрів було проведено ряд досліджень.

Для експериментальних досліджень використовувався тестовий мовленнєвий сигнал тривалістю 32 с, оцифрований з частотою дискретизації 22050 Гц і з рівномірним квантуванням 16 біт. Дія реверберації моделювалась відповідно до співвідношення (1.3), тобто шляхом згортки вихідного чистого мовленнєвого сигналу з ІХ приміщення (використовувалися ІХ трьох приміщень з часом реверберації = 0,74; 0,89 і 1,1 с).

Зіставлення результатів оцінювання PESQ для спотворених реверберацією сигналів, а також для сигналів, відкоригованих з використанням алгоритму logMMSE [85], показало, що придушення пізньої реверберації дозволяє на 2-10% підвищити якість мовленнєвого сигналу.

Результати перевірки справедливості співвідношення (1.11), наведені на рис. 3.5, свідчать, що оптимальним, в сенсі максимуму PESQ, є значення $\eta_z^d \approx 0,5$. При цьому оптимальне значення параметра T_l укладено в інтервалі 15, ..., 30 мс.

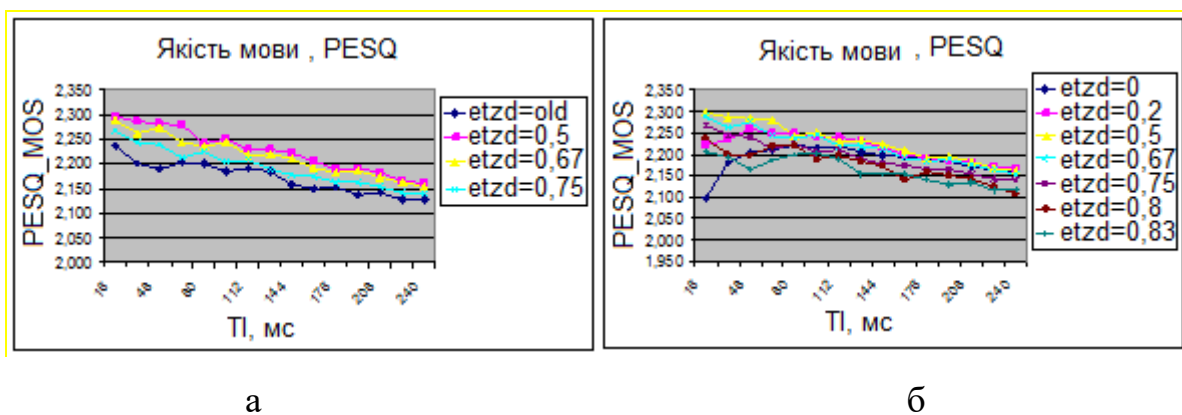


Рис. 3.5 – Оцінки PESQ для η_z^d за формулою (1.11) ($\text{etzd} = \text{old}$), і для

$$\eta_z^d = 0, 1, \dots, 0, 9$$

Разом з тим, дослідження показали, що значення показника PESQ практично не залежать від вибору параметра η_z^a (рис. 3.6).

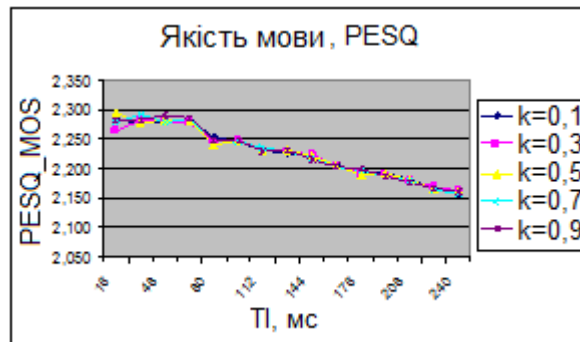


Рис. 3.6 – Значення PESQ для $\eta_z^d = 0,5$ і для $\eta_z^a = k \eta_z^d$, $k = 0,1, \dots, 0,9$

Тим самим показано, що замість співвідношень (1.9)-(1.11) доцільно використовувати просте і разом з тим більш ефективно в плані максимізації показника PESQ співвідношення:

$$\hat{\lambda}_y(l, k) = \eta_y \cdot \hat{\lambda}_y(l-1, k) + (1 - \eta_y) \cdot |Y(l, k)|^2, \quad \eta_y \approx 0,5. \quad (3.1)$$

Результати досліджень опубліковані в роботі [86].

Висновки до розділу 3

1. Представлено результати комп'ютерного моделювання алгоритму сліпого вимірювання часу реверберації, що дозволило оптимізувати параметри алгоритму.
2. Показано, що при довжині підфрейму 0,07-0,085 с забезпечується найменша похибка оцінювання ЧР.
3. Показано принципову можливість розширення діапазону вимірювання ЧР за реверберованим мовленнєвим сигналом з 0,2..1,2 с до 0,2..1,8с.
4. Експериментальні дослідження показали принципову можливість отримати середню похибку вимірювання ЧР 0,06с сліпим методом (відносно оцінки прямим методом) на всьому діапазоні вимірювання ЧР 0,2..1,8с.
5. Показано можливість подальшої оптимізації параметрів алгоритму придушення реверберації при використанні оцінок ЧР, отриманих сліпим методом.

4. АВТОМАТИЗАЦІЯ ОБ'ЄКТИВНОГО І СУБ'ЄКТИВНОГО ОЦІНЮВАННЯ ЯКОСТІ І РОЗБІРЛИВОСТІ МОВЛЕННЄВОГО СИГНАЛУ

4.1. Автоматизація артикуляційних випробувань для суб'єктивного оцінювання розбірливості мовлення

Експериментальні дослідження, проведені в рамках освітнього науково-дослідного проекту, впливу шуму та реверберації на розбірливість мовлення впроваджувалися протягом 2017-2020 років на кафедрі акустичних та мультимедійних систем Національного технічного університету "Київський політехнічний інститут імені Ігоря Сікорського". Для цих досліджень був розроблений спеціальний інструментарій програмного забезпечення. Цей інструментарій дозволив, по-перше, автоматизувати процедуру перевірки артикуляції, а по-друге, дозволив студентам проводити такі тести вдома, тобто в максимально комфортних умовах, за допомогою персональних комп'ютерів. За матеріалами дослідження була опублікована стаття [6] за співавторством дисертанта.

Під час реалізації досліджень було виявлено, що розробка та тестування відповідних програмних засобів неминуче супроводжуються різними неточностями та помилками. Одним із прикладів таких неточностей є недостатньо добре врахування особливостей української фонетики при створенні наборів перевірочних слів [29]. Інший приклад – недосконалість розробленого програмного забезпечення. Крім того, на результати експериментів можуть суттєво вплинути недоліки організації експериментальних досліджень, такі як, наприклад, недостатня підготовка учасників експерименту або, навпаки, надмірне інформування учасників експерименту під час його проведення [87].

Цілком природно очікувати, що режими прослуховування також впливають на результати експерименту. Оцінювачі розбірливості мовлення для викладу діотичної мови (один і той же сигнал направляється в різні гучномовці)

через пару акустичних моніторів та через навушники були наведені в [88]. Діхотичне прослуховування (різні стимули спрямовані в різні вуха) через навушники можна розглядати як бінауральний режим з усіма його перевагами порівняно з монофонічним та діотичним режимами. Таким чином, суб'єктивна оцінка зрозумілості мовлення для діхотичного слухання становить особливий інтерес, і ці результати також будуть представлені нижче.

Підготовка текстових та звукових таблиць. Таблиці з односкладними звукосполученнями стандарту [89], що передбачає можливість автоматизації суб'єктивного тестування артикуляції, були взяті за прототип для розробки текстових таблиць. Ці таблиці були адаптовані до української відповідно до особливостей українського правопису та вимови. Важливою особливістю стандарту [89] є використання неповних приголосний-голосний-приголосний односкладових звукосполучень. У цьому випадку досліджувані повинні додати пропущену букву в складі за допомогою клавіатури комп'ютера. Значне скорочення часу навчання суб'єктів – головна перевага такого підходу. Однак цей підхід викликає сумніви щодо достовірності результатів через його значну відмінність від "класичного" [90] тесту на артикуляцію.

Три опорні версії кожної текстової таблиці були створені шляхом копіювання та редагування першої (основної) версії таблиці. Таким чином, була врахована можливість подвійної інтерпретації окремих звуків, таких як /г/ та /х/, /з/ та /с/ тощо. Більш того, таким чином враховувалась також варіативність введення сприйнятих односкладових клавіатур з клавіатури. Наприклад, склад «м'яр» можна ввести з клавіатури двома способами, як «мйар» або як «м''яр». Перший варіант простіший, оскільки дозволяє комусь обійти проблему введення апострофу в розкладці української клавіатури. Друга версія виглядає трохи дивно, але вона відповідає вимогам синтаксису MATLAB.

Звукові таблиці були створені шляхом запису голосу диктора в заглушеній кімнаті. Для запису використовували мікрофон Superlux ECM 999, зовнішню USB звукову карту PRESONUS AudioBox та аудіоредактор і диктофон Audacity

версії 2.1.3. Запис проводився з частотою дискретизації 44100 Гц і точністю 16 біт. Склади були вбудовані у речення «Запишіть ___ тепер» і їх промовляли 7 чоловіків і 2 жінки. Наприклад, склад бон читався як «Запишіть бон тепер».

Моделі сигналів. При моделюванні спотвореної мови враховувалася можливість спотворення звуку через шум та реверберацію:

$$y(t) = x(t) \otimes h(t) + n(t), \quad (4.1)$$

де $x(t)$ – чистий мовленнєвий сигнал, $h(t)$ – ІХ кімнати, $n(t)$ – шум, \otimes – символ згортки.

Коли впливом реверберації можна знехтувати, шум додається до чистого сигналу $x(t)$ з ваговим коефіцієнтом k :

$$y(t) = x(t) + k \cdot n(t), \quad (4.2)$$

де $k = 10^{0.05(SNR - SNR_0)}$, $SNR = 10 \lg(D_x/D_n)$ – "початкове" співвідношення сигнал/шум для чистого сигналу $x(t)$ з "початковою" дисперсією D_x і шуму $n(t)$ з "початковою" дисперсією D_n , SNR_0 – це бажане значення SNR для суміші (4.2).

Зауважте, що зони існування мовленнєвого сигналу необхідно визначити перед розрахунком SNR. Детектор голосової активності є інструментом, який зазвичай використовується для вирішення цієї проблеми, і для тестування програмного забезпечення був розроблений простий детектор голосової активності.

Якщо не можна нехтувати ефектом реверберації, слід зазначити, що раннє відбиття може значно посилити прямий звук [22], і значення SNR можна розрахувати з урахуванням цієї обставини. Дійсно, ІХ $h(t)$ можна представити у вигляді суми початкових та хвостових відрізків $h_e(t)$ і $h_l(t)$ відповідно:

$$h(t) = h_e(t) + h_l(t), \quad (4.3)$$

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

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

З огляду на це, «початкове» SNR може бути представлено як $SNR = 10 \lg D_{x_e}/D_n$, де D_{x_e} – дисперсія посиленого сигналу $x_e(t) = x(t) \otimes h_e(t)$. Потім

параметри SNR_0 та k можуть бути обчислені відповідно до (4.2).

Комп'ютерна програма. Блок-схема розробленої програми управління показана на рис. 4.1. Програма управління починається із запиту на вибір режиму роботи або навчання. Програма також запитує прізвище суб'єкта та номери односкладних таблиць. Після підготовки папок, де будуть збережені результати оцінки розбірливості мовлення, запускається спеціальна програма під назвою «помічник» і забезпечує бездоганне продовження роботи у разі припинення робочого сеансу. Зауважимо, що таке періодичне припинення роботи необхідно, щоб уникнути надмірної втоми слухачів.

Після завантаження звукової таблиці чиста мова спотворюється відповідно до (4.1) і представляється слухачеві, який записує сприйнятий склад за допомогою клавіатури комп'ютера. Коли процес прослуховування звукової таблиці закінчується, проводиться розрахунок розбірливості мови та результати обчислення записуються у спеціальний файл, який розміщується у попередньо підготовлених папках на диску. Склад вважається правильно розпізнаним, якщо його символічний запис повністю збігається принаймні з одним із трьох опорних текстових варіантів артикуляційних таблиць.

Формування спотвореного сигналу забезпечується за допомогою чотирьох програм: 1) моделювання детектора голосової активності для подальшого правильного розрахунку бажаного значення SNR, 2) створення паралельної комбінації смугових октавних фільтрів, 3) використання цих фільтрів для отримання білого шуму, кольорового шуму зі ступінчастими спектрами (швидкість падіння спектральної щільності 3 дБ/октава та 6 дБ/октава для рожевого та коричневого шуму відповідно) у смузі частот від 90 Гц до 11 кГц, 4) спотворення мовленнєвих сигналів у відповідно до (4.1)-(4.2).

Обробка результатів тестування окремих слухачів та усереднення цих результатів для всіх слухачів виконується за допомогою ще двох комп'ютерних програм.

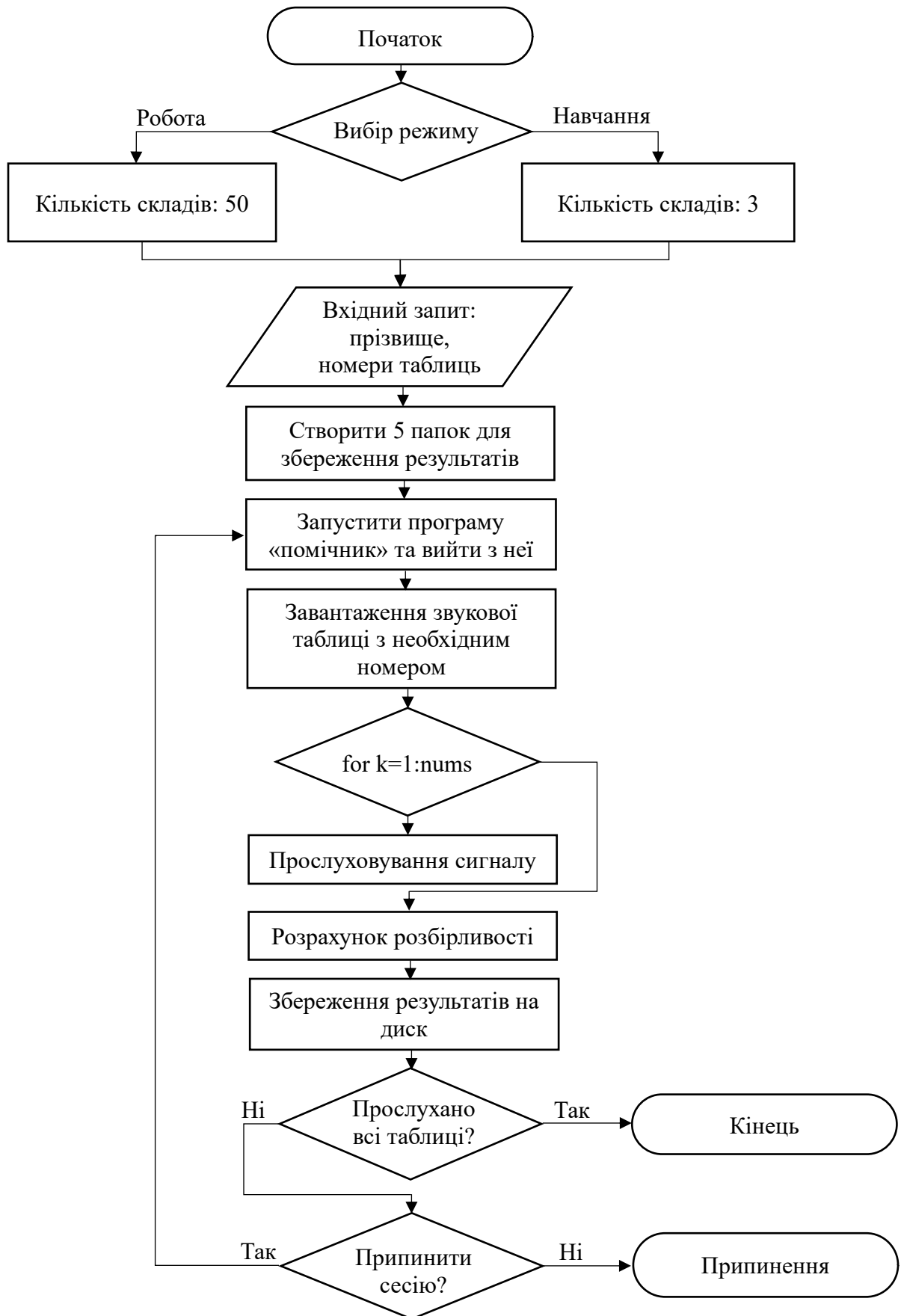


Рис. 4.1 – Структурна схема програми управління

Таким чином, розроблений набір програмних засобів дозволяє вирішити дві важливі завдання: 1) створення текстових та звукових таблиць, які є основою артикуляційних випробувань, 2) проведення артикуляційних випробувань.

4.2. Порівняння об'єктивних показників якості мови з суб'єктивною оцінкою розбірливості

При оцінці розбірливості та якості мовлення використовуються різні кількісні показники, оскільки якість мови та її розбірливість не тотожні, хоча й тісно пов'язані між собою поняття. Тим часом у ряді робіт представлені результати досліджень, спрямованих на оцінку розбірливості мовлення за допомогою показників якості мовлення [91], [92], [93]. Спроби використання показників, розроблених для іноземних мов, зазвичай пояснюються низькою ефективністю або недостатньою достовірністю "рідних" показників. Прикладом такої спроби є пропозиція використати показник якості PESQ для оцінки розбірливості мови в каналах зв'язку з низькошвидкісними кодеками [92].

Результати оцінки розбірливості мовлення у вихідному сигналі слухового апарату за наявності фонових завад у формі розмови між двома людьми представлені в [91]. Тестовими сигналами були записи 100 речень, тобто оцінювалася так звана фразова розбірливість. Об'єктивну оцінку якості мовлення проводили за допомогою широкосмугової версії вимірювання PESQ (WB-PESQ). Значення коефіцієнта кореляції між балами PESQ та суб'єктивними оцінками розбірливості дуже високі (0,91-0,99).

Розбірливість мовлення оцінювали також за допомогою техніки артикуляційного тесту на спотворення, властиві вокодеру [92]. У цьому випадку мовленнєві сигнали записувались як 50 звукових сполучень приголосний-голосний-приголосний (CVC), що вимовляються всередині речення. Оцінки розбірливості мовлення були у вигляді відносної кількості правильно зрозумілих звукових сполучень, а об'єктивну оцінку якості мовлення проводили за допомогою вимірювання PESQ. Значення коефіцієнта кореляції між PESQ та

суб'єктивними оцінками розбірливості виявилось дуже високим і близько 0,86-0,95.

Очевидно, що дослідження інших, більш простих для розрахунків, показників якості мовлення також становили б значний інтерес. На жаль, у [91] та [92] такі альтернативні заходи не розглядалися.

Ця вада була усунена в [93], де були представлені результати дослідження дуже великого переліку об'єктивних показників якості мовлення [94] на предмет їх придатності для ролі показників розбірливості мовлення. Додатково до вузькосмугового показника PESQ (NB-PESQ) були розглянуті SSNR, fwSNR, логарифмічний коефіцієнт правдоподібності (LLR), відстань Ітакура-Сайто (IS), відстань кепструма (CEP), зважена спектральна відстань (WSS). При оцінці розбірливості приголосних звуків та розбірливості речень, лише показники PESQ ($r = 0.77 \dots 0.79$) та fwSNR ($r = 0.68 \dots 0.81$) були названі помірно ефективними. Показники CEP, LLR та SSNR були не дуже ефективними, а параметри IS та WSS – найменш ефективними.

На жаль, легкі для розрахунку та добре зрозумілі показники LSD [95], [96] та BSD [97] не розглядалися в [93]. Недоліком також є відсутність результатів для "стандартного" білого, рожевого та коричневого шуму, що ускладнює порівняння різних об'єктивних показників. Нарешті, вкрай дискусійно використовувати коефіцієнт кореляції як міру надійності об'єктивних показників якості.

Суб'єктивна оцінка розбірливості мови, спотвореної шумом та реверберацією, важлива для поліпшення імітації слухової системи людини [16], [17], калібрування систем оцінки об'єктивної розбірливості мови [18], розробки та сертифікації пристроїв та каналів зв'язку, аудиторій та конференц-залів, слухових апаратів та кохлеарних імплантатів [18], [19], [20].

Метою роботи [2] є порівняння оцінки якості та розбірливості мовленнєвих сигналів, спотворених білим, рожевим та коричневим шумом різної інтенсивності. Оцінка розбірливості мовлення здійснюється за

допомогою техніки вимірювання артикуляції складів. Оцінка якості мовлення реалізується як суб'єктивно, так і об'єктивно. В останньому випадку використовуються показники SSNR, LSD, fwSNR, BSD та WB-PESQ.

20 слухачів у 22-річному віці виконали суб'єктивну оцінку розбірливості односкладних звукосполучень типу приголосний-голосний-приголосний, спотворених білим, рожевим та коричневим шумом. Оцінка була здійснена за допомогою спеціальної комп'ютерної програми [29]. Кожен слухач прослухав через навушники 3 таблиці, кожна з яких містила 50 складів. Склади програвались у випадковому порядку. Середні оцінки розбірливості наведені на рис. 4.2а, а суб'єктивні оцінки якості мовлення для безперервної мови (4 чоловіки та 4 жінки), спотворені білим, рожевим та коричневим шумом, наведені на рис. 4.2б [98].

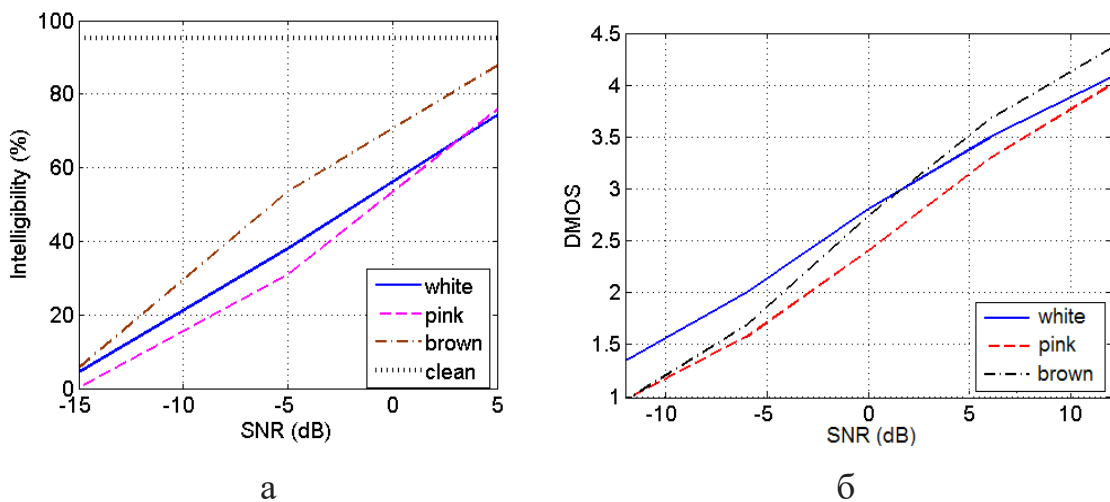


Рис. 4.2 – Середні оцінки розбірливості мовлення (а) та якості мови [98] (б)

Шкала деградації середньої суб'єктивної оцінки (DMOS) [97] була використана під час тестування якості мовлення, проведеного 10 слухачами середнього віку 22 років. Записи спотворених сигналів слухачам програвали в довільному порядку.

Як видно з рис. 4.2а, розбірливість мови найнижча для рожевого шуму, хоча існує тенденція наближення результатів білого до рожевого шумів для $SNR > 5$ дБ. Водночас розбірливість мови виявилася найвищою для коричневого

шуму.

Порівнюючи рис. 4.2а та 4.2б, можна зробити висновок, що поведінка функцій є подібною у випадку середніх та великих значень SNR ($\text{SNR} > 0$ дБ). Ця подібність втрачається у разі малих значень SNR ($\text{SNR} < 0$ дБ).

Порівняння суб'єктивних оцінок якості та розбірливості мовлення, спотвореного білим, рожевим та коричневим шумом, свідчить про те, що результати узгоджені для невеликого діапазону значень SNR, а саме для $0 < \text{SNR} < 5$ дБ. Тому при вивченні впливу умов прослуховування на розбірливість мови в широкому діапазоні значень SNR недоцільно використовувати якість мовлення як міру розбірливості мови.

Об'єктивну оцінку якості мовлення проводили за тим самим набором з 8 мовленнєвих сигналів (4 чоловічих та 4 жіночих), як і у випадку суб'єктивної оцінки якості мовлення [29]. Аналіз проводили для сегментів сигналу тривалістю 15 с, використовуючи 5 об'єктивних показників якості.

Різниця в часовій області (тобто форма хвилі) між чистим і погіршеним сигналами використовується для вимірювання SSNR:

$$SSNR = \frac{1}{M} \sum_{m=1}^M 10 \lg \left[\frac{\sum_{n=R(m-1)+1}^{Rm} x^2(n, m)}{\sum_{n=R(m-1)+1}^{Rm} [x(n, m) - y(n, m)]^2} \right], \quad (4.4)$$

де $x(n, m)$ та $y(n, m)$ – n -ний відлік m -того фрейму чистого сигналу $x(n)$ та спотвореного сигналу $y(n)$ відповідно; M – кількість фреймів, R – кількість відліків у фреймі. При обчисленні SSNR прийнято враховувати лише ті фрейми, для яких значення SNR знаходиться в інтервалі $[-10, +35]$ дБ [97].

Показник LSD відноситься до групи показників, що розраховуються у частотному домені:

$$LSD = \frac{2}{RM} \sum_{m=1}^M \sum_{j=1}^J |G\{X(j, m)\} - G\{Y(j, m)\}|, \quad (4.5)$$

$$G\{X(j, m)\} = \max \{20 \lg(|X(j, m)|), \delta\},$$

$$\delta = \max_{l, k} \{20 \lg(|X(j, m)|)\} - 50$$

де $X(j, m)$ і $Y(j, m)$ – m -ті фрейми дискретних перетворень Фур'є сигналів $x(n)$ і $y(n)$ відповідно, j індекс відліку частоти, J – це кількість спектральних відліків.

Група показників частотної області представлена показником fwSNR:

$$fwSNR = \frac{10}{M} \sum_{m=1}^M \frac{\sum_{k=1}^K W(k, m) \lg \frac{|X(k, m)|^2}{(|X(k, m)| - |Y(k, m)|)^2}}{\sum_{k=1}^K W(k, m)}, \quad (4.6)$$

де m і k – індекси відліку та критичної смуги відповідно, M – кількість фреймів, K – кількість критичних смуг, $|X(k, m)|$ і $|Y(k, m)|$ – амплітудні спектри фреймів $x(n, m)$ та $y(n, m)$, розраховані за допомогою вікна Гауса.

Показник BSD належить до групи перцептивних показників, які враховують особливості слухової системи людини:

$$BSD = \frac{\sum_{m=1}^M \sum_{k=1}^K [B_x(k, m) - B_y(k, m)]^2}{\sum_{m=1}^M \sum_{k=1}^K [B_x(k, m)]^2}, \quad (4.7)$$

де $B_x(k, m)$ і $B_y(k, m)$ – барк-спектри m -того фрейму сигналів $x(n)$ і $y(n)$ відповідно, k – індекс критичної смуги.

Алгоритм обчислення міри PESQ дуже громіздкий. Існує дві версії індикатора PESQ: вузькосмуговий [99] та широкосмуговий [100]. У роботі [2] було використано широкосмуговий варіант PESQ.

Оцінки вимірювань SSNR показані на рис. 4.3.

Очікується реакція SSNR на значення SNR, оскільки якість сигналу, спотвореного шумом, повинна монотонно зростати зі збільшенням SNR. Однак міра SSNR є єдиною мірою серед розглянутих, для якої графіки не залежать від спектральних властивостей шуму. Така поведінка пояснюється змістом показника (4.4).

Дійсно, якщо в знаменнику (4.4) була використана суміш $y(n) = x(n) + \xi(n)$, SSNR не повинен залежати від спектрів процесів $x(n)$ і $y(n)$, оскільки оцінки співвідношення дисперсій $\tilde{D}_x / \tilde{D}_\xi$ сигналу $x(n)$ та шуму $\xi(n)$ розраховуються на

кожному m -тому сегменті. Однак, оскільки нас цікавить різниця у формі сигналів $x(n)$ і $y(n)$ їх потрібно нормалізувати за їх стандартними відхиленнями. Тому рівняння $SSNR \approx SNR$ виконується лише для $SNR \geq 10$ дБ і лише тоді, коли сигнал стаціонарний. Це чітко підтверджується графіком рис. 4.3б, побудованим для сигналу у вигляді стаціонарного коричневого шуму.

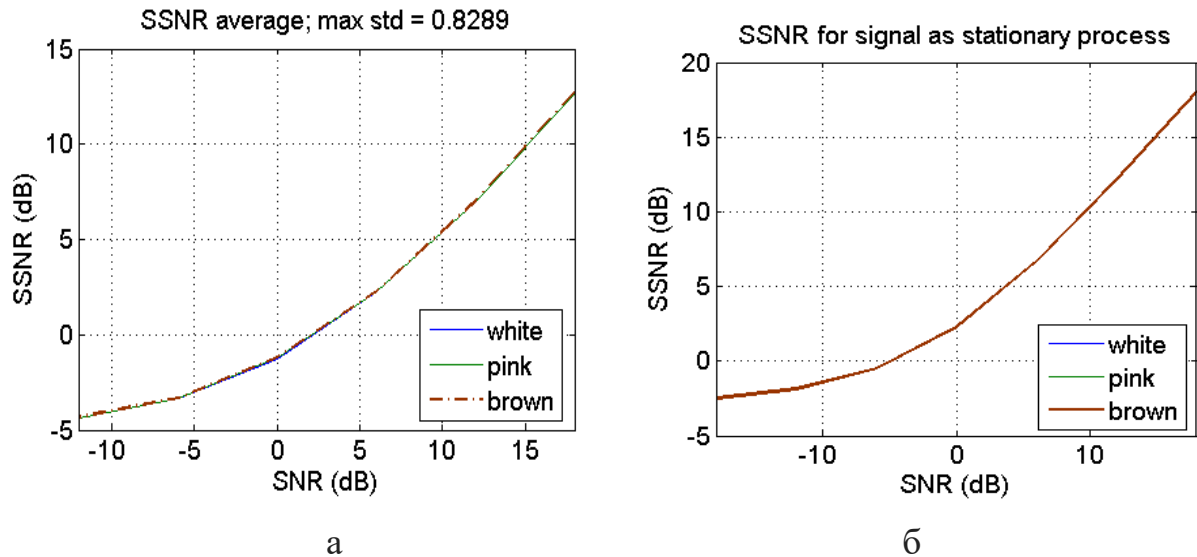


Рис. 4.3 – Оцінки показника SSNR для мови (а) та коричневого шуму як мовленнєвого сигналу (б).

Порівнюючи рис. 4.3а та 4.3б, ми бачимо, що значення SSNR для реальної мови (рис. 4.3а) менші від значень SSNR для стаціонарного сигналу (рис. 4.3б) для всіх розглянутих значень SNR. Ось чому рівняння $SSNR \approx SNR$ не виконується у разі нестационарного мовленнєвого сигналу навіть для $SNR \geq 10$ дБ. Для значень $SNR < -2$ дБ ситуація змінюється на рис. 4.3а, і виконується нерівність $SSNR > SNR$. Це можна пояснити тим, що переважну роль відіграє несхожість чистих і шумних сигналів для малих значень SNR. Дійсно, ми бачимо у випадку стаціонарного випадкового сигналу (рис. 4.3б), що нерівність $SSNR > SNR$ стає помітно вираженою при $SNR \leq 5$ дБ. Таким чином, інтервал $-2 \leq SNR \leq 5$ дБ розмежовує зони впливу на поведінку залежності $SSNR(SNR)$ таких "конкуруючих" факторів, як нестационарний характер мовленнєвого

сигналу та відносний рівень шуму.

Графіки для показників LSD та BSD (рис. 4.4а та 4.4б відповідно) дуже схожі.

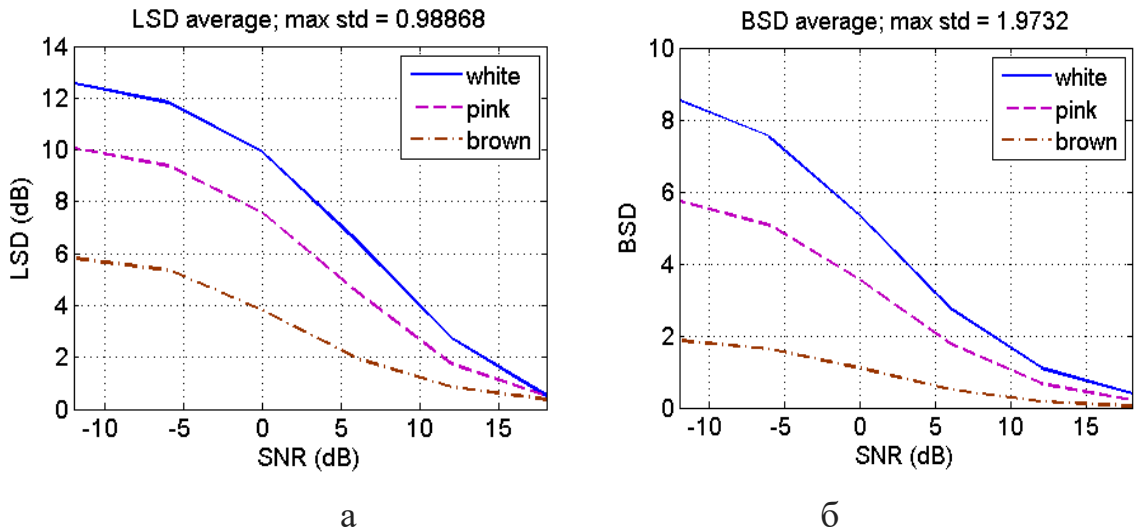


Рис. 4.4 – Оцінки показників LSD (а) та BSD (б)

Коричневий шум викликає найменше погіршення якості сигналу, тоді як білий шум є найбільш руйнівним у всьому діапазоні розглянутих значень SNR. Порівняння цих графіків з результатами суб'єктивної оцінки якості мовлення та розбірливості мови (рис. 4.2) свідчить про те, що графіки узгоджуються не повністю.

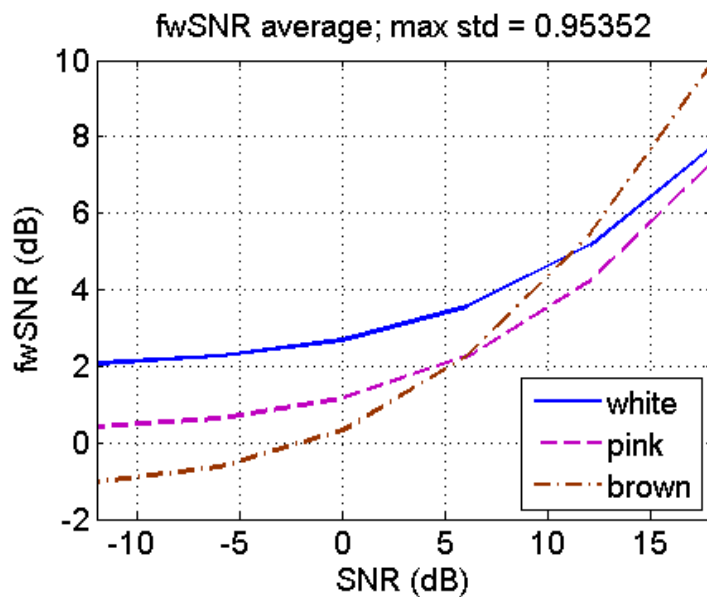


Рис. 4.5 – Оцінки показника fwSNR

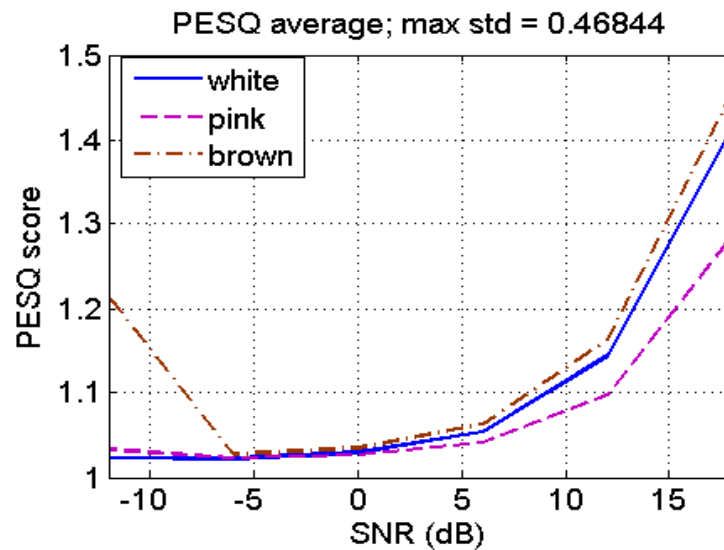


Рис. 4.6 – Оцінки показника PESQ

Поведінка графіків $f_{wSNR}(SNR)$ (рис. 4.5) є досить суперечливою для різних підінтервалів параметра SNR. Однак існують погодження з суб'єктивними показниками якості мови та розбірливості для $SNR > 10$ дБ, а саме: коричневий шум оцінюється як найменш руйнівний, а рожевий – як найбільш руйнівний.

Поведінка графіків PESQ (рис. 4.6) добре узгоджується з суб'єктивними оцінками якості мови та розбірливості при дБ (рис. 4.2). У той же час, показник PESQ не реагує належним чином на ситуацію сильного спотворення сигналу адитивним шумом, а саме: значення PESQ зменшуються зі збільшенням SNR при $SNR < -5$ дБ.

Слід зазначити, що значення об'єктивних оцінок якості мовлення були отримані для тривалості фрейму 32 мс. Для інших значень тривалості фреймів ці результати можуть дещо змінитися, але ці зміни навряд чи будуть суттєвими.

4.3. Аналіз впливу бінаурального ефекту на розбірливість мовлення у приміщенні

Режими прослуховування. Випробування на артикуляцію проводили для

трьох режимів: 1) діотична репрезентація мови через навушники, 2) діотична репрезентація мовлення через комп'ютерні динаміки, 3) діхотична репрезентація мовлення через навушники.

Перший випадок є найпростішим; його використання є «природним» через простоту вимірювання одноканальної ІХ. Тестування запропонованих артикуляційних таблиць та комп'ютерних програм проводилося у 4 етапи: 1) тестування з використанням чистих мовленнєвих сигналів, 2) тестування із сигналами, спотвореними білим, рожевим або коричневим шумом при -10 дБ, 0 дБ та +10 дБ значеннях SNR, 3) тестування з спотвореними сигналами з часом реверберації $T_{20} = 0,3; 0,6; 0,9; 1,1; 1,4; 2; 2,7$ с, 4) тестування з використанням сигналів, спотворених комбінованим ефектом рожевого шуму при значеннях SNR -10 дБ, 0 дБ та +10 дБ та з часом реверберації = 0,3; 0,9; 1,4 с.

Другий випадок з'явився “випадково” під час тестування у першому випадку через необережне виконання інструкцій слухачами, які ігнорували вимогу слухати звук лише через навушники. Слід зазначити, що цей інцидент виявився дуже важливим і повчальним, оскільки отримані результати змусили звернути увагу на деякі вади в організації експерименту, а також на особливу роль ранніх відбиттів. Таким чином, перевірку артикуляції проводили на тих же 4 етапах, що і в першому випадку.

Третій випадок був складнішим, оскільки для його реалізації були потрібні бінауральні ІХ. Для вирішення цього завдання було використано шість бінауральних ІХ з бази даних Aachen Impulse Response [77]. Ці ІХ належать до лекційної кімнати 10,8x10,9x3,15 м. У кімнаті покриття підлоги – паркет, 3 скляні вікна, бетонні стіни та дерев'яні меблі (столи та стільці). Гучномовець був розміщений біля лекційного столу, а мікрофони - біля студентських столів із збільшенням відстані $d = 2,25, 4, 5,56, 7,1, 8,68$ та 10,2 м між мікрофонами та гучномовцем. Час реверберації мало залежав від відстані і становив $RT_{60} = 0,7-0,83$ с. Значення співвідношення енергії ранніх до пізніх відбиттів C_{50} [11] та користь ранніх відбиттів ERB [22] за формулами

$$C_{50} = 10 \lg \left\{ \int_0^{0.05} h^2(t) dt \middle/ \int_{0.05}^{\infty} h^2(t) dt \right\},$$

$$ERB = 10 \lg \left\{ \int_0^{0.05} h^2(t) dt \middle/ \int_0^{0.01} h^2(t) dt \right\},$$

представлені на рис. 4.7 для всіх шести ІХ. Як видно з рис. 4.7а, шкідливий вплив пізньої реверберації помітно знижує якість мовлення та розбірливість із збільшенням відстані від 2,25 м до 5,56 м, але якість та розбірливість мови залишається приблизно такою ж і цілком прийнятною з подальшим збільшенням відстані. Дійсно, відомо, що розбірливість складів не нижче 80% досягається за умови $C_{50} \geq -2$ дБ [101]. Таким чином, розглядаючи лише вплив реверберації на розбірливість мови, ми можемо сказати, що розбірливість мови визначається протиставленням прямого звуку та ранніми відбиттям з одного боку, та пізньою реверберацією з іншого боку. У той же час ранні відбиття повинні сприятливо впливати на розбірливість мови (рис. 4.7б). Енергія прямого звуку зменшується зі збільшенням відстані, тому рівень прямої мови буде неприпустимо низьким для великих відстаней. Цей недолік прямої звукової енергії компенсується додаванням енергії раннього відбиття. Однак зауважимо, що ERB зростає зі збільшенням відстані набагато повільніше, ніж зазначено у [23], і становить лише 3,5 дБ на відстані 8,68 м.

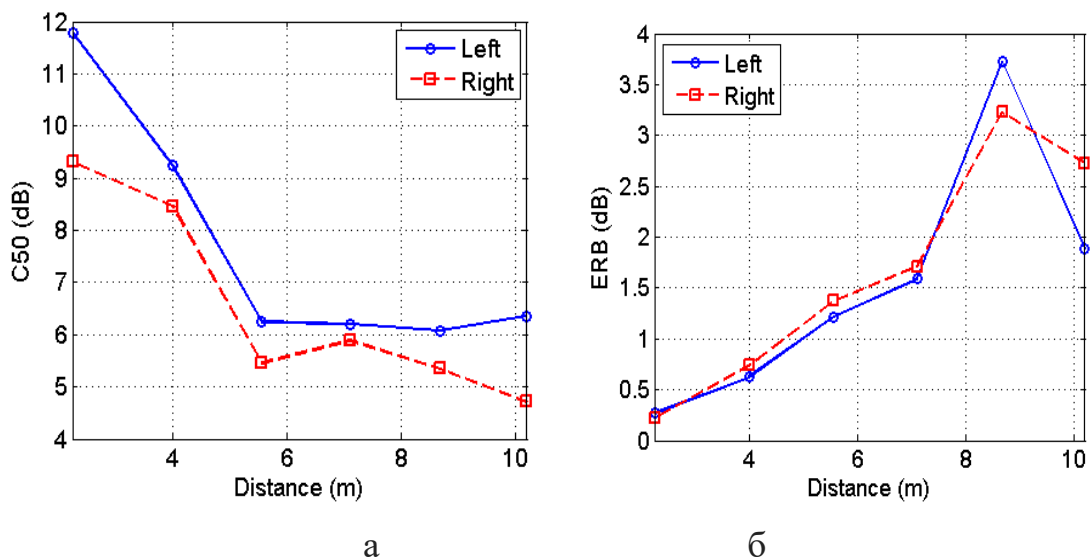


Рис. 4.7 – Значення параметрів C_{50} (а) та ERB (б)

У третьому випадку була використана інша модель сигналу, що прослуховується. Ця модель була сформована як згортка бінауральної ІХ $h(t)$ з адитивною сумішшю чистої мови та шуму:

$$y(t) = [x(t) + n(t)] \otimes h(t), \quad (4.8)$$

Модель (4.8) здається зручною, коли приміщення розглядається як свого роду шумопоглинаючий фільтр. Концепція «співвідношення прямої мови до шуму» [18]

$$SNR = 10 \lg D_x / D_n, \quad (4.9)$$

була використана для моделі (4.8).

Презентація діотичного мовлення через навушники [29].

Артикуляційний тест проводили 26 студентів із нормальним слухом середнього віку 22 років. Рідною мовою для суб'єктів була українська. Прослуховування проводилося через навушники, як того вимагають стандарти [89], [90].

Середні результати суб'єктивної оцінки розбірливості складів представлені на рис. 4.8. Відносне стандартне відхилення для всіх оцінок знаходилося в межах 4-15%.

Криві на рис. 4.8а добре узгоджуються з результатами [102], [103] для помірних та великих значень SNR (0..10 дБ), але узгодження не дуже добре для малих значень SNR (менше -5 дБ). У [103] показано, що при SNR < -5 дБ білий шум практично такий же ефективний, як і рожевий. Навпаки, у [102] зазначено, що білий шум гірший не тільки за рожевий, але й за коричневий. Враховуючи отримані результати, у [29] було рекомендовано повторити суб'єктивне тестування для випадку SNR < -5 дБ після аналізу можливих причин вищезгаданих розбіжностей.

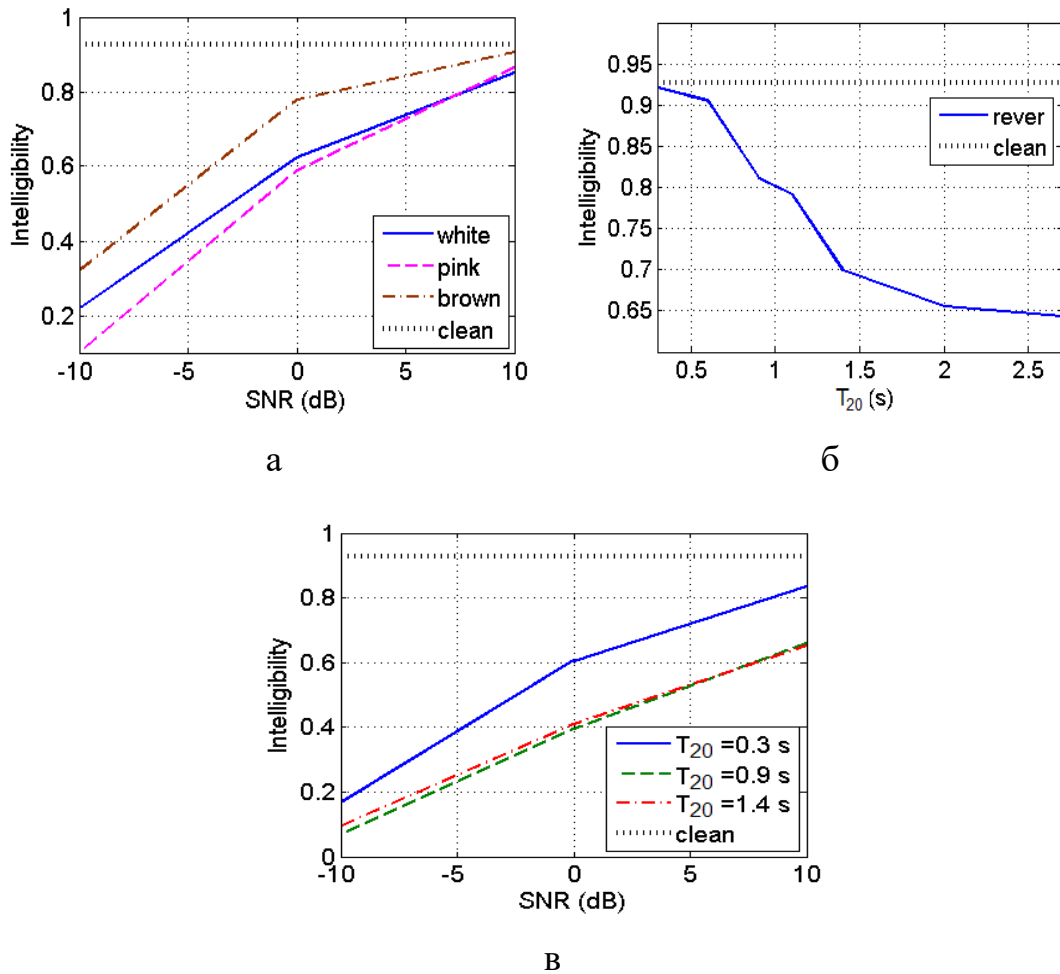


Рис. 4.8 – Оцінки розбірливості для сигналів, спотворених шумом (а), реверберацією (б) та спільною дією шуму та реверберації (в) [29]

Що стосується кривих, показаних на рис. 3б та 3в, вони задовільно узгоджуються з відомими результатами досліджень впливу шуму та реверберації на розбірливість мови [10], [22].

Презентація діотичного мовлення через комп'ютерні динаміки [88]. Цікаве явище було виявлено в [29] при прослуховуванні шумових і ревербераційних сигналів у кімнатах. На відміну від 16 студентів, згаданих вище, троє студентів пройшли тест з відхиленням від вимог, прослухавши сигнали через динаміки комп'ютера або акустичні монітори.

Виразність мовлення виявилася набагато вищою, ніж при прослуховуванні через навушники (рис. 4.9).

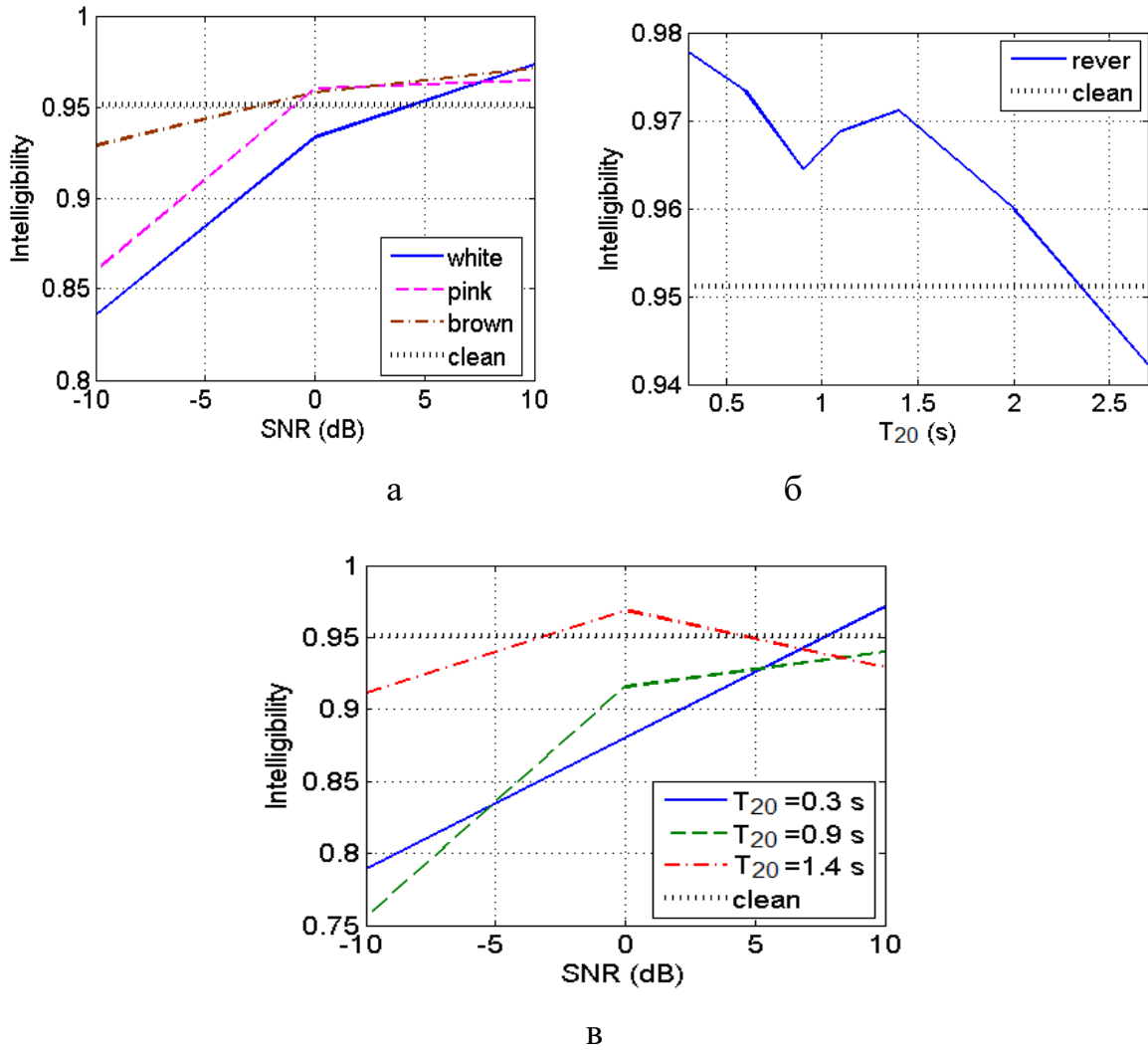


Рис. 4.9 – Оцінки розбірливості для сигналів, спотворених шумом (а), реверберацією (б) та спільною дією шуму та реверберації (в) для випадку прослуховування через акустичні монітори [88]

Частково ці результати можна пояснити дією раннього відбиття в кімнатах. У [22] було показано, що ранні відбиття можуть збільшити SNR до 9 дБ. Водночас таке значне збільшення розбірливості мови не можна пояснити виключно дією ранніх відбиттів. Дійсно, складність збільшення SNR через ефект ранніх відбиттів була показана в [23]. По-перше, було зазначено, що коли ранні відбиття були згруповані на початку ІХ, розбірливість мовлення була значно вищою, ніж тоді, коли перші відбиття були просторово розподілені. Таким чином, лише «пряма» частина енергії ранніх відбиттів вважалася

корисною для розбірливості мови. По-друге, був відзначений сприятливий ефект бінаурального режиму прослуховування, перевага якого становила близько 2..3 дБ.

Щодо «прямої» частини енергії ранніх відбиттів, на жаль, у [23] не було пояснень, як з фізичної точки зору ранні відбиття можуть виходити з прямого звукового напрямку в типовій ситуації, коли оратор і слухачі дивляться один на одного. Можливо, згаданий ефект у реальних кімнатах можна пояснити принципово іншою причиною, а саме - різницею фаз та амплітуд сигналів раннього відбиття, сприйнятих лівим і правим вухом слухача.

Як бачимо, згадане в [29] суттєве збільшення розбірливості мови не можна пояснити виключно фізичними причинами. Ретельний аналіз найбільш ймовірних причин виявленого явища виявив значні недоліки розробленого програмного забезпечення. По-перше, не був реалізований випадковий порядок подання елементів звукових таблиць, що сприяло запам'ятовуванню цих елементів. По-друге, занадто детальна інформація про помилки розпізнавання сприйнятих слів була представлена на екрані монітора, що також сприяло запам'ятовуванню елементів звукових таблиць.

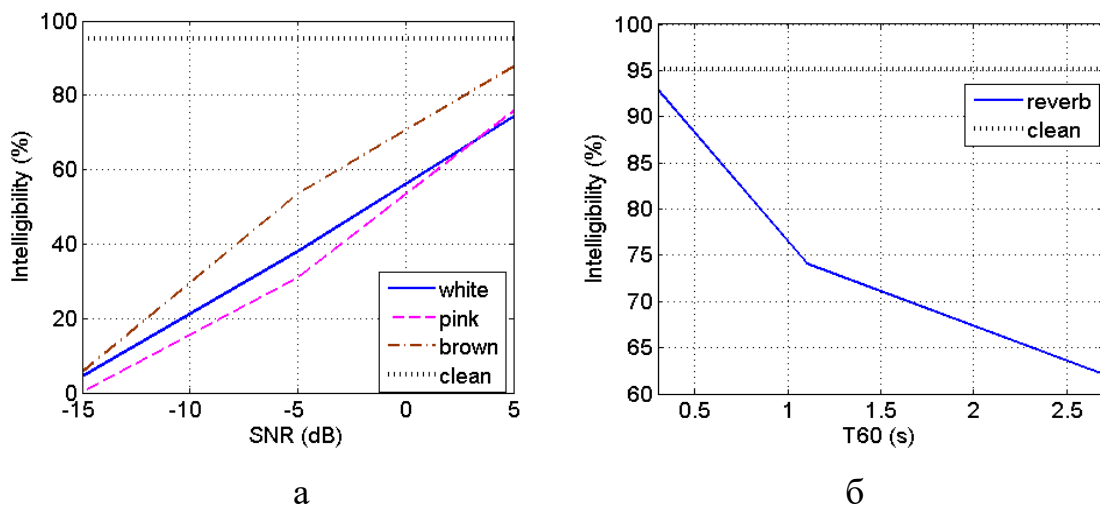


Рис. 4.10 – Оцінка розбірливості, навушники: завади у вигляді шуму (а) та реверберації (б) [88]

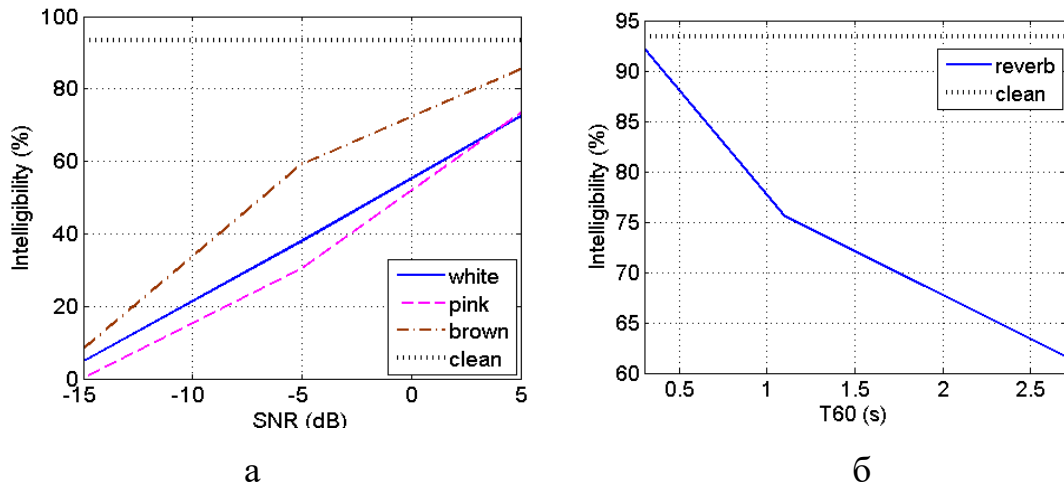


Рис. 4.11 – Оцінка розбірливості, динаміки ПК: завади у вигляді шуму (а) та реверберації (б) [88]

Після виправлення програмного забезпечення експеримент був повторений і до нього були залучені нові випробовувані (20 студентів у віці 20 років із нормальним слухом). Кожен слухач двічі прослуховував одні й ті ж артикуляційні таблиці: через навушники та через динаміки комп'ютера. Результати експерименту для випадків мовлення, спотвореного шумом або реверберацією, наведені на рис. 4.10 і 4.11 [88].

Як видно, показники розбірливості практично не відрізняються для навушників та динаміків, хоча для динаміків спостерігається легка тенденція до збільшення балів. Це невелике збільшення можна пояснити ефектом ранніх відбиттів, які є слабкими (щодо прямого звуку) через малу відстань між слухачем та джерелом звуку (близько 0,6..1 м). Водночас значення розбірливості мовлення на рис. 4.10 і 4.11 кардинально нижчі за ті, що показані на рис. 4.9, і близькі до тих, що показані на рис. 4.8. Таким чином, цей факт свідчить про значне підвищення надійності переглянутого пакету програмного забезпечення та результатів, отриманих з ним.

Очевидно, що отримані результати не в повній мірі описують ситуацію, коли мова представлена через комп'ютерні динаміки або іншу подібну звукову

систему. Тому закономірно поставити таке питання: як зміниться зрозумілість шумової мови в кімнаті на відстані, що перевищує значення 0,6..1 м?

Презентація діхотичного мовлення через навушники. Щоб відповісти на питання про вплив кімнати на розбірливість зашумленого мовлення, було проведено нове дослідження, де бінауральні ІХ [77] використовувалися як двоканальний фільтруючий імпульсний відгук для фільтрації зашумленого мовлення. Відфільтрована мова сприймається за допомогою прослуховування через навушники. У цьому випадку була використана модель (4.8) сприйманого сигналу.

У тестуванні взяли участь 15 слухачів у віці від 19 до 22 років, з них 11 чоловіків та 4 жінок із нормальним слухом. Середня оцінка розбірливості для всіх слухачів показана на рис. 4.12а. Значення стандартного відхилення цих оцінок показано на рис. 4.12б.

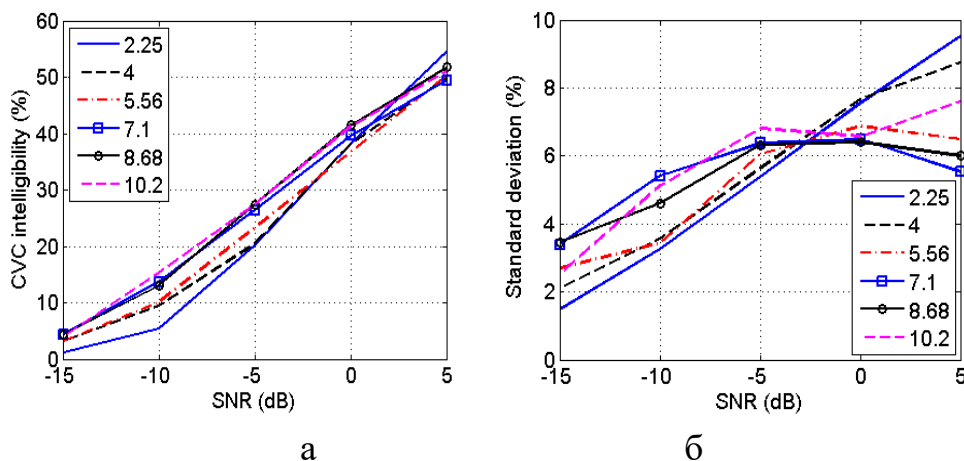


Рис. 4.12 – Оцінки розбірливості мовлення (а) та їх стандартні відхилення (б)

Як видно, для $SNR = 10 \dots -5$ дБ, розбірливість мови на відстанях 7..10 м була приблизно на 10% вищою, ніж на відстані 2,25 м. Для $SNR > -3.5$ дБ розбірливість мови на відстанях 4..10 м нижча, ніж на відстані 2,25 м.

Таким чином, ефект кімнати як шумопоглинаючого фільтра найбільше проявився при поєднанні невеликого значення SNR і вдосконаленої структури ранніх відбиттів. Для помірних і великих значень SNR згубний вплив пізніх відбиттів переважає над прибутковим ефектом ранніх відбиттів. З цього

приводу слід бути обережним, говорячи про виняткову корисність ранніх відбиттів. Слід також врахувати їх руйнівний вплив на мовленнєвий спектр, хоча цей ефект частково компенсується бінауральним прослуховуванням [16].

Важливою особливістю дослідження [6] було те, що відстань між оратором та випробуваними варіювалась від малих значень до великих. Для фіксованої відстані SNR також змінювався від малого до великого. Таким чином, були створені передумови для того, щоб слухачі запам'ятали елементи таблиць артикуляції під час продовження експерименту. Тому для підвищення достовірності результатів у майбутньому необхідно було б дещо модернізувати умови експериментальних досліджень, представивши слухачам елементи артикуляційних таблиць у довільному порядку для різних відстаней між оратором та слухачами.

4.4. Порівняння впливу ранньої та пізньої реверберації на суб'єктивну оцінку розбірливості мовлення.

У розділі 1.5 показано попередні досягнення у дослідженні впливу шуму, ранніх та пізніх відбиттів. У роботах [22] та [23] задіяли дороге обладнання і програмні комплекси. Експериментальна установка для нашого випадку була набагато простішою та дешевшою. Згідно організації дослідження, більшість досліджень проводилося вдома, де студенти використовували просте обладнання, таке як навушники та комп'ютер, та відносно просте комп'ютерне програмне забезпечення [29].

Дев'ять дикторів (7 чоловіків та 2 жінки) взяли участь у записі неспотвореної мови, яка проходила у заглушеній кімнаті. Мікрофон Superlux ECM 999, зовнішня звукова карта USB PRESONUS AudioBox та програмне забезпечення Audacity використовувалися для запису чистого мовлення з частотою дискретизації 44,1 кГц і розрядністю 16 бітів. Були записані односкладові зразки мови типу приголосний-голосний-приголосний. Наприклад, “був” /був/ читали як “Запишіть був зараз”.

Тестові стимули синтезували двома способами.

При використанні першого способу розраховували згортку бінауральної ІХ $h(t)$ з адитивною сумішшю чистого мовлення та шуму:

$$x(t) = [s(t) + n(t)] \otimes h(t), \quad (4.10)$$

де \otimes – символ згортки. У (1) до неспотвореного мовленнєвого сигналу $s(t)$ додається стаціонарний білий гаусівський шум $n(t)$ для забезпечення необхідного значення SNR у діапазоні від -15 дБ до $+5$ дБ. Було використано поняття "відношення прямої мови до шуму"

$$\text{SNR} = 10 \lg D_s / D_n, \quad (4.11)$$

де D_s і D_n – це дисперсії чистого мовлення $s(t)$ та шуму $n(t)$ відповідно.

За другим способом обчислюється згортка першої частини бінауральної ІХ $h(t)$ до 50 мс з адитивною сумішшю чистого мовлення та шуму:

$$x(t) = [s(t) + n(t)] \otimes h_e(t), \quad (4.12)$$

де "корисна" початкова частина $h(t)$:

$$h_e(t) = h(t), t \in 0 \dots 50 \text{ ms.}$$

Мова, оброблена спеціальним програмним забезпеченням [6], прослуховувалася через навушники, а склади ідентифікувалися слухачами та фіксувалися за допомогою комп'ютерної клавіатури під управлінням того ж самого програмного забезпечення.

9 слухачів у віці від 19 до 22 років з нормальним слухом, рідною мовою яких є українська, взяли участь у першому етапі тестування, де використовувалася модель (1). Ще 20 слухачів взяли участь у другому етапі тестування, де використовувалася модель (3).

Як уже згадувалося, тестування проводилося у домашніх умовах за допомогою навушників, комп'ютерів та спеціально розробленого програмного забезпечення. Сигнали подавалися слухачам у довільному порядку. Слухачі слухали 3 набори з одного складу; кожен набір містив 50 складів, для 5 різних значень SNR та для 6 різних RIR, що відповідають 6 відстаням до динаміка. Слухачі виконували кожен етап тестування протягом 10-15 днів, щоб уникнути

ВТОМИ.

Під час тесту на розбірливість слухачі мали визначити склади та записати їх за допомогою клавіатури. Набір складів є важливою частиною перевірки розбірливості, від якої істотно залежить достовірність результату. В якості прототипу були взяті набори складів стандарту ГОСТ Р 50840-95 [89] і розроблено дев'ять українських наборів розбірливості мовлення [88]. Було розроблено три варіанти кожного набору з урахуванням можливості неточного сприйняття та запису складів.

Автоматичне випробування артикуляції складається з чотирьох кроків, і лише третій крок не автоматизований:

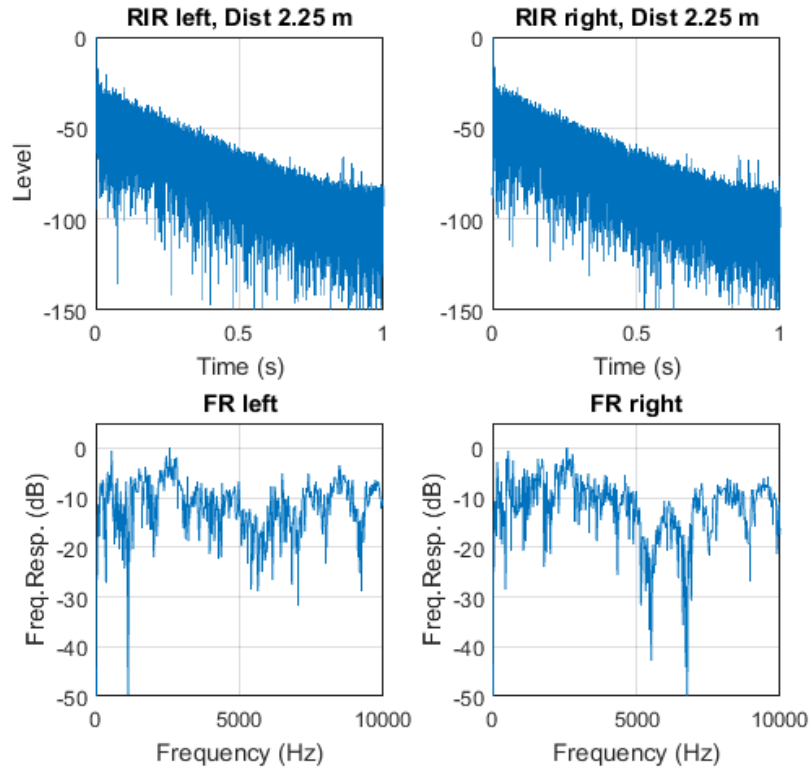
- моделювання спотвореної мови;
- програвання спотвореної мови слухачеві;
- запис слухачем сприйнятих складів;
- розрахунок розбірливості мовлення.

Спотворені склади програвались у випадковому порядку. Слухачі мали змогу виправити неправильно записаний з клавіатури склад.

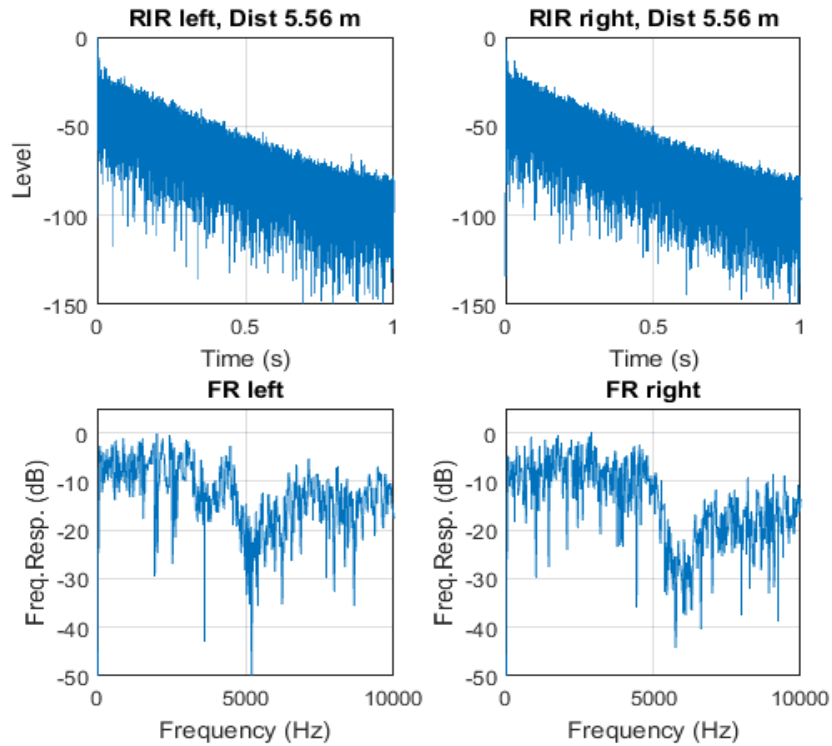
Для досліджень було використано шість бінауральних ІХ з бази даних Aachen Impulse Response (AIR) [78], [77], які належать до аудиторії 10,8 м, 10,9 м, 3,15 м. У кімнаті було 3 скляні вікна, поверхня підлоги - паркет, бетонні стіни, стільці та дерев'яні столи. Гучномовець був розташований біля місця викладання, а мікрофон був розміщений у різних місцях розміщення студентів на 6 відстанях, що варіювалися від 2,25 м до 10,2 м до динаміка. Час реверберації RT_{60} для ІХ збільшився з 0,70 с до 0,72 с для $d = 2.25 - 4.00m$ і з 0,79 с до 0,83 с для $d = 5.56 - 10.2m$. Форми хвиль бінауральних ІХ та частотні характеристики на відстанях 2,25 м та 5,56 м наведені на рис. 4.13а та 4.13б відповідно.

Значення співвідношення енергії ранніх та пізніх відбиттів $C_{50} = 10lg(E_0^{50}/E_{50}^{\infty})$ та енергії раннього відбиття $ERB = 10lg(E_0^{50}/E_0^{10})$ [22] для всіх шести ІХ $h(t)$ показано на рис. 4.13а та 4.13б відповідно

$$(E_a^b = \int_a^b h^2(t) dt).$$



a



б

Рис. 4.13 – Бінауральні ІХ кімнати та АЧХ для $d = 2,25$ м (а) та $d = 5,56$ м (б).

Відомо, що розбірливість складів не нижче 80% досягається, коли $C_{50} \geq -2\text{dB}$ [101]. Таким чином, з рис. 4.13а випливає, що всі ІХ здатні забезпечити хорошу розбірливість мови.

Крім того, з рис. 4.13а випливає, що значення C_{50} дуже мало змінюються для відстаней 5-10 м, тобто можна припустити, що вплив ранніх відбиттів на розбірливість мови для цих відстаней майже однаковий. Зазначимо, що результати досліджень, наведені в статті [7], спростовують це припущення.

Що стосується залежності ERB (енергії ранніх відбиттів) від відстані, то з рис. 4.13б випливає, що відносна (до прямого звуку) інтенсивність ранніх відбиттів монотонно зростає із збільшенням відстані.

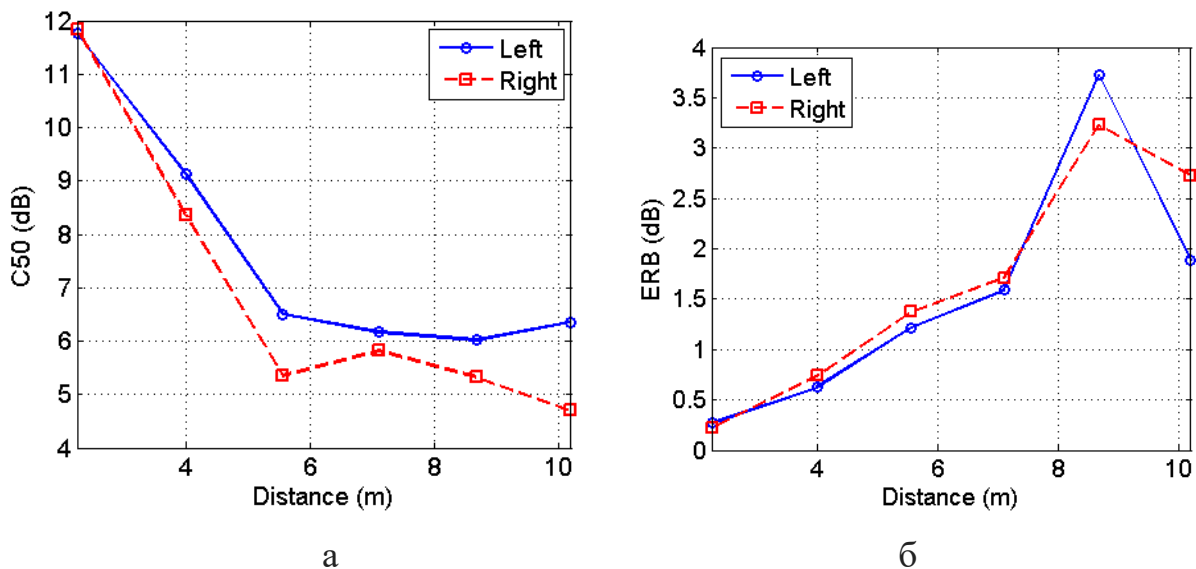


Рис. 4.14 – Значення C_{50} та ERB для бінауральних ІХ.

Результати випробувань розбірливості на першому етапі в середньому для 9 слухачів наведені на рис. 4.15а, б, а оцінки стандартного відхилення цих результатів наведені на рис. 4.15в. Як видно, для всіх розглянутих значень SNR ($\text{SNR} = -15\dots+5$ дБ) розбірливість мови є найвищою для відстані 2,25 м і найнижчою для відстані 5,56 м майже для всіх значень SNR.

Результати артикуляційних випробувань другого етапу в середньому для 20 слухачів наведені на рис. 4.16а, б, а оцінки стандартного відхилення цих результатів наведені на рис. 4.16в. Цікаво, що результати не сильно змінилися

якісно, але вони помітно змінилися кількісно. Дійсно, розбірливість мови збільшилася на 4-7% у діапазоні значень SNR від -10 дБ до +5 дБ. Водночас розбірливість мови все ще є найвищою для відстані 2,25 м і найнижчою для відстані 5,56 м для всіх значень SNR.

Різниця оцінок розбірливості між графіками рис. 4.14 та 4.15 сягає 8-10 дБ при SNR=5 дБ, тобто при малому рівні фонового шуму. При менших SNR ця різниця не така помітна. Можна зробити висновок, що вплив пізньої реверберації стає найбільш помітним для слухачів при SNR \geq 5 дБ.

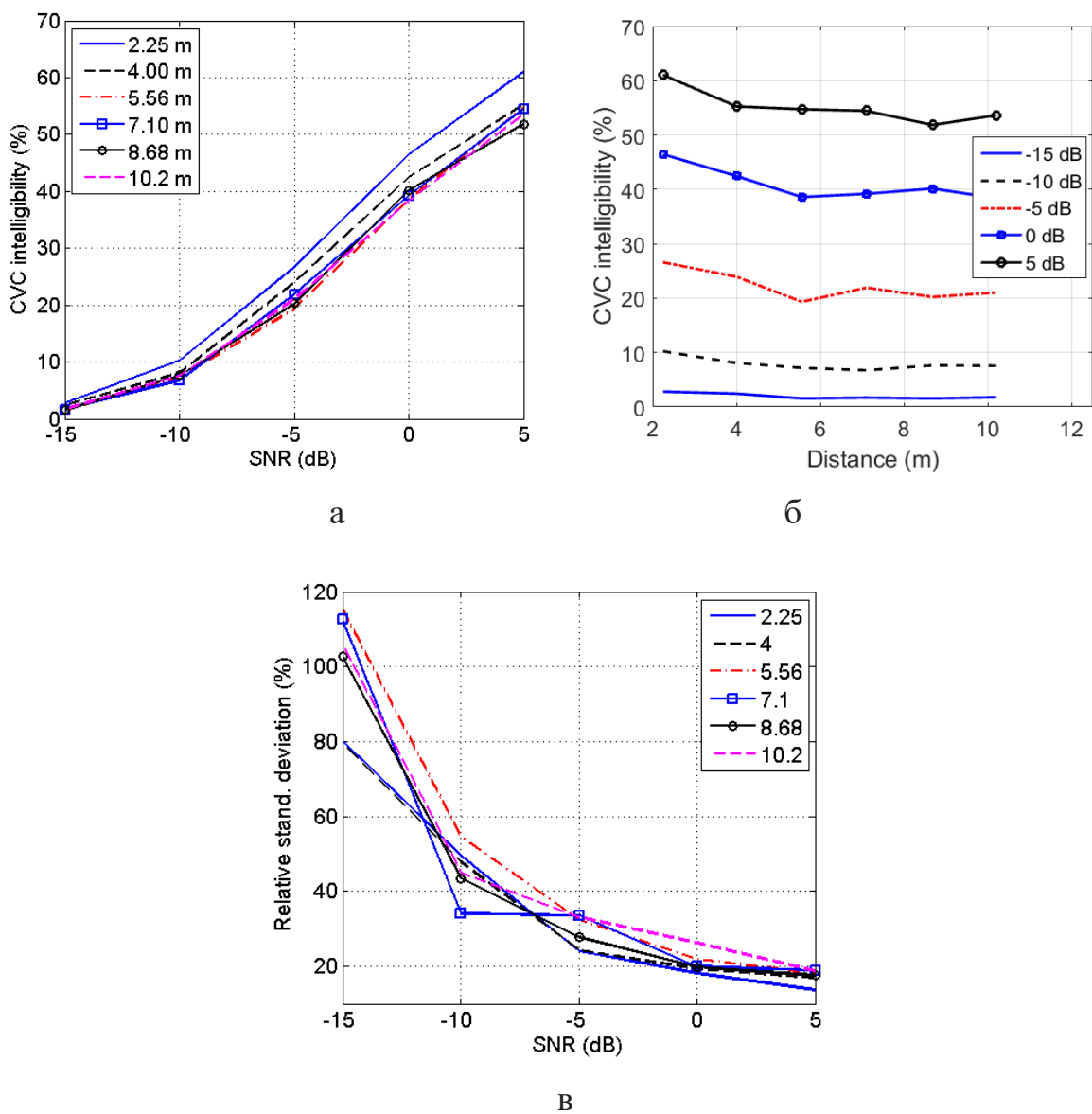


Рис. 4.15 – Середні оцінки розбірливості мови (а, б) та стандартне відхилення (в) для різних відстаней d для впливу ранніх та пізніх відбиттів.

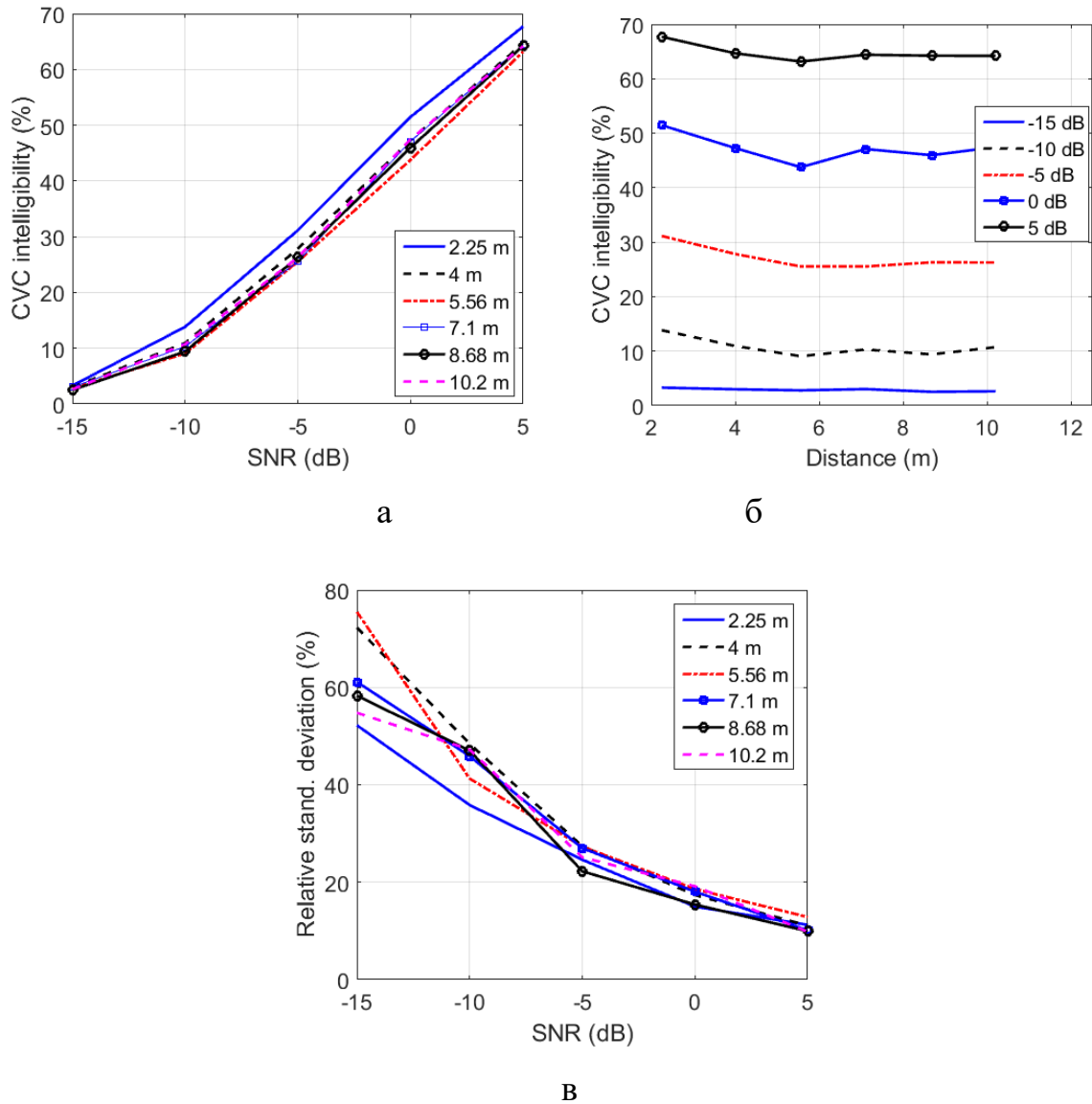


Рис. 4.16 – Усереднені оцінки розбірливості мови (а, б) та стандартне відхилення (в) для різних відстаней d та для впливу ранніх відбиттів.

Таким чином, слід бути обережним, говорячи про корисність ранніх відбиттів, враховуючи їх руйнівний вплив на спектр мовлення, хоча цей ефект частково компенсується бінауральним сприйняттям [16].

Результати оцінки розбірливості мови об'єктивними модуляційним та формантно-модуляційним методами представлені на рис. 4.17а, де показані значення розбірливості слів. Подібні результати для об'єктивного модуляційного методу (значення STI індексу передачі мови) показані на рис. 4.17б.

Коли переважає шум, використовується форматний метод [103], і розбірливість мови розраховується у два етапи. Спочатку розраховується розбірливість артикуляції A :

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

де p_k – розподіл ймовірностей формант у діапазонах частот, $P_k(E_k)$ і E_k відповідно – коефіцієнти сприйняття мовлення та ефективні рівні сприйняття мовленнєвого сигналу у k -тій смузі частот Δf_k відповідно. Існують деякі види $P_k(E_k)$, запропоновані Н. Покровським, М. Сапожковим та кафедрою акустичних та мультимедійних систем Національного технічного університету України (Київський політехнічний інститут ім. Ігоря Сікорського) [104].

Другим кроком розраховується розбірливість слів $W = W(A)$, де $W(A)$ – відома залежність між W і A [103].

Коли переважає реверберація, використовується формантно-модуляційний метод [105], а E_k можна обчислити як

$$E_k = \frac{1}{14} \sum_{i=1}^{14} SNR_{k,i},$$

$$SNR_{k,i} = 10 \lg \frac{m_{k,i}}{1 - m_{k,i}},$$

$$m_{k,i} \approx \left| \frac{\int_0^{\infty} h_{ek}^2(t) \exp(-j2\pi F_i t) dt}{\int_0^{\infty} h_{ek}^2(t) dt} \right|,$$

$$h_{ek}(v) = \int_{-\infty}^{\infty} h(v) h_{k0}(z - v) dv,$$

де $h_{k0}(v)$ – імпульсна характеристика k -того смугового фільтра.

Індекс передачі мовлення (STI) можна обчислити як

$$STI = \sum_{k=1}^7 \alpha_k \cdot T_k - \sum_{k=1}^6 \beta_k \cdot \sqrt{T_k \cdot T_{k+1}} \square$$

$$T_k = \begin{cases} 0, E_k \leftarrow 15; \\ (E_k + 15)/30, -15 < E_k < +15; \\ 1, E_k > +15; \end{cases}$$

де α_k і β_k є ваговими коефіцієнтами [106].

Як видно з рис. 5а, б, побудованих з використанням вищенаведених формул, найнижча розбірливість мови у разі переважної реверберації повинна бути на відстані 7,1 м, але не на відстані 5,56 м.

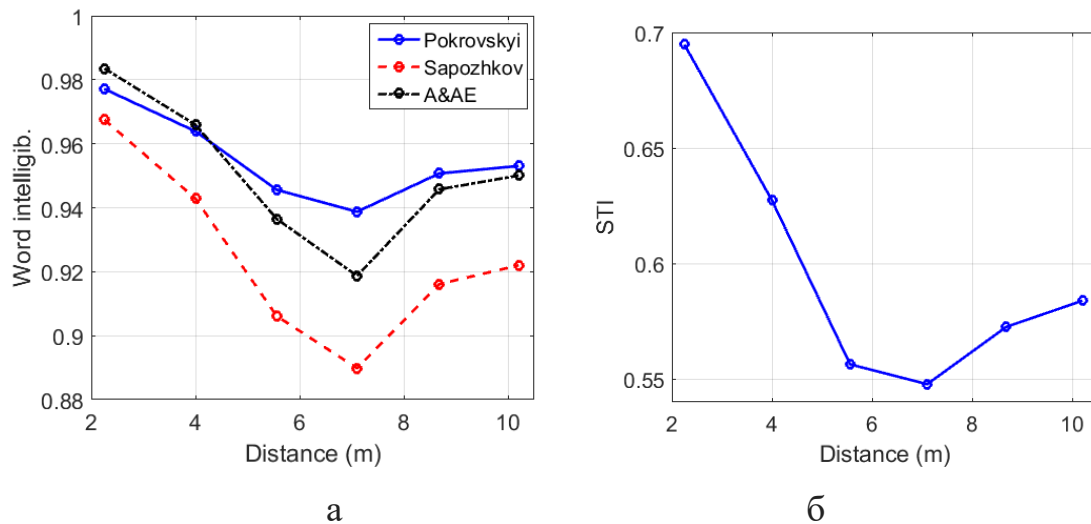


Рис. 4.17 – Оцінка розбірливості мовлення.

Виявлену різницю в результатах суб'єктивної та об'єктивної оцінки розбірливості мовлення можна пояснити тим, що об'єктивна оцінка розбірливості використовувала спрощений підхід, при якому ефект шумових перешкод не враховувався. Звичайно, це припущення доцільно перевірити у майбутніх дослідженнях.

Очевидно, що виявлена відмінність набагато менш важлива порівняно з виявленим фактом, що наявність ранніх відбиттів призводить до зниження розбірливості мови. Як видно, це зниження не може бути компенсоване навіть за допомогою бінаурального прослуховування. Цей результат добре узгоджується з висновками [23] про те, що збільшення інтенсивності ранніх відбиттів не еквівалентно збільшенню інтенсивності прямого звуку. На жаль, в [23] не було зроблено припущень щодо можливої причини цього явища. Виходячи з графіків рис. 4.13, ми можемо спробувати усунути цей недолік і припустити, що причиною цього явища є значна нерівномірність амплітудно-частотної характеристики приміщення, що обумовлено наявністю ранніх

відбиттів, що призводить до зниження якості мовленнєвого сигналу. Оскільки це припущення ґрунтується на результатах досліджень лише одного лекційного кабінету, у майбутньому доцільно перевірити його обґрунтованість для кількох кабінетів одного типу.

4.5. Моделювання впливу ранніх відбиттів на розбірливість мови

В результаті кількох звукових відбиттів від стелі, підлоги, стін та інших поверхонь мовленнєвий сигнал на виході мікрофона відрізняється від чистого мовленнєвого сигналу, що негативно впливає на розбірливість мови і, як наслідок, на продуктивність системи автоматичного розпізнавання мови. Таким чином, проблема впливу реверберації на розбірливість мови в системах голосового управління є актуальною.

Дослідження впливу ранніх звукових відбиттів на розбірливість мови триває тривалий час. У [37] показано, що слухова система людини інтегрує дію ранніх звукових відбиттів з дією прямого звуку, якщо відбиття надходять до слухача із затримкою не більше 30-40 мс щодо прямого звуку. У [26] було виявлено, що ранні відбиття із затримкою 0-95 мс позитивно впливають на розбірливість мови, тоді як пізніші - погіршують розбірливість, діючи як шумова завада. Цей факт дозволив використати параметр C_{95} як міру розбірливості мови [107]. У [108] розглядається індекс ясності C_{50} , подібний до параметра C_{95} . Також зазначається, що C_{50} набагато краще, ніж час реверберації характеризує вплив реверберації на продуктивність систем автоматичного розпізнавання мови. У [107] було показано, що індекс передачі мовлення (STI) є не менш ефективним показником розбірливості мови порівняно з C_{50} .

Оскільки різні частини ІХ кімнати (RIR) по-різному впливають на розбірливість мови, дослідники часто використовують техніку «вирізання» досліджуваної частини ІХ та згортання цієї частини з мовленнєвим сигналом. Цей прийом був використаний, наприклад, у [109] при аналізі впливу ранніх та пізніх відбиттів на роботу систем автоматичного розпізнавання мови. Подібний

прийом був використаний у [110] при аналізі впливу реверберації на слухову систему людей з кохлеарними імплантатами, а також у [7] при аналізі сприйняття реверберованої мови слухачами з нормальним слухом.

Однак замість того, щоб вирізати потрібну частину ІХ, цю частину можна синтезувати за допомогою комп'ютерного моделювання [111], [112], [113]. У цьому підході зазвичай використовуються фізичні [26] або математичні [22], [23] моделі реальних кімнат.

У той же час використання надто точних і складних моделей реальних кімнат може бути непродуктивним, коли дослідники змушені мати справу з виявленими явищами, не даючи переконливих пояснень. На жаль, праці [22], [23], присвячені аналізу ранніх відбиттів, не позбавлені цього недоліку.

Модель ІХ у вигляді імпульсного потоку:

$$h(t) = [\delta(t) + \sum_{n=1}^N a_n \delta(t - t_n)]e^{-\alpha t}, \quad (4.13)$$

де $\delta(t)$ – дельта -функція Дірака, a_n і t_n є випадковими величинами, N – це кількість імпульсів у часовому інтервалі $[0, T_{60}]$, T_{60} – це час реверберації, $\alpha = \ln(10^3)/T_{60}$ – це параметр, що визначає швидкість спаду ІХ. Зауважимо, що модель (4.13) є частиною більш загальної моделі ІХ під назвою «гібридна крива спаду енергії» [114].

У приміщеннях у формі коробки щільність імпульсного потоку d (4.13) з плином часу збільшується відповідно до квадратичного закону [35]:

$$d(t) = 4\pi c^3 t^2 / V \quad (4.14)$$

У статті [5] представлені дві математичні моделі раннього відображення, простіші за модель (4.13)-(4.14). Перша модель являє собою єдине відбиття, а друга модель – це потік імпульсів постійної щільності.

Оскільки ситуація з одним відбивачем вивчена досить добре [37], розгляд першої моделі дасть змогу судити про правильність обраного методу аналізу ранніх відбиттів. Щодо моделі у вигляді потоку імпульсів постійної щільності, її можна розглядати як перше наближення до повної моделі (4.13)-(4.14).

Модель ІХ з одним відбиттям можна описати простим виразом

$$h(t) = \delta(t) + g\delta(t - t_0), \quad (4.15)$$

де t_0 – затримка часу, g ($0 \leq g \leq 1$) - рівень відбиття. Графік функції (4.15) та відповідної амплітудно-частотної характеристики

$$|H(f)| = \sqrt{1 + g^2 + 2g \cos 2\pi f t_0}$$

показані на рис. 4.18 для $t_0 = 0,02$ с, $g = 0,9$.

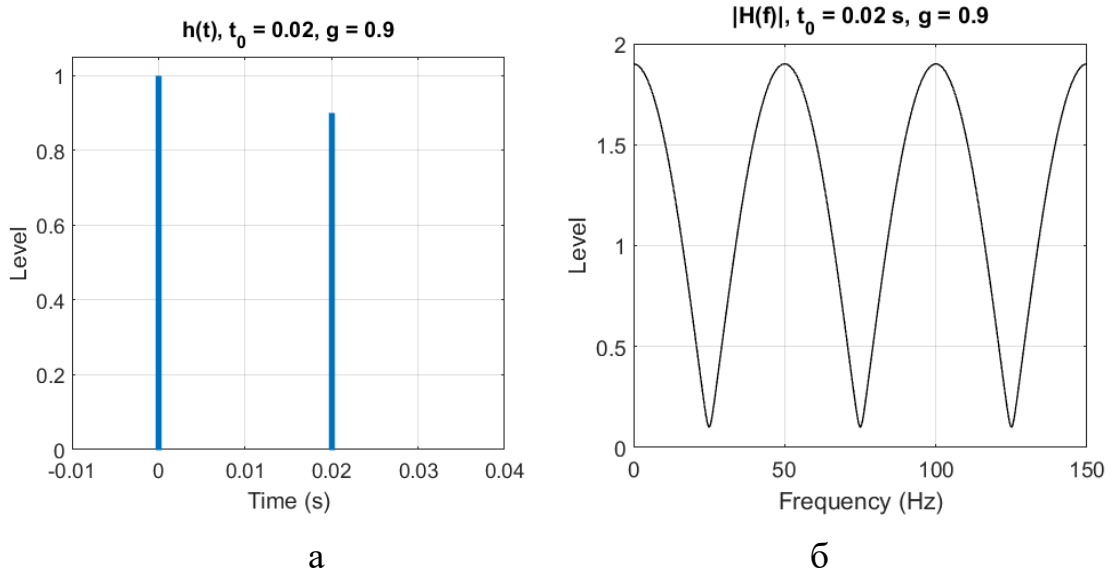


Рис. 4.18 – IX (а) та АЧХ (б) для першої моделі.

Подібні графіки для другої моделі IX, описаної (1.13), для $T_{60} = 1$ с і постійної частоти відбивачів $d = 160$ Гц, показані на рис. 4.19.

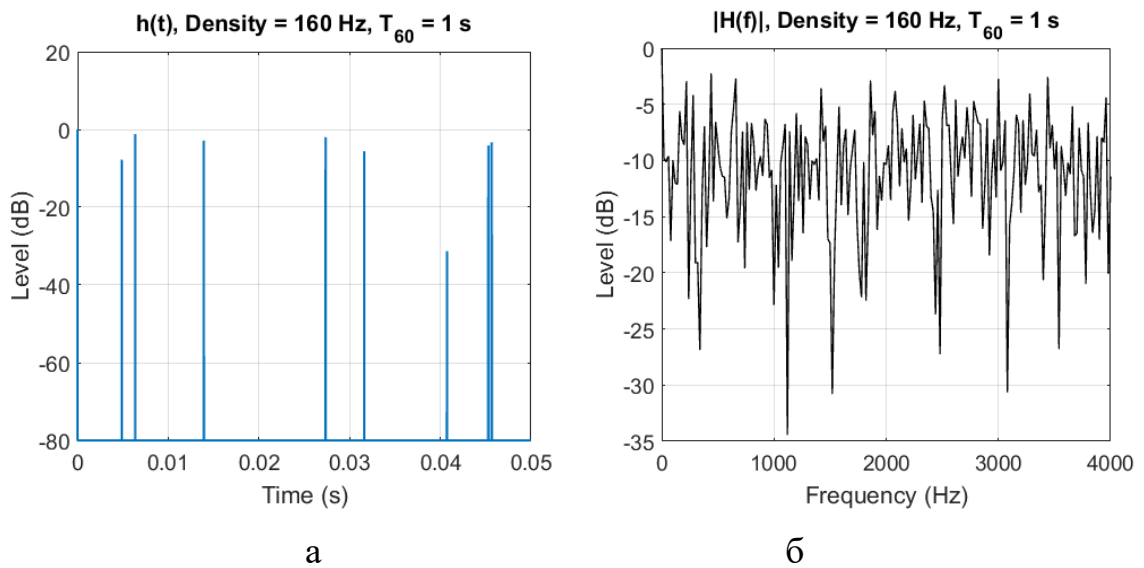


Рис. 4.19 – IX (а) та АЧХ (б) для другої моделі

Діаграми АЧХ, показані на рис. 4.18 і 4.19 чітко демонструють, що головною причиною погіршення якості мовленнєвого сигналу є значна нерівномірність частотної характеристики кімнатної. Оскільки поняття якості мови та розбірливості не є однозначно пов'язаними, графіків на рис. 4.18 та 4.19 недостатньо для пояснення причин погіршення розбірливості мови.

У роботі [5] індекс передачі мовлення (STI) використовується як міра виразності мовлення, спотворена ранніми відбиттями. Оцінки STI добре узгоджуються з результатами суб'єктивної оцінки розбірливості мовлення [107].

Відповідно до методу модуляції [115], [116], STI оцінюється наступним чином:

$$STI = \sum_{k=1}^7 \alpha_k \cdot MT_k - \sum_{k=1}^6 \beta_k \cdot \sqrt{MT_k \cdot MT_{k+1}} \quad (4.16)$$

$$MT_k = \frac{1}{14} \sum_{i=1}^{14} T_{ki}$$

$$T_{ki} = \begin{cases} 0, & E_{ki} < -15; \\ (E_{ki} + 15)/30, & -15 \leq E_{ki} \leq +15; \\ 1, & E_{ki} > +15; \end{cases}$$

$$E_{ki} = 10 \lg \frac{m_{ki}}{1 - m_{ki}}$$

$$m_{ki} = \left| \int_0^{\infty} h_k^2(t) \exp(-j2\pi F_i t) dt \right| / \int_0^{\infty} h_k^2(t) dt \quad (4.17)$$

де $h_k(t)$ – результат фільтрації $h(t)$ k -тим смуговим фільтром (використовуються 7 октавних фільтрів з центральними частотами від 125 Гц до 8 кГц); F_i – частота модуляції (на практиці використовується 14 значень F_i , в діапазоні від 0,63 Гц до 12,5 Гц); α_k і β_k є вагою та коефіцієнтами надмірності відповідно [115].

Для першої моделі ІХ можна записати

$$\begin{aligned} h_k(t) &= \int_{-f_{k2}}^{-f_{k1}} H(f) e^{j2\pi ft} df + \int_{f_{k1}}^{f_{k2}} H(f) e^{j2\pi ft} df = \\ &= 2[f_{k2} \text{Sa}(2\pi f_{k2} t) - f_{k1} \text{Sa}(2\pi f_{k1} t)] + \\ &+ 2g[f_{k2} \text{Sa}(2\pi f_{k2}(t-t_0)) - f_{k1} \text{Sa}(2\pi f_{k1}(t-t_0))] \end{aligned} \quad (4.18)$$

де f_{k1} і f_{k2} – нижня та верхня частоти відсікання k -того смугового фільтра з АЧХ прямокутної форми $H(f)$, $Sa(x) = \sin x / x$. Форма функції (4.18) для $k = 1$ і $k = 4$ для $t_0 = 0,029$ с і $g = 0,9$ показана на рис. 4.20 зі зміщенням вздовж осі часу на 0,1 с. Такий зсув необхідно здійснити в обчисленнях за допомогою співвідношення (4.16), щоб лівий хвіст функції був врахований.

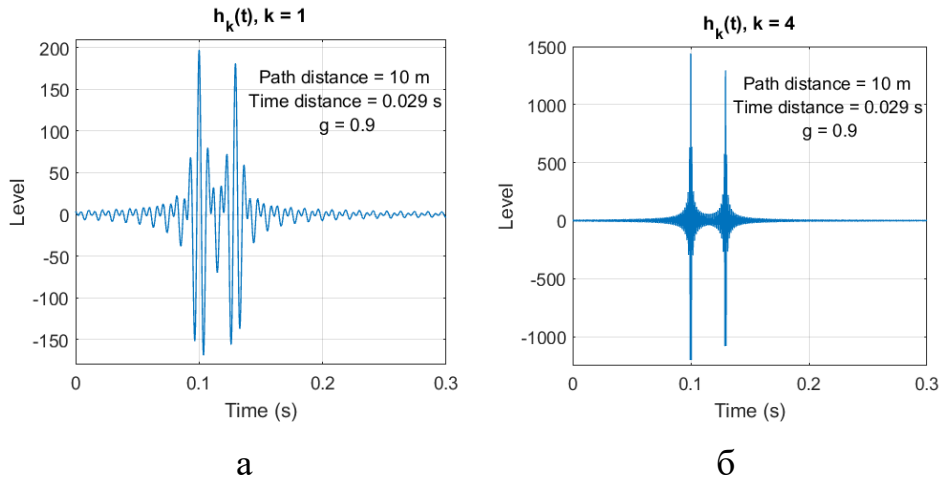


Рис. 4.20 – Форма функції (б) для $k = 1$ (а) та $k = 4$ (б).

Узагальнення (4.18) до другого випадку моделі RIR можна записати так:

$$h_k(t) = 2[f_{k2} Sa(2\pi f_{k2}t) - f_{k1} Sa(2\pi f_{k1}t)] + 2 \sum_{n=1}^N a_n [f_{k2} Sa(2\pi f_{k2}(t-t_n)) - f_{k1} Sa(2\pi f_{k1}(t-t_n))] \quad (4.19)$$

Як можна помітити, розрахунки STI легко реалізувати за допомогою (4.15), (4.16), (4.18) та (4.19).

Результати моделювання залежності STI від різниці прямого та відбитого звукових шляхів (часів приходу) наведені на рис. 4.21 для різних відносних рівнів відбитого звуку. Як бачимо, у разі сильного відбиття ($g = 0,9$), якщо різниця у часі прибуття близька до 30 мс, рейтинг STI зменшується на 1 бал (Таблиця 4.1) від «відмінно» до «добре». Подальше збільшення різниці в часі до 40 мс призводить до зменшення розбірливості мови ще на 1 бал, аж до відмітки «задовільно». У той же час значення STI близьке до 1, якщо різниця в часі не перевищує 10-13 мс.

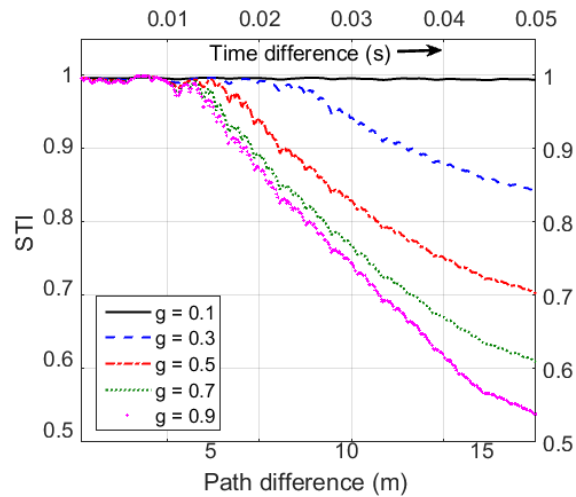


Рис. 4.21 – Оцінки STI для першої моделі IX.

Таблиця 4.1 – Оцінка STI та рейтинг зрозумілості [116]

STI	Оцінка розбірливості
> 0.75	Відмінно
0.60-0.75	Добре
0.45-0.60	Задовільно
0.30-0.45	Незадовільно
< 0.30	Погано

У випадку відбиттів помірної сили ($g = 0,5$), оцінка розбірливості зменшується з «відмінно» до «добре», якщо різниця на шляху близька до 13,6 м (різниця в часі 40 мс).

Залежності середнього значення STI від щільності ранніх відбиттів в інтервалі 0-50 мс наведені на рис. 4.20. Усереднення було проведено по 100 оцінкам STI для значень часу реверберації 0,4 с, 0,7 с та 1 с. Кінці вертикальних ліній відповідають межах 95% довірчого інтервалу.

На рис. 4.22 розглянуто два випадки. У першому випадку (рис. 4.22а) щільність встановлено на рівні 40 Гц, 375 Гц, 1030 Гц та 2060 Гц. Можна помітити, що середня розбірливість мінімальна при щільності, близькій

до 375 Гц. Наведений на рис. 4.22б, подібний графік для щільності від 40 Гц до 320 Гц дозволяє зробити висновок, що мінімум середньої розбірливості розташований у $d \approx 375$ Гц.

Наведені на рис. 4.22 графіки можна інтерпретувати також з точки зору кількості відбиттів N у часовому інтервалі 0-50 мс. Дійсно, швидке зменшення розбірливості зі збільшенням щільності з 40 Гц до 160 Гц (рис. 4.21а) відповідає збільшенню середньої кількості відбиттів з 2 до 8. Повільне зниження розбірливості продовжується до 360-375 Гц (18-19 відбиттів). Подальше збільшення N з 18 до 100 навіть дещо збільшує середню розбірливість. Однак це збільшення настільки незначне, що можна сказати, що розбірливість мовлення стабілізувалася на мінімальному рівні.

З графіків на рис. 4.22 випливає, що збільшення T_{60} з 0,4 с до 1 с призводить до незначного погіршення розбірливості мови. Цей результат також можна пояснити властивостями першої моделі. Дійсно, збільшення T_{60} призводить до збільшення впливу пізніх відбиттів і, як наслідок, погіршується розбірливість мови.

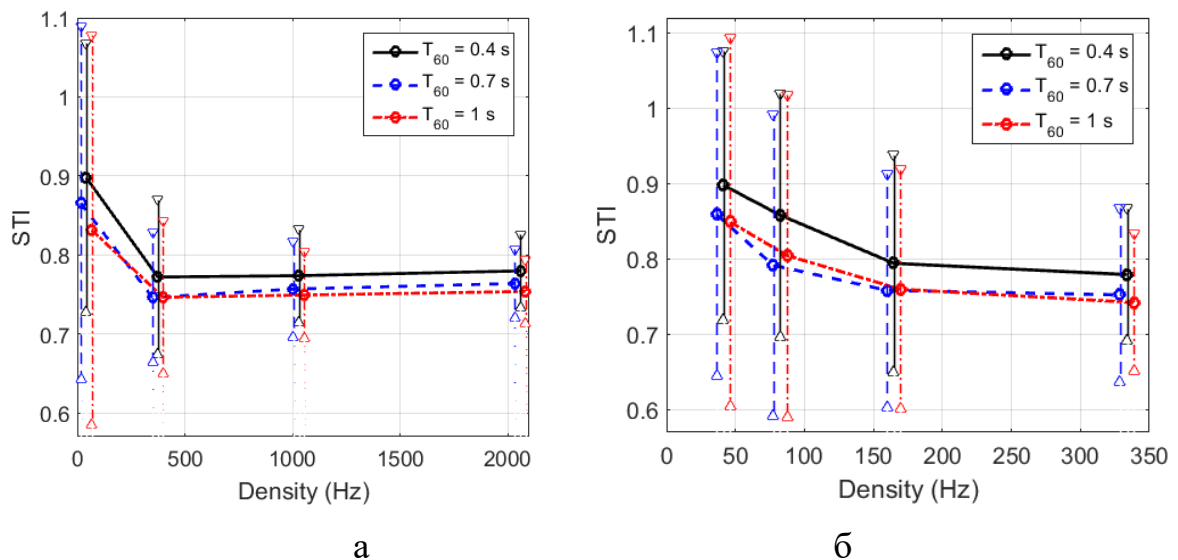


Рис. 4.22 – Оцінки STI для другої моделі ІХ

Узагальнюючи результати, отримані для другої моделі, можна прийти до таких висновків:

- збільшення середньої кількості відбиттів з 2 до 8 призводить до швидкого погіршення розбірливості мовлення з “відмінно” на “добре” за час реверберації від 0,4 с до 1 с;

- існує мінімум середньої розбірливості мови, що знаходиться близько до межі між «хорошим» і «задовільним», коли середня кількість відбиттів близька до 18;

- подальше збільшення середнього числа відбиттів до 100 призводить до ледь помітного підвищення розбірливості мови;

- для невеликої кількості ранніх відбиттів (від 2 до 8) розбірливість мовлення може змінюватись у широкому діапазоні; цей діапазон помітно звужується із збільшенням кількості відбитків;

- збільшення T_{60} з 0,4 с до 1 с призводить до незначного зниження розбірливості мови.

Отримані результати для першої моделі добре узгоджуються з відомим висновком [37], що перше сильне відбиття незначно погіршує розбірливість мови за умови, що затримка відбитого звуку щодо прямого звуку не перевищує 30-40 мс. Крім того, ці результати узгоджуються з висновком [117], що шкідлива дія першого відбиття зростає зі збільшенням його амплітуди та затримки.

Добре узгоджується також друга модель у випадку невеликої (від 2 до 8) кількості відбиттів в інтервалі 0-50 мс. Ця кількість відбиттів близька до реальної ситуації, коли слухач у кімнаті у формі коробки сприймає 6 первинних відбиттів. Подальше збільшення кількості відбиттів можливе шляхом додавання спеціальних відбивачів, а також шляхом додавання вторинних та третинних відбиттів.

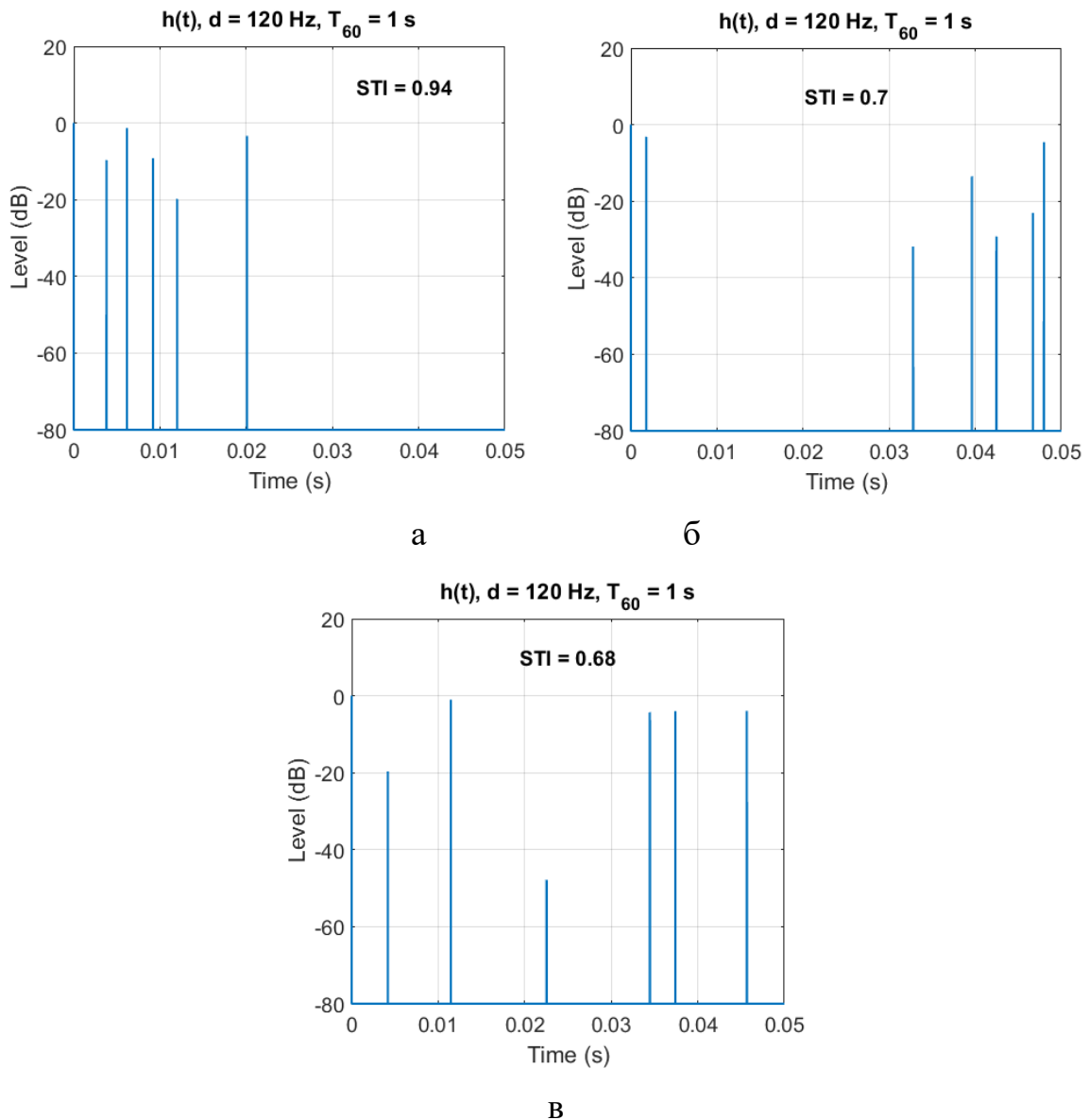


Рис. 4.23 – ІХ для випадку невеликої кількості відбиттів: (а) всі відбиття зосереджені на початку інтервалу часу 0-50 мс; (б-в) сильні відбиття зосереджені в кінці часового інтервалу.

Цікаво, що для меншої кількості відбиттів значно більший розкид значень розбірливості мови, ніж для великої кількості відбиттів. Цей результат можна пояснити, використовуючи результати першої моделі (рис. 4.21). Дійсно, при невеликій кількості відбиттів можливі екстремальні ситуації, коли більшість сильних відбиттів зосереджені або на початку інтервалу 0-50 мс (рис. 4.23а), або в його кінці (рис. 4.23б). У першому випадку розбірливість висока, тоді як у

другому – низька. З великою кількістю відбиттів такі ситуації мало ймовірні, тому розкид значень розбірливості порівняно невеликий.

Зауважте, що надзвичайно високі та низькі значення розбірливості мови також можуть бути отримані за відносно рівномірного розподілу відбиттів у інтервалі 0-50 мс. У цьому випадку важливо місце розташування сильних відбиттів. Діаграма на рис. 4.23в показує випадок, коли відбиття розподілені відносно рівномірно, але сильні відбиття зосереджені в кінці інтервалу 0-50 мс, що призводить до зменшення розбірливості мови [4].

4.6 Акустична паспортизація навчальних приміщень

При роботі над проектом з акустичної паспортизації навчальних приміщень була поставлена мета розробити форми акустичного паспорта для акустичної експертизи вже наявних навчальних приміщень. При цьому ставилося завдання представлення інформації у формі, зручній як для адміністрації навчального закладу, так і для архітекторів, що займаються реконструкцією приміщень.

Акустичну паспортизацію приміщень слід здійснювати як на етапі проектування нових приміщень, так і на етапі реконструкції існуючих приміщень [118]. У першому випадку акустичний паспорт являє собою зведення результатів акустичного моделювання, розрахунків та рекомендацій щодо коригування технічного і дизайн-проектів з урахуванням вимог акустичного проектування. У другому випадку це зведення результатів вимірювань акустичних характеристик наявних приміщень і рекомендацій щодо коригування цих характеристик.

В обох випадках ключовим є перелік характеристик, що визначають кінцеві, тобто «споживчі» властивості приміщення. Очевидно, склад таких переліків в першу чергу визначається функціональним призначенням приміщень. Так, наприклад, при паспортизації приміщень для прослуховування музики загальна кількість оцінюваних характеристик і параметрів приміщення

може досягати двох десятків, хоча в кінцевому рахунку вдається скоротити цей перелік до чотирьох-п'яти базових показників [119], [120], [101].

Перелік характеристик і параметрів приміщень, призначених для прослуховування мови, також може бути зроблений досить компактним, що робить його зручним для огляду та аналізу. Акустичну паспортизацію таких приміщень проводять, прагнучи вирішити такі завдання:

- забезпечення максимальної розбірливості мови для збереження інформаційної складової переданого повідомлення;
- збереження природного тембру голосу для збереження естетичного компонента;
- забезпечення правильної локалізації звукового образу (суміщення його із зоровим образом) для підвищення фразової (смісловий) розбірливості мови.

Перелік оцінюваних параметрів.

Дисертант запропонував до складу акустичного паспорта ввести параметри, однаково інформативні як для адміністрації навчальних закладів, так і для архітекторів і дизайнерів. Це параметри, що характеризують рівень шуму приміщення, розбірливість мови і якість мовленнєвого сигналу [3].

При оцінюванні рівня шуму в навчальному приміщенні, доцільно оцінювати як середній, так і максимальний рівні шуму. Очевидно, такі вимірювання доцільно проводити під час навчальних занять, оскільки в перервах рівень шуму може значно зростати, що є типовим та природнім для навчальних закладів.

Для судження про розбірливість мови в приміщенні в роботі [3] оцінювалися показники у вигляді коефіцієнтів чіткості (ясності) C_{50} і коефіцієнту складової розбірливості.

Доцільність оцінювання показника C_{80} може бути обґрунтована тим, що досліджувані приміщення належать кафедрі акустичних та мультимедійних електронних систем, де питанням комп'ютерної обробки музики приділяється значна увага.

Мірою якості мовленнєвого сигналу служив коефіцієнт низького тону.

Крім зазначених вище показників, оцінювалася також залежність часу реверберації від частоти. Корисність такої залежності в тому, що вона дозволяє архітекторам і дизайнерам скласти більш повне уявлення про вплив конструктивних особливостей приміщень на акустичні характеристики.

Організація експериментальних досліджень.

Вимірювання шуму в навчальних приміщеннях проводилися з використанням цифрового шумоміра SL-814.

Оцінювання часу реверберації здійснювалося відповідно до рекомендацій ISO 3382-2 «Вимірювання акустичних параметрів в приміщеннях» [71]. При цьому використовувався імпульсний метод, згідно з яким на диск комп'ютера записувався відгук приміщення на імпульсний вплив у вигляді звуку вибуху гумової кульки. Отриманий запис потім піддавався обробці для оцінювання часу реверберації.

Запис звуку проводилася з використанням мікрофона МКФ МД-186 і зовнішньої звукової карти M-Audio FastTrackPro. Запис і обробка записаних сигналів здійснювалася в середовищі Matlab 2009b.

Дамо коротку характеристику навчальних приміщень, акустичні параметри і характеристики яких оцінювалися. Аудиторія 438 призначена для проведення лекцій та практичних занять, 224 і 229 – для проведення нарад та індивідуальних консультацій, 233, 212а – для лабораторних занять, 231, 212 – як для лабораторних, так і для практичних занять. Стіни приміщень оштукатурені і обклеєні шпалерами (224, 229) або пофарбовані. На підлозі аудиторій 212, 212а, 229 укладений паркет, в інших аудиторіях – лінолеум.

Результати експериментів.

Результати оцінювання основних акустичних параметрів навчальних приміщень представлені в таблиці 1. Як випливає з наведених даних, середній рівень шуму в усіх приміщеннях перевищує допустимий санітарними нормами

СРСР від 1984 р значення 40 дБА [121]. У приміщеннях 212а і 231 це перевищення є значним і становить 7-8 дБА. Зауважимо, однак, що ці приміщення відповідають санітарним нормам України від 1999 року, де нормальним для навчальних приміщень вважається рівень шуму 50 дБА [122].

Максимальний рівень шуму лише в двох приміщеннях перевищує норму [121]. При цьому в приміщенні 438 зафіксовано перевищення норми на 6 дБА.

Значення коефіцієнта чіткості C_{50} для всіх приміщень укладені в допустимих межах, забезпечуючи 80% складову і 95% фразову розбірливість [101]. Значення коефіцієнта C_{80} також знаходяться в дозволеному діапазоні.

Таблиця 4.2

Параметри	Оцінки параметрів							Норма
	212	212а	224	229	231	233	438	
Номер приміщення	212	212а	224	229	231	233	438	
Середній рівень шуму, дБА	44	48	42	43	47	44	44	$\leq 40/50^*$
Макс. рівень шуму, дБА	48,6	51	47	47,5	55,1	48,6	60,7	$\leq 55/65^*$
Час приходу першого відбиття, с	0,01	0,027	0,012	0,02	0,01	0,017	0,015	$\leq 0,05$
C_{50} , дБ	4,59	3,68	1,57	-0,95	-0,82	4,40	-0,95	≥ -2
C_{80} , дБ	8,38	7,02	4,49	1,81	2,39	8,32	1,81	≥ -3
k_c	0,76	0,79	0,84	0,89	0,89	0,77	0,89	$> 0,8$
k_{HT}	0,84	0,90	0,81	0,90	0,87	0,98	0,90	0,9-1/1-1,3**

* Числа перед символом «/» означають норму для лекційних приміщень, а числа після символу «/» означають норму для лабораторних і комп'ютерних приміщень [121], [122]

** числа перед символом «/» означають норму для мови, а числа після символу «/» означають норму для музики

Разом з тим, значення коефіцієнта розбірливості мови в приміщеннях 212 і 233 виявилися на 4-5% нижче нормативного значення, що свідчить про дещо підвищеному рівні дифузних відображень звуку.

Судячи зі значень коефіцієнта низького тону, якість мовленнєвого сигналу є найбільш низькою в приміщеннях 212 і 224. Решта приміщень можна визнати задовільними в плані звучання мовленнєвих повідомлень, і лише одна лабораторія 233 може бути визнана однаково придатною як для мови, так і для музики.

Результати побудови залежностей часу реверберації від частоти представлені на рис. 4.24.

Представлені результати дозволяють скласти досить компактний опис акустичних властивостей обстежених приміщень, а також свідчать, що деякі з обстежених приміщень мають досить низькі акустичні характеристики.

Так, наприклад, рівень шуму в декількох аудиторіях помітно, на 7-8 дБ перевищує санітарні норми СРСР, хоча за санітарними нормами України шумність цих приміщень знаходиться в межах норми.

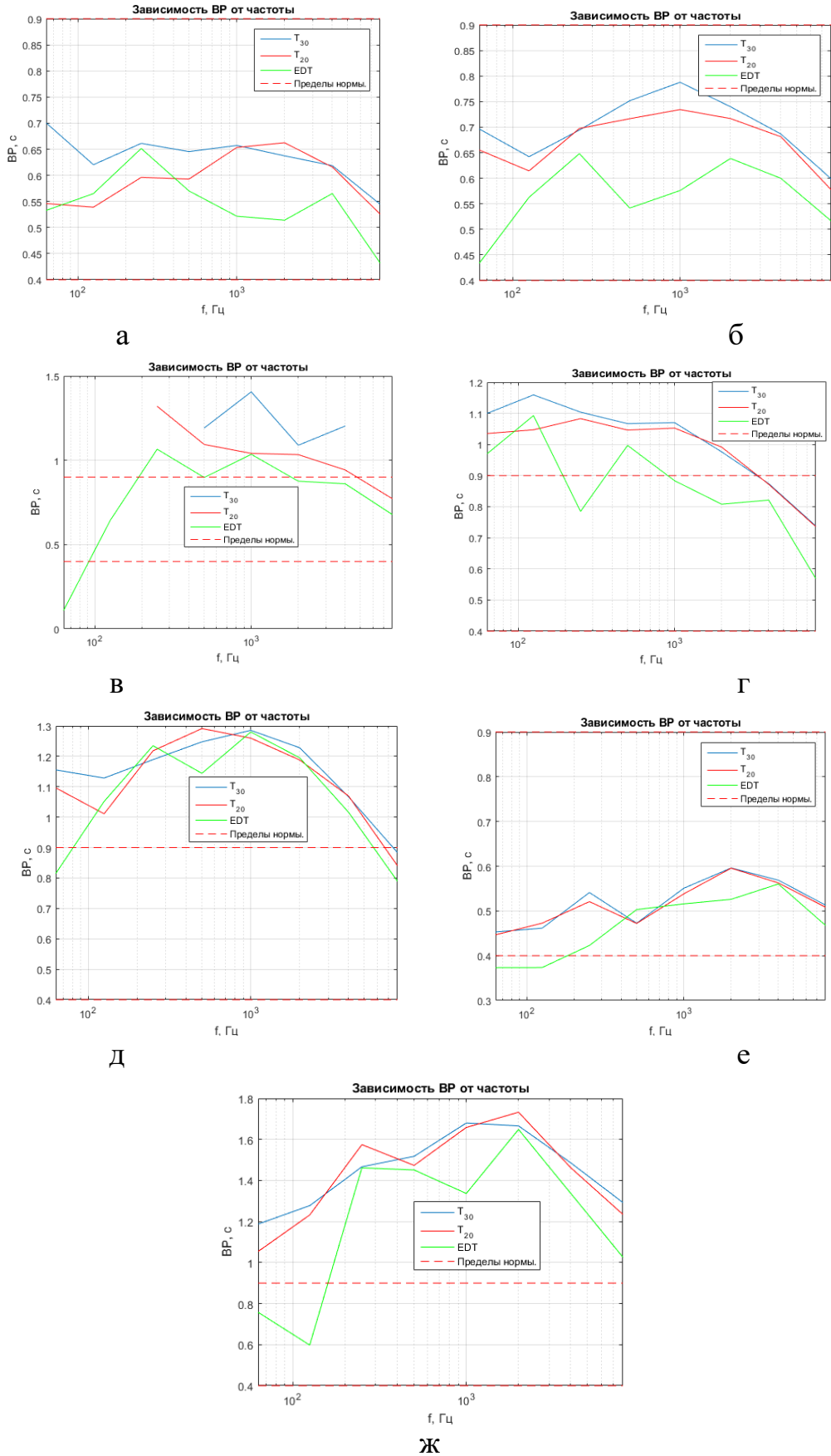


Рис. 4.24 – Залежність оцінок часу реверберації від частоти для приміщень 212 (а), 212а (б), 224 (в), 229 (г), 231 (д), 233 (е), 438 (ж)

Отримані оцінки розбірливості мови на основі лише ревербераційних властивостей приміщень можуть розглядатися лише як попередні, оскільки розбірливість мови залежить від спільної дії шуму і реверберації. Тому в подальшому доцільно використовувати показник у вигляді відношення корисної і шкідливої складових звуку [11]

$$U_{50} = 10 \lg \frac{E_{50}}{L_{50} + N},$$

де N – рівень фонового шуму. Індекс розбірливості мови STI може бути наближено обчислений через U_{50} на основі співвідношення [123]

$$U_{50} \cong 31 \cdot STI - 16.$$

Приклад акустичного паспорту приміщення приведено у додатку Г.

4.7. Автоматизоване оцінювання розбірливості мовлення у приміщенні в залежності від розташування слухача

Представлена в [77] та [78] база даних, що містить записи бінауральних імпульсних характеристик кількох приміщень різного призначення, на жаль, є досить обмеженою. Зокрема, вона містить записи бінауральних ІХ, зроблених із використанням штучної голови, лише для одного лекційного приміщення середнього розміру (приблизно 370 м³). Тому використання зазначених записів не дозволяє зробити певне узагальнення акустичних властивостей лекційних приміщень й таким чином зробити ще один крок в напрямі паспортизації приміщень навчальних закладів [3]. Крім цього, є низка інших задач, розв'язання яких потребує наявності бази бінауральних ІХ. Серед таких задач, зокрема, дослідження просторових властивостей слуху [16], оцінювання розбірливості мови [6, 5], дослідження алгоритмів дереверберації [80]. Оскільки під дереверберацією розуміють часткову нейтралізацію дії так званих пізніх відбиттів, неминуче виникає питання щодо корисності ранніх (до 50-80 мс) відбиттів. Таку корисність доведено, зокрема, в [25] та [22], хоча результати

подальших досліджень показали, що цінність ранніх відбиттів не слід перевищувати [23].

Після обчислення бінауральної ІХ приміщення, стає можливим оцінювання розбірливості мови модуляційним [124], [115] або формантно-модуляційним [125], [126] методом. Першим кроком такого оцінювання є обчислення, за формулою Шредера [127], так званих коефіцієнтів передачі модуляції:

$$m_{ki} = \frac{\left| \int_0^{\infty} h_k^2(t) \exp(-j2\pi F_i t) dt \right|}{\int_0^{\infty} h_{ki}^2(t) dt}, \quad (4.20)$$

де $h_k(t)$ – результат фільтрації функції $h_r(t)$ смуговим фільтром (на практиці використовують 7 октавних фільтрів із центральними частотами від 125 Гц до 8 кГц); F_i – частота модуляції (на практиці використовують 14 значень F_i , в діапазоні від 0,63 Гц до 12,5 Гц).

Наступними кроками є обчислення ефективних відношень сигнал-шум для кожної смуги та кожної модуляційної частоти:

$$E_{ki} = 10 \lg \frac{m_{ki}}{1 - m_{ki}}, \quad (4.21)$$

середнього індексу передачі модуляції

$$MT_k = \frac{1}{14} \sum_{i=1}^{14} T_{ki}, \quad (4.22)$$

$$T_{ki} = \begin{cases} 0, & E_{ki} < -15; \\ (E_{ki} + 15)/30, & -15 \leq E_{ki} \leq +15; \\ 1, & E_{ki} > +15; \end{cases}$$

Останнім кроком, згідно модуляційного методу [115], є обчислення індексу передачі мови:

$$STI = \sum_{k=1}^7 \alpha_k \cdot MT_k - \sum_{k=1}^6 \beta_k \cdot \sqrt{MT_k \cdot MT_{k+1}}, \quad (4.23)$$

де α_k – вагові коефіцієнти, β_k – коефіцієнти надмірності, значення яких для октавних смуг із центральними частотами f_0 наведено в табл. 4.2.

Таблиця 4.2 – Вагові коефіцієнти та коефіцієнти надмірності для октавних смуг

f_{0k} , Гц	125	250	500	1000	2000	4000	8000
α_k	0,085	0,127	0,230	0,233	0,309	0,224	0,173
β_k	0,085	0,078	0,065	0,011	0,047	0,095	-

Згідно формантно-модуляційного методу [125], на останньому кроці обчислюють словесну розбірливість:

$$W = \begin{cases} 1,54 \cdot A^{0,25} [1 - \exp(-11 \cdot A)], & A < 0,15; \\ 1 - \exp\left(\frac{11 \cdot A}{1 + 0,7 \cdot A}\right), & A \geq 0,15, \end{cases} \quad (4.24)$$

$$A = \sum_{k=1}^7 p_k \cdot P_k(E_k), \quad (4.25)$$

Де A – формантна розбірливість; p_k – імовірність знаходження формант в k -й смузі частот; $P_k(E_k)$ – коефіцієнт сприйняття мови. В працях Н. Б. Покровського [128] та М. А. Сапожкова [129] коефіцієнти сприйняття відрізняються та вважаються незалежними від смуги частот: $P_k(E_k) = P(E_k)$. В [102] викладено результати досліджень причин такої ситуації та показано доцільність використання уточнених коефіцієнтів сприйняття, форма яких залежить від смуги частот.

Організація досліджень.

На відміну від німецького вимірювального комплексу, де використовувалася професійна аудіоапаратура [77], розроблений комплекс «Штучна голова» [8] містить електроакустичну апаратуру різної якості, а саме: побутову активну колонку Genius SP-HF 2.0 500 (14 Вт, 65 Гц - 20 кГц, сигнал-шум 73 дБ), всеспрямовані вимірювальні конденсаторні мікрофони Superlux

ЕСМ-999 (20 Гц – 20 кГц, динамічний діапазон 106 дБ, сигнал-шум 70 дБ), зовнішню звукову карту Steinberg UR242 (24 біт, 192 кГц).

Запис сигналів виконувався із частотою дискретизації 44,1 кГц та глибиною квантування 24 біти. Два мікрофони розташовувалися в зонах вушних раковин штучної голови на відстані 1 см від місцезнаходження вушного каналу.

Основою тестового сигналу слугувала mls-послідовність із ступенем 16 (тобто mls-послідовність містила 2^{16} вибірок), що відповідає довжині сигналу 1,49 с при частоті дискретизації 44,1 кГц. Ця mls-послідовність при випромінненні повторювалася 17 разів, що дозволяло при обчисленні ІХ приміщення усереднювати 16 останніх сплесків оцінки ВКФ для підвищення відношення сигнал-шум на 12 дБ.

Для вимірювань було обрано 2 аудиторії, малого та середнього розміру (рис. 4.25).

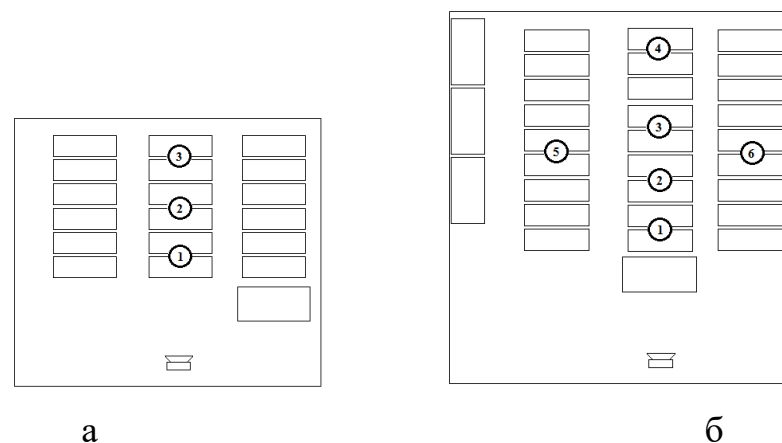


Рис. 4.25 – Плани малої аудиторії 438 (а) та середньої аудиторії 209 (б)

Характеристики малої аудиторії №438: розміри 7х7х3,1 м, два вікна, двері, задня стінка приміщення зашклена, три ряди із 6 парт в кожному ряді та стіл викладача, наповненість приміщення під час запису сигналів - 11 студентів та викладач. Дистанції від гучномовця до номерів 1, 2 та 3 точок, де розміщувалася ШГ, 2 м, 4 м та 6 м, відповідно.

Характеристики середньої аудиторії №209: розміри 10х15х3,1 м, чотири вікна, двері, дві книжних шафи, шафа-гардероб, три ряди із 9 парт в кожному

ряді, стіл викладача та 2 додаткових вільних столи, наповненість приміщення під час запису сигналів - 9 студентів та викладач. Дистанції від гучномовця до номерів 1, 2, 3, 4, 5 та 6 точок, де розміщувалася ШГ, 3 м, 6 м, 9 м, 14 м, 9 м та 9 м, відповідно.

Результати досліджень.

На рис. 4.26 наведено графіки ІХ розглянутих приміщень, із результатами вимірювання часу реверберації. Графіки наведено лише для лівого каналу, оскільки для правого каналу вони є подібними. Обвідну ІХ, побудовану за методом Шредера [55], використано для обчислення часу реверберації, оцінки якого представлені на рис. 4.26.

Порівнюючи ІХ аудиторій 438 (рис. 4.26а) та 209 (рис. 4.26б), бачимо, що в аудиторії 209 спостерігається біля 3-х потужних ранніх відбиттів на інтервалі часу до 30 мс, та ще один виразний сплеск в момент часу 0,27 с, тоді як спад енергії в аудиторії 438 є більш плавним. Час реверберації в обох приміщеннях виявився досить великим та близьким до 1 с, що у випадку аудиторії 209 пояснюється слабким заповненням порівняно великого приміщення, а у випадку аудиторії 438 пояснюється наявністю скляної задньої стінки.

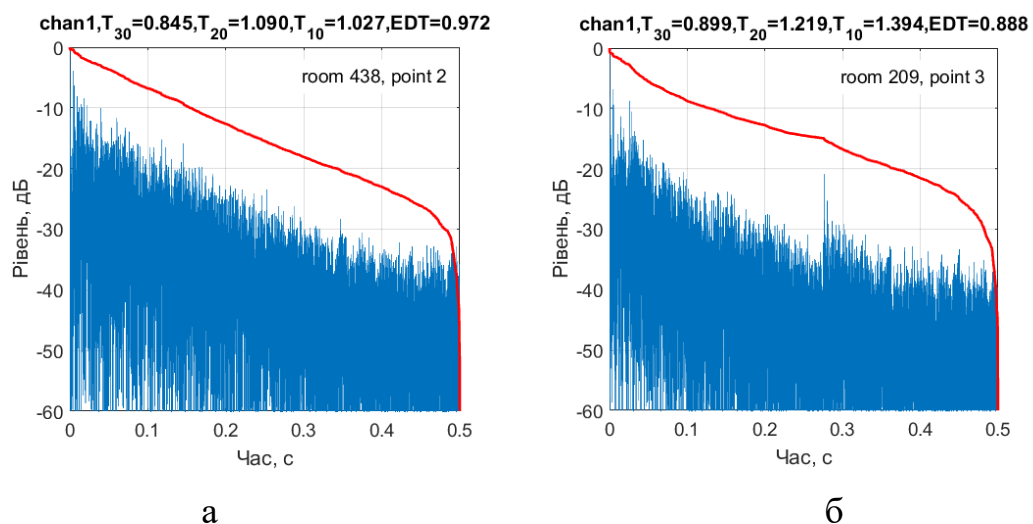


Рис. 4.26 – ІХ та час реверберації для аудиторій 438 (а) та 209 (б)

Результати оцінювання характеристик $C_{50} = 10\lg(E_0^{50}/E_{50}^\infty)$ та $ERB = 10\lg(E_0^{50}/E_0^{10})$ наведено на рис. 4.27 та 4.28, відповідно, де $E_a^b = \int_a^b h^2(t)dt$, a та b – час в мілісекундах. Оскільки для $t > 0,5$ с рівень шумової завади переважає над рівнем реверберації (рис. 4.26), при обчисленнях C_{50} в якості ∞ використано значення 0,5 с.

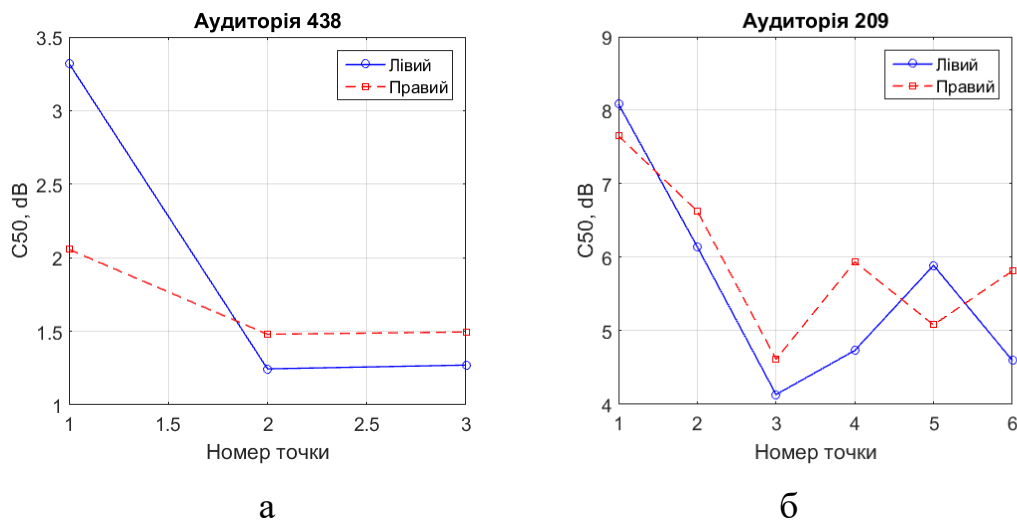


Рис. 4.27 – Оцінки параметру C_{50} для аудиторій 438 (а) та 209 (б)

Згідно отриманим оцінкам C_{50} , чіткість мовлення в аудиторії 209 має бути помітно вищою, ніж в аудиторії 438, причому ця чіткість має швидко спадати із наближенням до середини кімнати, після чого має стабілізуватися на певному рівні. Цікаво, що, згідно оцінкам C_{50} , в аудиторії 209 чіткість мови біля задньої стінки (точка 4) є помітно вищою, ніж в середині кімнати (точка 3). В аудиторії 438 подібний ефект також спостерігається, але є ледь помітним.

Поведінка показника ERB, що характеризує співвідношення енергій ранніх відбиттів та прямого звуку, також підтверджує існування зазначеного ефекту. Дійсно, якщо вважати, що прямий звук є кращим з точки зору розбірливості мови, порівняно із ранніми відбиттями, тоді, згідно рис. 4.28а, бачимо, що в аудиторії 438 розбірливість в лівому каналі є вищою біля задньої стінки (точка 3), порівняно із серединою приміщення (точка 2).

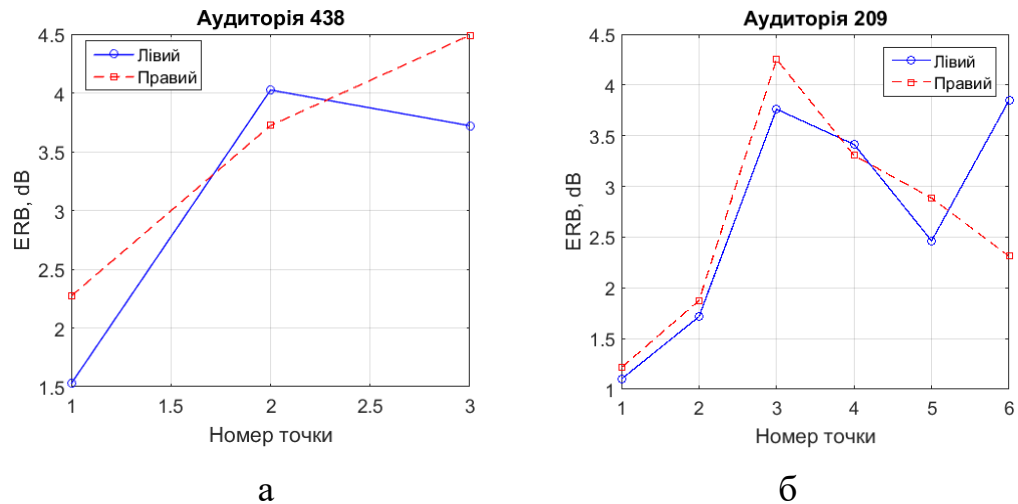


Рис. 4.28 – Оцінки параметру ERB для аудиторій 438 (а) та 209 (б)

Зважаючи на певну неоднозначність отриманих оцінок C50 та ERB, які є непрямими мірами розбірливості мови, цікаво оцінити розбірливість мови у вказаних аудиторіях модуляційним та формантно-модуляційним методами із використанням (2.5)-(2.7).

Результати оцінювання розбірливості мови модуляційним методом представлено на рис. 4.29, а формантно-модуляційним методом – на рис. 4.30, при цьому оцінки, отримані формантно-модуляційним методом, були розраховані із використанням уточнених коефіцієнтів сприйняття, форма яких залежить від смуги частот [102].

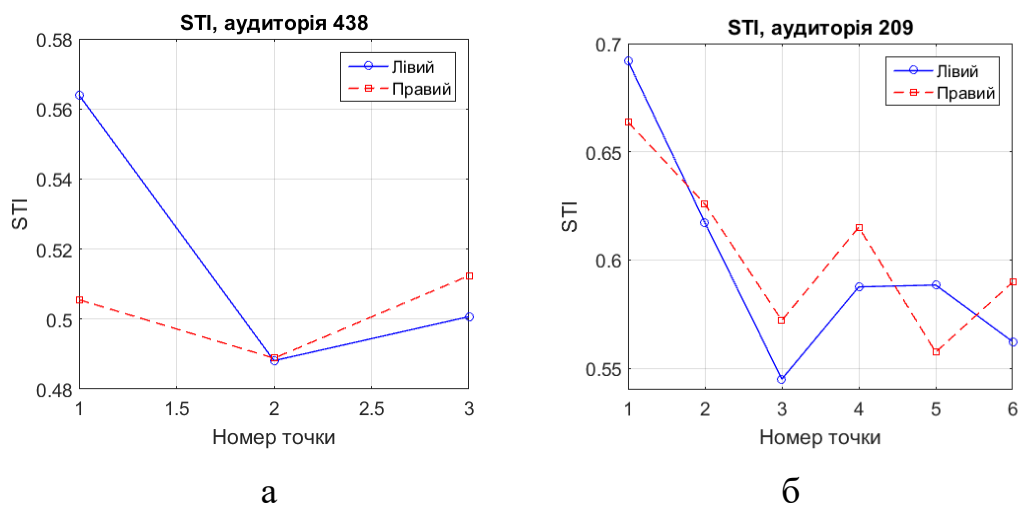


Рис. 4.29 – Оцінки індексу STI для аудиторій 438 (а) та 209 (б)

Неважко бачити хороше узгодження оцінок розбірливості мови, отриманих модуляційним та формантно-модуляційним методами, а також узгодження їх із оцінками параметру C_{50} (рис. 4.27). Слід, однак, зазначити, що ефект підвищення розбірливості мови біля задньої стінки аудиторії 438, порівняно із серединою приміщення, є значно виразнішим для оцінок розбірливості мови (рис. 4.28а та 4.29а), ніж для коефіцієнту C_{50} (рис. 4.27а). Що стосується аудиторії 209, де було виконано вимірювання ще й біля бокових стінок (точки з номерами 5 та 6), бачимо, що біля бокових стінок аудиторії розбірливість мови є також вищою, ніж в центрі приміщення. Зазначимо, що ефект підвищення розбірливості мови біля стінок приміщення було виявлено також в [21] шляхом суб'єктивного оцінювання розбірливості мови.

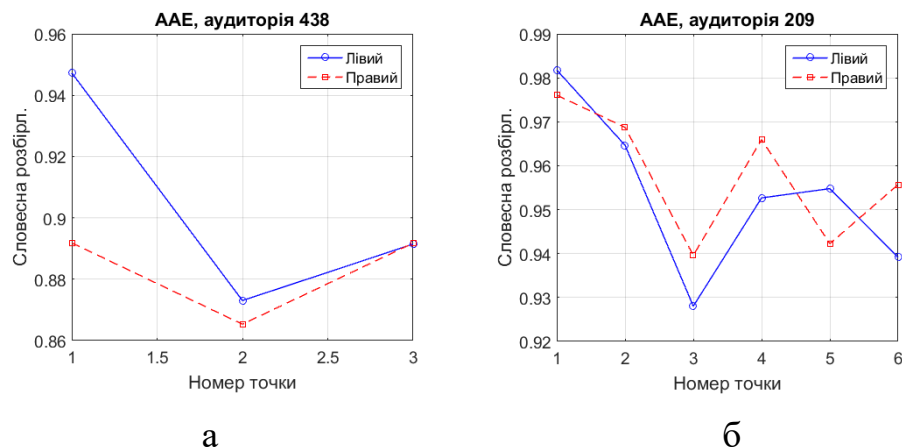


Рис. 4.30 – Оцінки словесної розбірливості для аудиторій 438 (а) та 209 (б)

Даному ефекту можна дати досить просте фізичне пояснення. Дійсно, до вух людини, що сидить біля стінки, майже одночасно (в наших випадках різниця в часі становила 3-5 мс) приходять прямий та відбитий від стінки звуку. При цьому потужність відбитого від стінки звуку є ненабагато меншою, порівняно із потужністю прямого сигналу. На низьких частотах (100-200 Гц) сигнали є практично когерентними, тому при близьких рівнях прямого та відбитого звуку на цих частотах можна очікувати підвищення рівня результуючого звуку до 6 дБ, а на високих частотах, де сигнали є практично некогерентними, рівень звуку може підвищитися до 3 дБ. Зауважимо, що хоча ці оцінки є дуже приблизними,

проте вони добре узгоджуються із результатами, наведеними в [25], [22]. Що стосується середини приміщення, відбитий від стінки звук приходиться із більшою затримкою відносно прямого звуку (в наших випадках 15-20 мс) та є послабленим за рахунок розповсюдження в просторі. Оскільки про когерентність прямого та відбитого звуків в даному випадку вже не йдеться, то очевидно, що рівень сумарного сигналу має бути меншим, ніж біля стінки. Крім того, слід вказати ще й на те, що через помітну затримку між прямим та відбитим сигналом кімната фактично перетворюється в гребінчастий фільтр [130] із суттєво нерівномірною амплітудно-частотною характеристикою. Таким чином, сумарний звук є ще й спотвореним, що також не може сприяти хорошій розбірливості мови в середині кімнати.

Можна сказати, що в аудиторіях, подібних розглянутим в даній роботі, за критерієм розбірливості мови користь ранніх відбиттів є різною в різних частинах приміщення. Біля стін приміщення вона є більшою, ніж в середині приміщення.

4.8 Спосіб виявлення кліпування мовленнєвого сигналу

В цій дисертації основними причинами падіння розбірливості мовленнєвого сигналу у приміщенні розглянуто фоновий шум і реверберацію. Однак при записі такого сигналу існує ризик виникнення нелінійних спотворень, зокрема, кліпування.

Максимальне використання динамічного діапазону каналу передачі мови є дуже бажаним, оскільки він забезпечує найвище значення SNR. Однак існує ризик нелінійних спотворень через кліпування сигналу, при якому великі миттєві значення сигналу замінюються деякими постійними значеннями.

У той же час невелике значення кліпування супроводжується невеликими нелінійними спотвореннями, які можуть бути прийнятними для слухачів. Тому доцільно побудувати алгоритм виявлення кліпування так, щоб йому передувала алгоритм оцінки значення кліпування.

Алейнік С. В., Матвеев Ю. Н. і Раєв А. Н. запропонували для виявлення кліпування використовувати параметр у вигляді нормованої дистанції між центральним та крайніми піками амплітудного розподілу [131]. Коефіцієнт кліпування дозволяє виявити сильне кліпування, але не його рівень. Недоліками цього способу є необхідність попереднього оцінювання кількох сотень значень амплітудного розподілу сигналу, а також нездатність до виявлення слабого кліпування. У той же час, позитивною властивістю цього параметра є скінченність області значень, що робить його зручним у використанні в інженерних цілях.

Автором дисертації у співавторстві було побудовано і запатентовано алгоритм виявлення кліпування [1]. Запропонований алгоритм забезпечує можливість виявлення не тільки сильного, але й слабого кліпування шляхом того, що замість кількох сотень значень амплітудного розподілу сигналу використовують лише один параметр амплітудного розподілу акустичного сигналу, при цьому попереднє оцінювання амплітудного розподілу сигналу є непотрібним.

Для виявлення кліпування у якості параметра, що порівнюють, використовують кумулятивний коефіцієнт статистичного розподілу миттєвих значень акустичного сигналу або функціональне перетворення кумулятивного коефіцієнта.

Кумулятивний коефіцієнт четвертого порядку обчислюють за формулою:

$$\gamma = \frac{\mu_4}{\mu_2^2} - 3,$$

де μ_k – центральний момент k -го порядку.

Можна використовувати такі функціональні перетворення як нормований центральний момент четвертого порядку:

$$\chi = \gamma + 3 = \frac{\mu_4}{\mu_2^2},$$

обернене значення нормованого центрального моменту четвертого порядку:

$$\beta = \frac{1}{\chi} = \frac{\mu_2^2}{\mu_4},$$

корінь квадратний від оберненого значення нормованого центрального моменту четвертого порядку:

$$\delta = \sqrt{\beta} = \frac{\mu_2}{\sqrt{\mu_4}}.$$

У роботі [132] були отримані карти відповідності між запропонованими об'єктивними мірами і суб'єктивними оцінками. Ці карти узгодження є монотонними функціями, які не містять різких зламів, що дозволяє зробити висновок, що запропоновані параметри придатні як міри якості кліпованого мовленнєвого сигналу.

Висновки до розділу 4

1. Розроблено апаратно-програмний комплекс для проведення автоматизованих артикуляційних випробувань, спрямованих на оцінювання розбірливості українського мовлення в різних заводових умовах.
2. З використанням розробленого апаратно-програмного комплексу отримано наступні результати:
 - зіставлено об'єктивні та суб'єктивні оцінки розбірливості мовлення за різних заводових умов (шум, рання та пізня реверберація), що дозволило перевірити достовірність результатів інструментального оцінювання акустичних параметрів приміщення;
 - підтверджено результати попередніх досліджень щодо помітного, на 5-7%, зростання розбірливості мовлення при діотичному прослуховуванні, порівняно із діхотичним.
 - порівняння впливу ранньої та пізньої реверберації на результати оцінювання розбірливості мовлення показало, що ранні відбиття можуть впливати на розбірливість мовлення не тільки позитивно, але й негативно, через погіршення якості мовленнєвих сигналів.
3. Показано, що поведінка об'єктивних показників якості мовленнєвого сигналу, такі як SSNR, LSD, fwSNR, BSD та PESQ. LSD і BSD частково узгоджується з поведінкою суб'єктивних оцінок розбірливості мовлення: показник PESQ добре узгоджується із суб'єктивними оцінками для $SNR > 0$ дБ, а fwSNR – для $SNR > 10$ дБ.
4. Артикуляційні випробування для бінаурального режиму прослуховування тестових таблиць показали, що приміщення можна розглядати як спеціальний фільтр. Такий фільтр при низьких значеннях SNR підвищує суб'єктивну оцінку розбірливості мовлення, а при помірних і високих значеннях SNR – дещо знижує її.
5. Дослідження суб'єктивної і об'єктивної оцінок розбірливості мовленнєвого сигналу як функції відстані до джерела свідчать про те, що

розбірливість мови є мінімальною посеред кімнати середнього розміру. Цей мінімум зберігається, навіть якщо вплив пізньої реверберації штучно придушується.

6. Запропоновано і запатентовано використання кумулятивного коефіцієнту четвертого порядку або його функціональних перетворень як параметру для виявлення кліпування сигналу.

ВИСНОВКИ

У дисертаційній роботі представлені результати дослідження акустичних параметрів приміщення як каналу мовленнєвої комунікації. Виконано огляд актуальних досліджень, розроблено програмне забезпечення, проведено комп'ютерне моделювання, проведено натурні експерименти та артикуляційні випробування, що дало змогу отримати такі результати:

1. Огляд сучасного стану задач вимірювання та моделювання імпульсної характеристики приміщення, прямого та сліпого методів оцінювання часу реверберації, впливу реверберації на якість та розбірливість мовлення дав змогу виявити прогалини у сучасній теорії та практиці акустичної експертизи приміщень. Зокрема, аналітичний огляд прямих та сліпих методів вимірювання часу реверберації дав змогу виявити границі їх застосування та обґрунтувати необхідність їх подальшого вдосконалення.

2. Вдосконалено алгоритми оцінювання часу реверберації прямим та сліпим методами, що дозволяє розширити границі застосування зазначених алгоритмів та підвищити ступінь автоматизації акустичних вимірювань.

3. Вперше вказано умови, за яких об'єктивне оцінювання якості мовленнєвих сигналів, що маскуються шумовою завадою, може бути застосовано замість суб'єктивного оцінювання розбірливості мовлення. Це дозволяє суттєво спростити акустичну експертизу приміщень, призначених для мовленнєвих презентацій.

4. Вперше вказано на можливість зниження розбірливості мови в окремих частинах приміщень через погіршення якості мовленнєвого сигналу, яке спричинене дією ранніх відбиттів звуку. Виявлення такої негативної дії ранніх відбиттів звуку дозволяє обґрунтувати рішення стосовно необхідності коригування акустики приміщення. Разом із тим, показано, що на невеликих (10-30 см) дистанціях від відбиваючих поверхонь ранні відбиття призводять до помітного, на 2-3%, підвищення індексу розбірливості мовлення. Тим самим обґрунтовано доцільність розширення переліку параметрів, що входять до

акустичного паспорту приміщення, шляхом включення до цього переліку оцінок розбірливості мови в різних точках приміщення.

5. Вперше запропоновано, при автоматизації оцінювання розбірливості українського мовлення артикуляційним методом, спосіб фіксації сприйнятих слухачем звукових сполучень, що полягає в наборі з клавіатури комп'ютера цих звукових сполучень типу "приголосний-голосний-приголосний", а також розроблено апаратно-програмну систему для реалізації такого оцінювання.

6. Показано, що вдосконалений алгоритм автоматичного оцінювання часу реверберації прямим методом може бути успішно використаний для акустичної паспортизації приміщень, призначених для мовленнєвих презентацій.

7. Виконано поглиблене дослідження питання акустичної паспортизації приміщень, запропоновано перелік акустичних параметрів для включення у відповідний паспорт, розроблено програмне забезпечення для акустичної паспортизації приміщень.

8. Розроблено рекомендації щодо коригування амплітудної частотної характеристики акустичної апаратури непрофесійного рівня якості при проведенні вимірювань імпульсних характеристик приміщення, а також при об'єктивному оцінюванні якості та розбірливості мовлення із використанням штучної голови.

9. Проведено низку артикуляційних випробувань із суб'єктивного оцінювання розбірливості мовлення, завдяки яким вдалось зіставити результати об'єктивного та суб'єктивного оцінювання розбірливості мовлення, дослідити вплив бінаурального сприйняття звуку, ранньої та пізньої реверберації на оцінки розбірливості мовлення.

Практичне значення отриманих результатів полягає у підвищенні достовірності результатів артикуляційних випробувань та у підвищенні ступеня автоматизації таких випробувань, спрямованих на суб'єктивну оцінку розбірливості українського мовлення в приміщеннях. Встановлення принципової можливості використання мір якості мовленнєвого сигналу

замість мір розбірливості мовлення дозволяє спростити процедуру оцінювання розбірливості мови. Розробка програмно-апаратної системи для артикуляційних випробувань каналів передачі української мови в умовах дії шумової завади та реверберації дозволяє автоматизувати такі випробування, підвищити їх достовірність та суттєво знизити ресурсомісткість випробувань. Вдосконалення алгоритмів оцінювання часу реверберації прямим та сліпим методами дозволяє розширити границі застосування зазначених алгоритмів та підвищити ступінь автоматизації акустичних вимірювань.

Результати, представлені у дисертації, можуть бути використані для експертизи та корекції акустичних характеристик приміщень шляхом розрахунку та вимірювання якості та розбірливості мовленнєвих сигналів у приміщеннях, що проектуються, будуються або піддаються реновації. Володіння інформацією щодо характеру та ступеня впливу шуму, ранніх відбиттів звуку та пізньої реверберації на якість та розбірливість мовлення в різних точках приміщень дозволить підвищити точність акустичної експертизи. Автоматизація вимірювань акустичних характеристик приміщення, в тому числі якості та розбірливості мовлення, дозволить підвищити точність вимірювань, знизивши при цьому витрати матеріальних, часових, фінансових та людських ресурсів.

Викладені у дисертації нові теоретичні та практичні результати досліджень знайшли застосування у освітньому процесі кафедри акустичних та мультимедійних електронних систем за спеціальністю 171 Електроніка, освітньою програмою “Акустичні електронні системи та технології обробки акустичної інформації”, Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”.

Список літератури

- [1] А. М. Продеус, І. В. Котвицький, М. В. Дідковська, В. С. Дідковський, К. А. Кухарічева, Д. Є. Моторнюк та О. О. Дворник, «Спосіб виявлення кліпування мовного та музичного сигналів». Україна Патент UA 144291 U, МПК G01R 23/20, 25 вересня 2020.
- [2] A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, D. Motorniuk and A. Khrapachevskyi, "Objective and Subjective Assessment of the Quality and Intelligibility of Noised Speech," in *Proc. of 2018 IEEE 5th Int. Sc.-Pract. Conf. "Problems of Infocommunications. Science and Technology" (PIC S&T`2018)*, Kharkiv, pp. 71-74. 9-12 October, 2018.
- [3] Ю. С. Костючок, Л. С. Мартинович, Д. Е. Моторнюк, В. А. Нечитайло, А. В. Храпачевский и А. Н. Продеус, «Акустическая паспортизация учебных помещений,» *Электроника и связь*, т. 21, № 2(91), pp. 63-70, 2016.
- [4] A. Prodeus, M. Didkovska, K. Kukharicheva and D. Motorniuk, "Modeling the Influence of Early Sound Reflections on Speech Intelligibility," in *Proc. of 2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC)*, Kyiv, 20-23 Oct. 2020. DOI: 10.1109/MSNMC50359.2020.9255657\.
- [5] A. Prodeus, M. Didkovska, K. Kukharicheva and D. Motorniuk, "Two Simplified Models of Early Sound Reflections in a Room," *Electronics and Control Systems*, vol. 3, no. 65, pp. 79-87, 2020.
- [6] A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi and D. Motorniuk, "Automated Subjective Assessment of Speech Intelligibility under Diotic and Dichotic Listening," in *Lecture Notes on Data Engineering and Communications Technologies, 2021, 48*, Cham, Springer, January 2021, pp. 1-16.

- [7] A. Prodeus, M. Didkovska, D. Motorniuk and O. Dvornyk, "The Effects of Noise, Early and Late Reflections on Speech Intelligibility," in *Proc. 2020 IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO)*, Kyiv, pp. 488 – 492. 22-24 April 2020. DOI: 10.1109/ELNANO50318.2020.9088854.
- [8] О. О. Dvornyk, D. I. Motorniuk, M. V. Didkovska та А. М. Prodeus, «Апаратно-програмний комплекс «Штучна голова». Частина 1 Коригування частотної характеристики тракту.» *Мікросистеми, Електроніка та Акустика*, т. 25, № 1, р. 56–64, 2020.
- [9] О. О. Dvornyk, D. I. Motorniuk, M. V. Didkovska та А. М. Prodeus, «Апаратно-програмний комплекс «Штучна голова». Частина 2. Оцінювання розбірливості мови в аудиторіях.» *Мікросистеми, Електроніка та Акустика*, т. 25, № 3, р. 48–55, 2020.
- [10] W. Yang and J. Bradley, "Effects of Room Acoustics on the Intelligibility of Speech in Classrooms for Young Children," *J. Acoust. Soc. Am*, vol. 125, no. 2, p. 922–933, 2009.
- [11] J. Bradley, R. Reich and S. Norcross, "On the Combined Effects of Signal-to-Noise Ratio and Room Acoustics on Speech Intelligibility," *J. Acoust. Soc. Am*, vol. 106, no. 4, Pt.1, pp. 820-1828, 1999.
- [12] J. Bradley and H. Sato, "Speech Intelligibility Test Results for Grades 1, 3 and 6 Children in Real Classrooms," in *ICA*, Kyoto, 2004.
- [13] K. Eggenschwiler, "Lecture Halls - Room Acoustics and Sound Reinforcement," in *Forum Acusticum*, Budapest, pp. 1-6, 2005.
- [14] C. D. Marino, D. Fuerth, D. Gignac, A. Lunardi, C. Novak, R. P. Pikul and A. Simone, "Acoustic Enhancement of Proposed Grand Lecture Hall using Computer Simulation," *Canadian Acoustics*, vol. 39, no. 1, pp. 43-48, 2011.

- [15] R. Youssef, D. Bard, F. Mahmoud and N. Esa, "Acoustical Quality Assessment of Lecture halls at Lund University, Sweden," in *Inter-noise*, Melburn, pp. 934-943. 6-19 November, 2014.
- [16] J. Blauert, *The Technology of Binaural Listening*, Berlin, Heidelberg, New York: Springer-Verlag, 2013.
- [17] S. Naida and O. Pavlenko, "Coupled Circuits Model in Objective Audiometry," in *2018 IEEE 38th Int. Conf. on Electronics and Nanotechnology (ELNANO)*, Kyiv, pp. 281-286. 24-26 April, 2018.
- [18] W. Jiang and H. Schulzrinne, "Recognition Performance as an Effective Perceived Quality Predictor," in *IEEE Int. Workshop on Quality of Service*, pp. 269-275, 2002.
- [19] S. van Wijngaarden, J. Verhave and H. Steeneken, *The Speech Transmission Index after four decades of development*, 2012.
- [20] S. Naida and O. Pavlenko, "Newborn Hearing Screening Based on the Formula for the Middle Ear Norm Parameter," in *2018 IEEE 38th Int. Conf. on Electronics and Nanotechnology (ELNANO)*, Kyiv, pp. 287-291. 24-26 April, 2018.
- [21] H. Sato and J. Bradley, "Evaluation of Acoustical Conditions for Speech Communication in Working Elementary School Classrooms," *Acoust. Soc. Am*, vol. 106, no. 4, Pt.1, p. 2064–2077, 2004.
- [22] J. Bradley, H. Sato and M. Picard, "On the importance of early reflections for speech in rooms," *J. Acoust. Soc. Am*, vol. 113, no. 6, pp. 3233-3244, 2003.
- [23] I. Arweiler, J. Buchholz and T. Dau, "Speech intelligibility enhancement by early reflections," in *ISAAR*, Marienlyst, Denmark, August 2009.
- [24] E. A. P. Habets, *Single- and multi-microphone speech dereverberation using spectral enhancement*, Ph.D. dissertation. Technische Universiteit Eindhoven, January, 2007.

- [25] G. Souloudre, N. Popplewell and J. Bradley, "Combined effects of early reflections and background noise on speech intelligibility," *Journal of Sound and Vibration*, vol. 135, no. 1, pp. 123-133, 1989.
- [26] J. Burger and J. Lochner, "The influence of reflections on auditorium acoustics," *J. Sound Vib*, vol. 1, pp. 426-454, 1964.
- [27] J. Bradley, "Speech intelligibility studies in classrooms," *J. Acoust. Soc. Am*, vol. 80, p. 846–854, 1986.
- [28] J. Bradley, "Predictors of speech intelligibility in rooms," *J. Acoust. Soc. Am*, vol. 80, p. 837–845, 1986.
- [29] A. Prodeus, K. Bukhta, P. Morozko, O. Serhiienko, I. Kotvytskyi and I. Shherbenko, "Automated System for Subjective Evaluation of the Ukrainian Speech Intelligibility," in *Proc. of IEEE 38th Int. Conf. on Electronics and Nanotechnology (ELNANO)*, Kyiv, pp. 533-538. 24-26 April, 2018.
- [30] H. Lopacz and R. Marczuk, "A New Approach to the Room Impulse Response Simulation," *Archives of Acoustics*, vol. 29, no. 1, pp. 61-85, 2004.
- [31] P. Malecki, "Spatial Impulse Response Assessment in Room Acoustics Auralization," *Acta Physica Polonica A*, vol. 128, no. 1-A, pp. 17-21, 2015.
- [32] D. Berkley and J. Allen, "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943-950, April 1979.
- [33] A. Kulowski, "Algorithmic representation of the ray tracing technique," *Applied Acoustics*, vol. 18, no. 6, pp. 449-469, 1985.
- [34] W. G. Gardner, "Reverberation Algorithms," in *Application of digital signal processing to audio and acoustics*, Norwell, MA, Kluwer Academic Publishers, 1998, pp. 85-132.
- [35] H. Kuttruff, *Room Acoustics*, 4th ed., London EC4P 4EE: Spon Press, 11 New Fetter Lane, 2000.

- [36] S. V. A. Gari, Investigations on the Influence of Acoustics on Live Music Performance using Virtual Acoustic Methods, Ph.D. dissertation, Germany: Detmold University of Music, 2017.
- [37] H. Haas, "Uber den Einfluss eines Einfachechos auf die Harsamkeit von Sprache," *Acustica*, no. 1, pp. 49-58, 1951.
- [38] Ш. Я. Вахитов, Ю. А. Ковалгин, А. А. Фадеев та Ю. П. Щевьев, Акустика: учебник для вузов, Москва: Горячая линия – Телеком, 2009. – 660 с..
- [39] F. Aigner and M. J. O. Strutt, "On a physiological effect of several sources of sound on the ear and its consequences in architectural acoustics," *Journal of the Acoustical Society of America*, vol. 6, no. 3, pp. 155-159, 1935.
- [40] V. M. A. Peutz, "Articulation loss of consonants as a criterion for speech transmission in a room," *Journal of the Audio Engineering Society*, vol. 19, no. 11, pp. 915-919, December 1971.
- [41] D. A. Berkley, "Normal Listeners in Typical Rooms - Reverberation Perception, Simulation, and Reduction," in *Acoustical factors affecting hearing aid performance*, Baltimore, University Park Press, 1980, p. 3–24.
- [42] J. B. Allen, "Effects of small room reverberation on subjective preference," *Journal of the Acoustical Society of America*, vol. 71, no. S1, p. S5, 1982.
- [43] R. H. Bolt and A. D. MacDonald, "Theory of speech masking by reverberation," *Acoustical Society of America*, vol. 21, p. 577–580, 1949.
- [44] A. K. Nabelek, T. R. Letowski and F. M. Tucker, "Reverberant overlap- and self-masking in consonant identification," *Journal of the Acoustical Society of America*, vol. 86, no. 4, p. 1259–1265, 1989.
- [45] А. Продеус, В. Дидковский и В. Овсяник, «Слепое измерение времени реверберации в системах автоматического распознавания речи,» *Системи обробки інформації*, т. 7, № 123, pp. 59-66, 2014.
- [46] P. Naylor and N. Gaubitch, *Speech Dereverberation*, London: Springer-Verlag, 2010.

- [47] В. С. Дідковський, М. В. Дідковська та А. М. Продеус, Акустична техніка. Т. 12. Комп'ютерна обробка акустичних сигналів. Навчальний посібник, Київ, 2010. – 430 с.
- [48] J. Steinberg, "Effects of distortion upon the recognition of speech sounds," *Journal of the Acoustical Society of America*, vol. 1, pp. 121-137, 1929.
- [49] M. Hodgson and E. Nosal, "Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms," *Journal of the Acoustical Society of America*, vol. 111, pp. 931-939, 2002.
- [50] J. Benesty, S. Makino and J. Chen, *Speech Enhancement*, Germany: Springer, 2005.
- [51] K. Lebart, J. M. Boucher and P. N. Denbigh, "A new method based on spectral subtraction for speech dereverberation," *Acta Acoustica*, vol. 87, no. 3, p. 359–366, 2001.
- [52] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, Vols. ASSP-32, no. 6, pp. 1109-1121, December 1984.
- [53] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Log- Spectral Amplitude Estimator," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, Vols. ASSP-33, no. 2, pp. 443-445, April 1985.
- [54] I. Cohen, "Optimal Speech Enhancement Under Signal Presence Uncertainty Using Log-Spectral Amplitude Estimator," *IEEE Signal Processing Lett*, vol. 9, no. 4, pp. 113-116, April 2002.
- [55] M. Schroeder, "New Method of Measuring Reverberation Time," *JASA*, vol. 37, no. 0, pp. 409-412, 1965.
- [56] ГОСТ 24146-89 Зрительные залы. Метод измерения времени реверберации.

- [57] Акустика, ISO 3382-2-2008. Измерение акустических параметров в помещениях. Часть 2. Время реверберации в обычных помещениях.
- [58] C. Zhixin, "Investigation on Simulation and Measurement of Reverberation for Small Rooms," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 2, no. 7, pp. 69-73, 2011.
- [59] J. B. Allen and S. T. Neely, "Invertibility of room impulse response," *J. Acoust. Soc. Am*, vol. 66, p. 165–169, 1979.
- [60] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech, Signal Process*, vol. 36, no. 2, p. 145–152, 1988.
- [61] F. Fricke and J. Nannariello, "The prediction of reverberation time using neural network analysis," *Appl. Acoust*, vol. 58, p. 305–325, 1999.
- [62] T. J. Cox, F. Li and P. Darlington, "Extracting room reverberation time from speech using artificial neural networks," *Audio Eng. Soc*, vol. 49, p. 219–230, 2001.
- [63] R. Ratnam, D. L. Jones, B. C. Wheeler and a. al., "Blind estimation of reverberation time," *J. Acoust. Soc. Am*, vol. 114, no. 5, pp. 2877-2892, 2003.
- [64] N. Gaubitch, H. W. Lollmann, M. Jeub and et.al, "Performance comparison of algorithms for blind reverberation time estimation from speech," in *IWAENC 2012*, pp. 1-4. 4-6 September, 2012.
- [65] J. Y. C. Wen, E. A. P. Habets and P. A. Naylor, "Blind estimation of reverberation time based on the distribution of signal decay rates," in *IEEE Intl. Conf. on Acoust., Speech, Signal Process. (ICASSP)*, Las Vegas, April 2008.
- [66] T. Falk, C. Zheng and W.-Y. Chan, "A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech," *IEEE Trans. Audio, Speech, Lang. Process*, vol. 18, no. 7, p. 1766–1774, September 2010.
- [67] V. Poor, *An Introduction to Signal Detection and Estimation*, New York: Springer-Verlag, 1994.

- [68] H. Lollmann and P. Vary, "Estimation of the reverberation time in noisy environments," in *Intl. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Seattle, September 2008.
- [69] R. Ratnam, J. D. and O. W., "Fast Algorithms for Blind Estimation of Reverberation Time," *IEEE signals processing letters*, vol. 11, no. 6, pp. 537-540, June 2004.
- [70] H. Lollmann, E. Yilmaz, M. Jeub and P. Vary, "An improved algorithm for blind reverberation time estimation," in *Intl. Workshop Acoust. Echo Noise Control (IWAENC)*, Tel-Aviv, August 2010.
- [71] ГОСТ Р ИСО 3382-2-2013. Акустика. Измерение акустических параметров помещений. Часть 2. Время реверберации обычных помещений. - Москва, Стандартинформ, 2014. - 20 с..
- [72] R. C. Heyser, "Acoustical Measurements by Time Delay Spectrometry," *Journal of the Audio Engineering Society*, vol. 15, no. 4, pp. 370-382, October 1967.
- [73] M. R. Schroeder, "Integrated-impulse method measuring sound decay without using impulses," *JASA*, August 1979.
- [74] Д. Моторнюк, «Автоматическое вычисление времени реверберации в условиях шумовой помехи,» в *Доклады на VIII Международной научно-технической конференции молодых ученых*, 15-17 апреля 2015.
- [75] "HMS II.3 Artificial Head Measurement System with Ear Simulator and Artificial Mouth," *Data Sheet, HEAD Acoustics GmbH*, pp. 1-4, 2015.
- [76] P. Søndergaard and P. Majdak, "The auditory modeling toolbox," in *The technology of binaural listening*, Berlin-Heidelberg-New York NY, Springer, 2013.
- [77] M. Jeub, M. Schafer and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *16th International Conference on Digital Signal Processing*, Santorini-Hellas, 5-7 July, 2009.

- [78] "Aachen Impulse Response Database," [Online]. Available: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>. [Accessed 24 07 2021].
- [79] "Sound Forge Audio Studio," [Online]. Available: <https://www.magix.com/ru/muzyka/sound-forge/sound-forge-audio-studio/>. [Accessed 25 07 2021].
- [80] E. Habets, N. Gaubitch and P. Naylor, "Temporal selective dereverberation of noisy speech using one microphone," in *2008 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, March-April 2008. pp. 4577-4580.
- [81] А. Н. Тихонов, «О некорректных задачах линейной алгебры и устойчивом методе их решения,» *ДАН СССР*, т. 163, № 3, p. 591—594, 1965.
- [82] M. Jeub, "Blind Reverberation Time Estimation. URL: <http://www.mathworks.com/matlabcentral/fileexchange/35740-blind-reverberation-time-estimation> (Дата звернення 10.08.2021)".
- [83] "AIR database," [Online]. Available: <http://www.ind.rwth-aachen.de/~air>. [Accessed 15 08 2021].
- [84] "TIMIT Acoustic-Phonetic Continuous Speech Corpus," [Online]. Available: <https://catalog.ldc.upenn.edu/ldc93s1>. [Accessed 17 08 2021].
- [85] P. Loizou, *Speech enhancement: Theory and Practice*, Boca Raton: CRC Press, 2007.
- [86] Д. Моторнюк и В. Нечитайло, «Оптимизация параметров алгоритма дереверберации,» в *Доклады на VII Международной научно-технической конференции молодых ученых*, Киев, 15-17 апрель 2014.
- [87] A. Prodeus, A. Vityk, O. Dvornyk, I. Kotvytskyi, O. Chaika and M. Yaroshenko, "Subjective Evaluation of the Speech Intelligibility on the Background of Noise and Reverberation," *Microsystems, Electronics and Acoustics*, vol. 23, no. 2, pp. 66-73, 2018.

- [88] A. Prodeus, K. Bukhta, P. Morozko, O. Serhiienko, I. Kotvytskyi and O. Dvornyk, "Automated Subjective Assessment of Speech Intelligibility in Various Listening Modes," *Microsystems, Electronics and Acoustics*, vol. 23, no. 3, pp. 49-57, 2018.
- [89] GOST R 50840-95: Transmission of Speech Through Communication Paths. Methods for Assessing Quality, Intelligibility and Recognizability. Gosstandart of Russia, Moscow, 1996.
- [90] GOST 16600-72: Transmission of Speech Through the Radiotelephone Communications. Requirements to Intelligibility of Speech and Methods of Articulatory Measurements. Standartinform, Moscow, 2007.
- [91] J. Beerends, E. Larsen, N. Iyer and J. Vugt, "Measurement of speech intelligibility based on the PESQ approach," in *Workshop Measurement of Speech and Audio Quality in Networks (MESAQIN)*, Prague, June 2004.
- [92] J. Beerends, S. Wijngaarden and R. Buuren, "Extension of ITU-T recommendation P.862 PESQ towards measuring speech intelligibility with vocoders," in *New Directions for Improving Audio Effectiveness, Proceedings of the RTO-MP-HFM-123, Neuilly-sur-Seine, France, 2005*. pp. 10-1–10-6..
- [93] J. Ma, Y. Hu and P. Loizou, "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions," *Acoust. Soc. Am*, vol. 125, no. 5, pp. 3387-3405, May 2009.
- [94] Y. Hu and P. Loizou, "Evaluation of objective quality measures for speech enhancement," *IEEE Transactions on Speech and Audio Processing*, vol. 16, no. 1, pp. 229-238, 2008.
- [95] R. M. Gray, A. Buzo, A. H. Gray and Y. Matsuyama, "Distortion measures for speech processing," *IEEE Transactions on Acoustics Speech and Signal Processing*, Vols. ASSP-28, no. 4, pp. 367-376, September 1980.

- [96] A. H. J. Gray and J. D. Markel, "Distance measures for speech processing," *IEEE Trans. Acoust., Speech, Signal Processing*, Vols. ASSP-24, no. 5, pp. 380-391, Oct. 1976.
- [97] N. Cote, *Integral and diagnostic intrusive prediction of speech*, Berlin: Springer-Verlag, 2011.
- [98] A. Prodeus, A. Vityk and D. Didenko, "Subjective evaluation of quality and intelligibility of speech distorted by synthesized noise," *Microsystems, Electronics and Acoustics*, vol. 22, no. 6, pp. 56-63, 2017.
- [99] "Telephone transmission quality, telephone installations, local line networks. Methods for objective and subjective assessment of quality. Perceptual evaluation of speech quality (PESQ).," *Recommendation P.862. Series P*, 2001.
- [100] "Perceptual Evaluation of Speech Quality (PESQ) ITU-T Recommendations P.862, P.862.1, P.862.2. Version 2.0," October 2005.
- [101] W. Ahnert та W. Schmidt, *Appendix to EASERA Manual: Fundamentals to Perform Acoustical Measurements*, Berlin, 2006. – P. 1 – 53..
- [102] А. Продеус, В. Дидковский и М. Дидковская, *Акустическая экспертиза и коррекция коммуникационных каналов. Монография.*, Saarbrücken, Deutschland: LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co. KG, 2017.
- [103] Y. Makarov and A. Khorev, "To the Evaluation of the Effectiveness of Protection of Acoustic (Speech) Information," *Special Technique*, no. 2, 2000.
- [104] A. Prodeus, V. Didkovskyi та M. Didkovska, *Acoustic Expertise of Voice Communication Channels: Monograph*, Kyiv: Imex-LTD, 2008.
- [105] A. Prodeus, «Formant-Modulation Method of Speech Intelligibility Evaluation: Measuring and Exactness,» в *VII Int. Conf. MEMSTECH 2011*, Polyana, Ukraine, 11-14 May 2011, pp.54-60.

- [106] T. Houtgast, H. J. M. Steeneken and S. J. van Wijngaarden, Past, present and future of the Speech Transmission Index, Soesterberg: TNO Technische Menskunde, 2002.
- [107] K. Jacob, "Correlation of Speech Intelligibility Tests in Reverberant Rooms with Three Predictive Algorithms," in *85th Convention of the Audio Engineering Society*, Los Angeles, November 3-6, 1988. pp. 1020-1030.
- [108] F. Xiong, S. Goetze and B. Meyer, "Estimating Room Acoustic Parameters for Speech Recognizer Adaptation and Combination in Reverberant Environments," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*, Florence, 4-9 May, 2014.
- [109] H. Golzer and M. Kleinschmidt, "Importance of early and late reflections for automatic speech recognition in reverberant environments," in *Elektronische Sprachsignalverarbeitung (ESSV)*, 2003.
- [110] Y. Hu and K. Kokkinakis, "Effects of early and late reflections on intelligibility of reverberated speech by cochlear implant listeners," *J. Acoust. Soc. Am.*, vol. 135, no. 1, January 2014.
- [111] M. R. Schroeder, B. S. Atal and C. Bird, "Digital computers in room acoustics," in *4th International Congress on Acoustics, ICA*, Copenhagen, 1962. p. M21.
- [112] J. Rindel, "The Use of Computer Modeling in Room Acoustics," *Journal of Vibroengineering*, vol. 3, no. 4, pp. 219-224, 2000.
- [113] M. Vorlander, "Computer simulations in room acoustics: Concepts and uncertainties," *J. Acoust. Soc. Am.*, vol. 133, no. 3, pp. 1203-1213, March 2013.
- [114] K. D. Jacob, T. K. Birkle and C. B. Ickler, "Accurate Prediction of Speech Intelligibility without the Use of In-Room Measurements," *Audio Eng. Soc.*, vol. 39, no. 4, pp. 232-242, April 1991.
- [115] H. J. M. Steeneken and T. Houtgast, "Validation of the revised STIr method," *Elsevier Speech Communication*, vol. 38, pp. 26-37, 2002.

- [116] H. J. M. Steeneken, «Forty years of speech intelligibility assessment (and some history),» в *40 Years anniversary Institute of Acoustics*, Birmingham, 15-16 October, 2014.
- [117] A. Warzybok, J. Rannies, S. Doclo та В. Kollmeier, «Influence of early reflections on speech intelligibility under different noise conditions,» в *Forum Acusticum 2011*, Aalborg, Denmark, June 27-July 1, 2011.
- [118] И. Алдошина и Р. Приттс, Музыкальная акустика, С-Пб: Композитор, 2006. – 719 с..
- [119] J. S. Bradley, "The Evolution of Newer Auditorium Acoustics Measures," *Canadian Acoustics*, vol. 18, no. 4, pp. 13-23, 1990.
- [120] J. S. Bradley, «Review of objective room acoustics measures and future needs,» в *International Symposium on Room Acoustics, ISRA 2010*, Melbourne, 29-30 August, 2010. pp. 1-8.
- [121] Санитарные нормы допустимого шума в помещениях жилых и общественных зданий и на территории жилой застройки. - Министерство здравоохранения СССР, Главное санитарно-эпидемиологическое управление. - 1984 г..
- [122] ДСН 3.3.6.037-99 Санітарні норми виробничого шуму, ультразвуку та інфразвуку. - Постанова ОЗ України N 37 від 01.12.99, м.Київ.
- [123] M. Long, *Architectural Acoustics*, Elsevier Academic Press, 2006. – 873 p..
- [124] H. J. M. Steeneken and T. Houtgast, "A physical method for measuring speech-transmission quality," *Acoust. Soc. Am*, vol. 67, pp. 318-326, 1980.
- [125] А. Н. Продеус, Л. Б. Дронжевская, В. А. Климов и Д. А. Шагитова, «Формантный и формантно-модуляционный методы оценки разборчивости речи. Часть 1. Унификация алгоритмов,» *Электроника и связь*, № 6, часть 2, pp. 117-124, 2010.
- [126] А. Н. Продеус, Б. Л. Дронжевская, В. А. Климов и Д. А. Шагитова, «Формантный и формантно-модуляционный методы оценки

разборчивости речи. Часть 2. Точность и скорость измерений,» *Электроника и связь*, т. 64, № 6, р. 16–24, 2011.

- [127] M. Schroeder, "Modulation Transfer Functions: Definition and Measurement," *Acustica*, vol. 49, pp. 179-182, 1981.
- [128] Н. Покровский, Расчет и измерение разборчивости речи, Москва: Связьиздат, 1962. - с. 390.
- [129] М. А. Сапожков, Речевой сигнал в кибернетике и связи, Москва: Связьиздат, 1963. – 472 с..
- [130] U. Zölzer, DAFX: Digital Acoustic Effects, 2nd ed., John Wiley & Sons, 2011.
- [131] С. В. Алейник, Ю. Н. Матвеев та А. Н. Раев, «Метод оценки уровня клиппирования речевого сигнала,» *Научно-технический вестник информационных технологий, механики и оптики*, т. 3, № 79, pp. 79-83, 2021.
- [132] A. Prodeus, I. Kotvytskyi, M. Didkovska та К. К., «Kurtosis and Normalized Variance as Measures of Speech Signals Clipping Value,» *Electronics and Control System*, № 4(62), pp. 24-32, 2019.
- [133] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, Numerical Recipes in C, Cambridge: Cambridge U.P, 1992.
- [134] Д. Є. Моторнюк, «Дослідження похибки вимірювання часу реверберації всліпу,» в *IX міжнародна науково-технічна конференція молодих вчених «Електроніка-2016»*, Київ, ст. 77-78. 2016.

ДОДАТОК А. Список опублікованих праць за темою дисертації

Статті в періодичних наукових виданнях інших держав або прирівняні до таких (публікації у виданнях категорії “А” або в закордонних виданнях, проіндексованих у базах даних Web of Science Core Collection та/або Scopus):

1. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, and D. Motorniuk, “Automated Subjective Assessment of Speech Intelligibility under Diotic and Dichotic Listening,” Lecture Notes on Data Engineering and Communications Technologies, 2021, 48, pp. 1–16, January 2021, Springer, Cham, ISSN 2367-4512, e-ISSN 2367-4520, DOI: 10.1007/978-3-030-43070-2_1
Scopus: <https://www.scopus.com/sourceid/21100975545>

Статті в наукових виданнях, включених до переліку наукових фахових видань України з присвоєнням категорії “Б”:

1. A. Prodeus, M. Didkovska, K. Kukharicheva, D. Motorniuk, "Two Simplified Models of Early Sound Reflections in a Room", Electronics and Control Systems, No. 3(65), 2020, pp. 79-87. DOI:10.18372/1990-5548.65.14991, ISSN 1990-5548 (Published: 2020-11-27)

2. O. Dvornyk, A. Prodeus, D. Motorniuk, M. Didkovska, "Hardware and Software System "Artificial Head". Part 2. Evaluation of Speech Intelligibility in Classrooms," Microsystems, Electronics and Acoustics, vol. 22, no. 3, 2020, pp. 48-55, DOI: 10.20535/2523-4455.me.209928 (Опубліковано: Dec 28, 2020), ISSN 2523-4447, e-ISSN 2523-4455

3. O. Dvornyk, A. Prodeus, M. Didkovska, D. Motorniuk, "Artificial Software Complex "Artificial Head". Part 1. Adjusting the Frequency Response of the Path," Microsystems, Electronics and Acoustics, vol. 22, no. 1, 2020, pp. 56-64, DOI: 10.20535/2523-4455.me.198431, ISSN 2523-4447, e-ISSN 2523-4455

4. Ю.С. Костючок, Л.С. Мартинович, Д.Е. Моторнюк, В.А. Нечитайло, А.В. Храпачевский, А.Н. Продеус. Акустическая паспортизация учебных помещений.

- ISSN 1811-4512. Electronics and Communications, 2016, Vol.21, No.2(91). – P.63-70

Доповіді на конференціях:

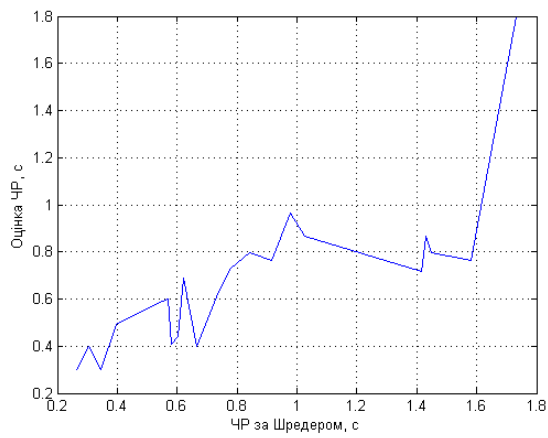
1. A. Prodeus, M. Didkovska, D. Motorniuk, O. Dvornyk, "The Effects of Noise, Early and Late Reflections on Speech Intelligibility" Proc. 2020 IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO), pp. 488 - 492, DOI: 10.1109/ELNANO50318.2020.9088854

2. A. Prodeus, M. Didkovska, K. Kukharicheva, D. Motorniuk, "Modeling the Influence of Early Sound Reflections on Speech Intelligibility," Proc. of 2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 20-23 Oct. 2020, Kyiv, Ukraine.
DOI: 10.1109/MSNMC50359.2020.9255657

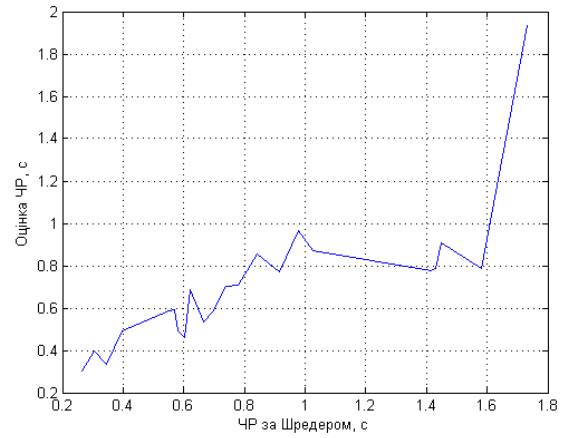
3. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, D. Motorniuk, and A. Khrapachevskyi, "Objective and Subjective Assessment of the Quality and Intelligibility of Noised Speech," Proc. of 2018 IEEE International Scientific-Practical Conference "Problems of Infocommunications. Science and Technology" (PIC S&T`2018), October 9-12, 2018, Kharkiv, Ukraine, pp. 71-74.
DOI: 10.1109/INFOCOMMST.2018.8632125.

А. М. Продеус, І. В. Котвицький, М. В. Дідковська, В. С. Дідковський, К. А. Кухарічева, Д. Є. Моторнюк, О. О. Дворник Спосіб виявлення кліпування мовного та музичного сигналів // **Патент** UA 144291 U, МПК G01R 23/20, опубл. 25.09.2020.

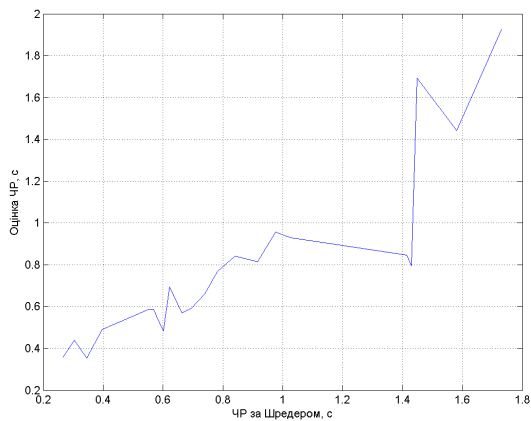
ДОДАТОК Б. Графіки відповідності ЧР виміряного методом МП дійсному ЧР



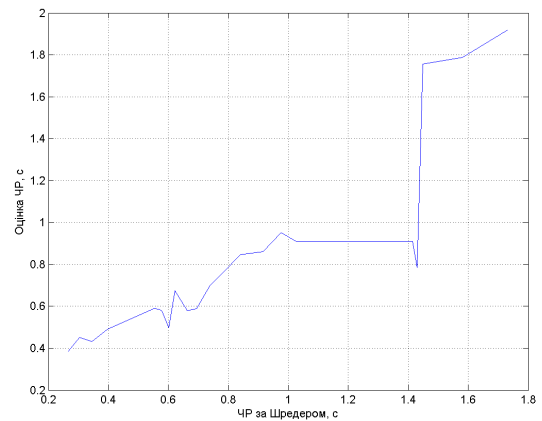
а



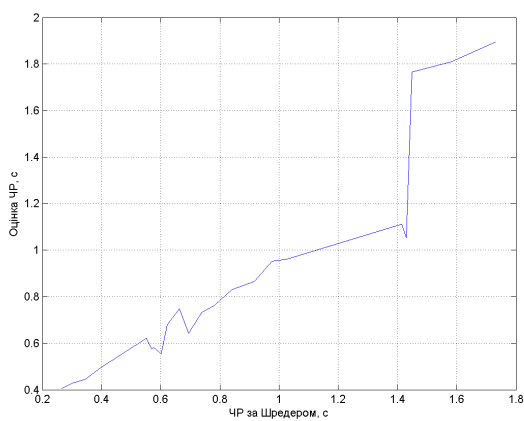
б



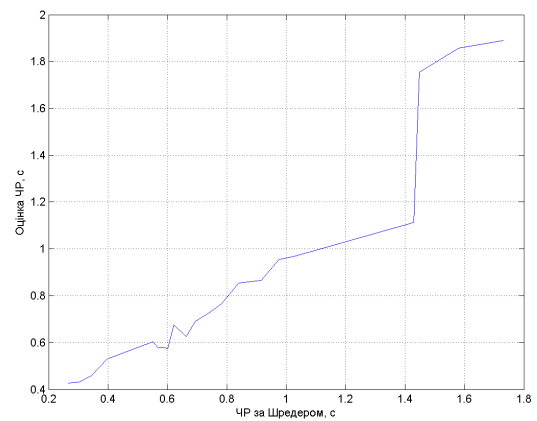
в



г

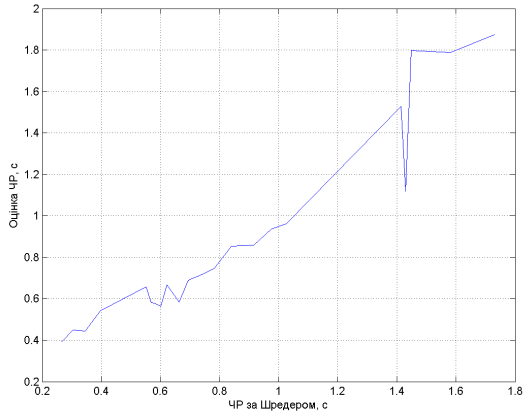


д

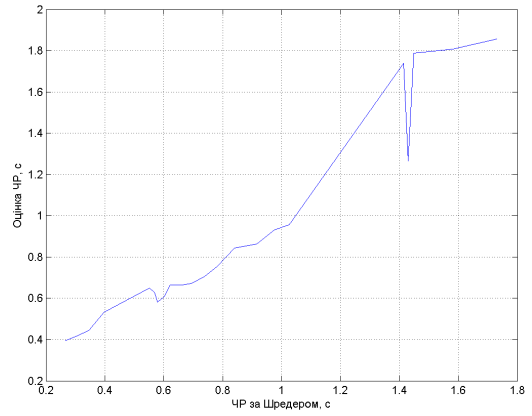


е

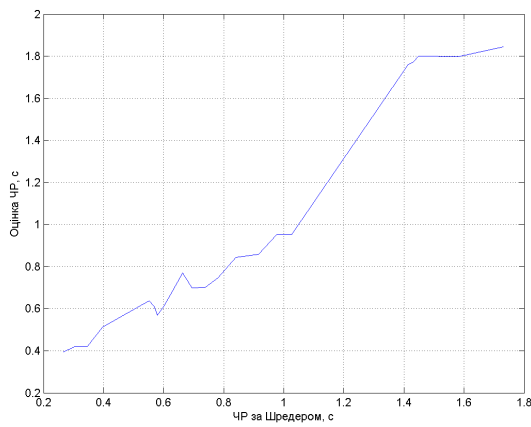
Рис. Б.1 Графіки відповідності ЧР виміряного методом МП дійсному ЧР для довжини підфрейму: (а) 0,04с, (б) 0,045 с, (в) 0,05 с, (г) 0,055 с, (д) 0,06 с, (е) 0,065 с.



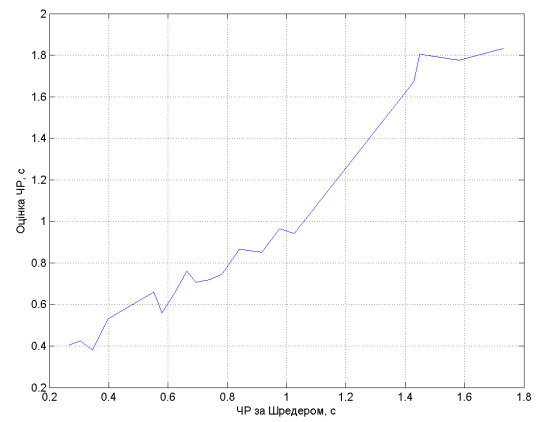
а



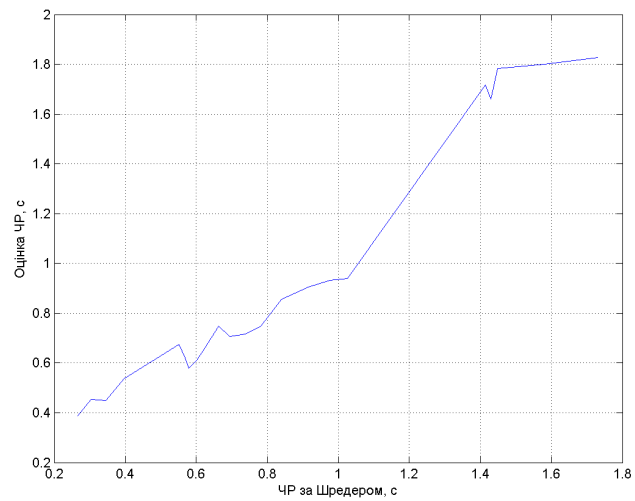
б



в



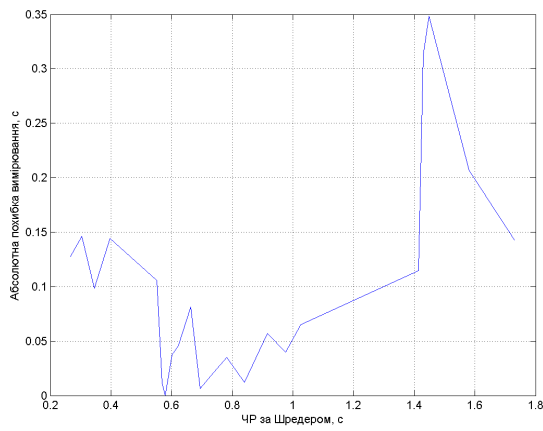
г



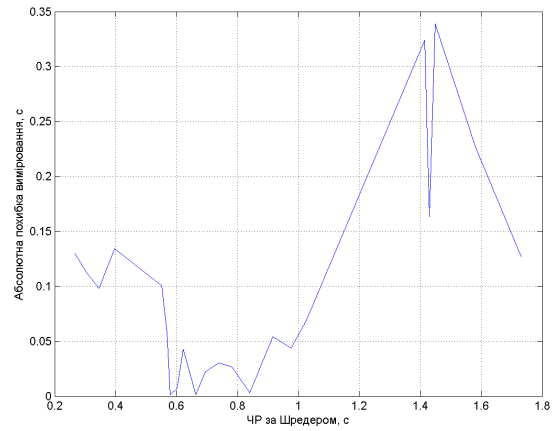
д

Рис. Б.2 – Графіки відповідності ЧР виміряного методом МП дійсному ЧР для довжини підфрейму: (а) 0,07с, (б) 0,075 с, (в) 0,08 с, (г) 0,085 с, (д) 0,09 с.

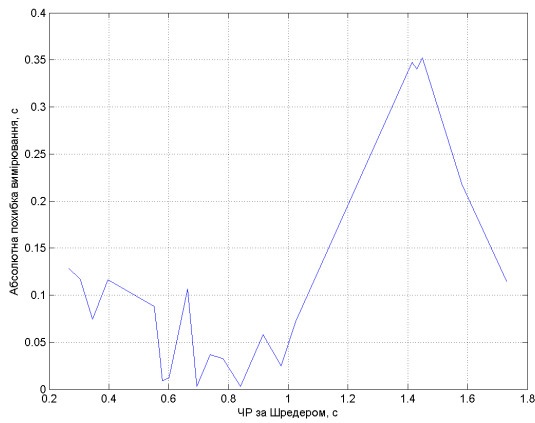
ДОДАТОК В. Графіки залежності абсолютної похибки вимірювання від часу реверберації



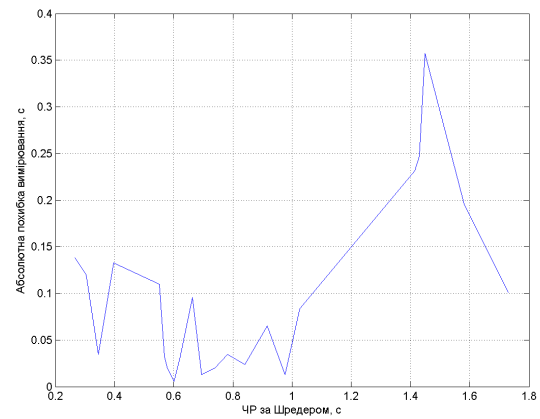
а



б



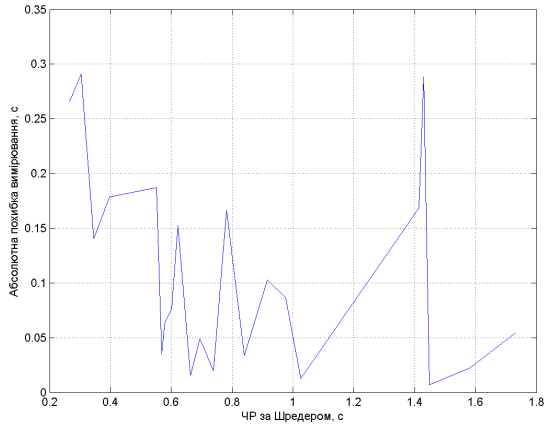
в



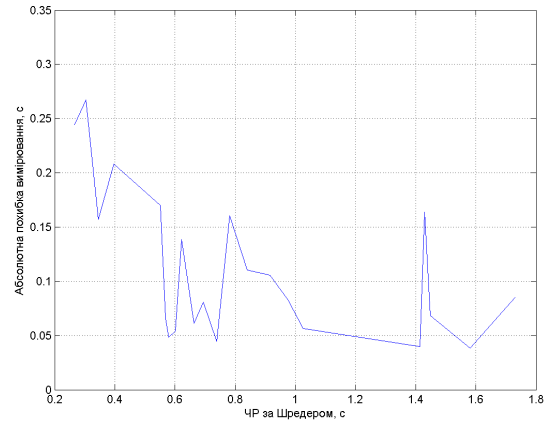
г

Рис. В.1. – Графік залежності абсолютної похибки вимірювання від часу реверберації для довжини підфрейму: (а) 0,07с, (б) 0,075 с, (в) 0,08 с, (г) 0,085

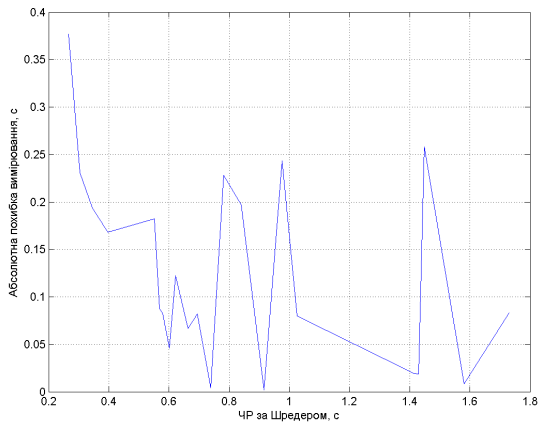
с.



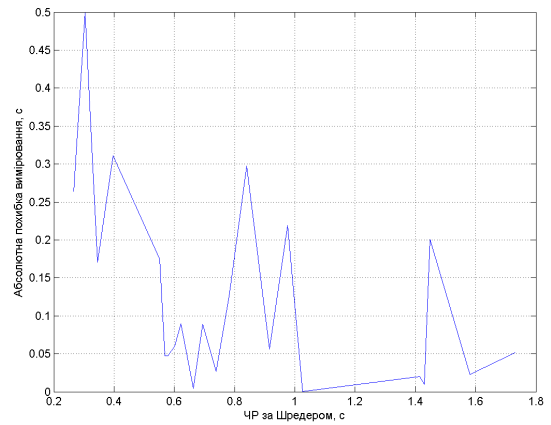
а



б



в



г

Рис. В.2. – Графік залежності абсолютної похибки вимірювання від часу реверберації для довжини підфрейму: (а) 0,07с, (б) 0,075 с, (в) 0,08 с, (г) 0,085 с. Алгоритм з двома гістограмами.

ДОДАТОК Г. Приклад запропонованого акустичного паспорту навчальної аудиторії

Акустичний паспорт аудиторії №438

1.	Дата і час	16.11.15 13:30-14:00
2.	Місце проведення вимірювань	Корпус № 12 НТУУ "КПІ ім. І. Сікорського", аудиторія 438: вул. академіка Янгеля, 16/9
3.	Склад експертної групи	Костючок Ю.С., Мартинович Л.С., Моторнюк Д.Є., Нечитайло В.О.

Виміряні параметри приміщення

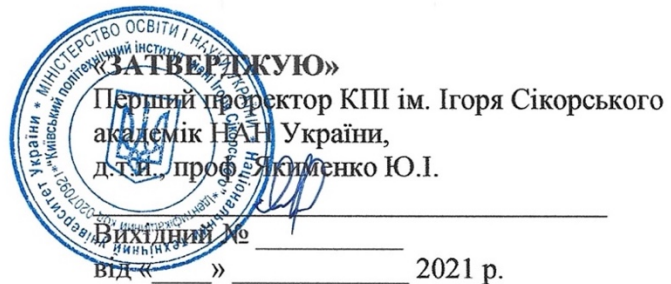
4.	Частота, Гц	ВР, с			
	63	1,1875	1,0540	0,7576	
	125	1,2773	1,2318	0,5978	
	250	1,4667	1,5744	1,4614	
	500	1,5176	1,4734	1,4512	
	1000	1,6802	1,6584	1,3365	
	2000	1,6660	1,7327	1,6487	
	4000	1,4862	1,4624	1,3374	
	8000	1,2925	1,2346	1,0261	

	Параметр	Виміряне значення	Норма
5.	Час приходу першого відбиття, с	0,015	≤0,05
6.	Середній рівень шуму, дБА	44	≤50
7.	Макс. рівень шуму, дБА	60,7	≤55
8.	Індекс чіткості C50, дБ	-0,9499	≥-2
9.	Індекс прозорості C80, дБ	1,8085	≥-3
10.	Коеф. складової розбірливості	0,8887	> 0,8
11.	Тембр (коефіцієнт низького тону)	0,8960	0,9 -1,0

Заключення експертної групи:

У приміщенні було зафіксовано максимальне значення рівня шуму, яке перевищує санітарні норми, що пояснюється приєднаної до аудиторії лаборантської кімнати, стіни і двері якої пропускають шум до аудиторії. Окрім цього, приміщення є гулким. Рекомендовано провести корекцію акустичних характеристик цієї аудиторії.

ДОДАТОК Д. Акт про використання результатів дисертаційної роботи



АКТ

про використання результатів дисертаційної роботи
 випускника аспірантури кафедри акустичних та мультимедійних електронних
 систем Національного технічного університету України «Київський політехнічний
 інститут імені Ігоря Сікорського» Моторнюк Д.Є. у навчальному процесі

Ми, що нижче підписалися, декан факультету електроніки КПІ ім. І. Сікорського д.т.н., проф. Жуйков В.Я., зав. кафедрою акустичних та мультимедійних електронних систем д.т.н., проф. Найда С.А., д.т.н., проф. Продеус А.М., к.т.н., доц. Богданов О.В., склали цей Акт про те, що результати дисертаційної роботи Моторнюк Д.Є. «Вимірювання часу реверберації в контексті задачі дереверберації мовного сигналу» впроваджені у навчальний процес кафедри акустичних та мультимедійних електронних систем, а саме, у дисципліні «Обробка сигналів електронних та акустичних систем» відображені методи та алгоритми цифрової обробки сигналів при оцінюванні акустичних характеристик приміщення, лекції за наступними темами:

- Штучна голова та вимірювання бінауральних акустичних характеристик
- Банк бінауральних імпульсних характеристик
- Проблеми автоматизації суб'єктивного оцінювання розбірливості мови
- Проблеми автоматизації об'єктивного оцінювання розбірливості мови

Декан факультету електроніки

Зав. каф. АМЕС

Професор каф. АМЕС

д.т.н., проф. Жуйков В.Я.

д.т.н., проф. Найда С.А.

д.т.н., проф. Продеус А.М.,