КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (27), 2025
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2025.27.617632
УДК 004.934.2: 534.442.6

Нужний Сергій Миколайович
кандидат технічних наук, доцент,
завідувач кафедри комп'ютерних технологій та інформаційної безпеки
Національний університет кораблебудування імені адмірала Макарова, м. Миколаїв, Україна
ORCID: 0000-0002-7706-0453
s.nuzhniy@gmail/com

УДОСКОНАЛЕННЯ АЛГОРИТМУ ВІДНОВЛЕННЯ ЛІНГВІСТИЧНОЇ СКЛАДОВОЇ МОВНОЇ ІНФОРМАЦІЇ ФОНЕМНО-ФОРМАНТНИМ МЕТОДОМ ДЛЯ ЗАДАЧ ОЦІНКИ РІВНЯ ЇЇ ЗАХИЩЕНОСТІ

Анотація. У статті розглянуто удосконалення алгоритму відновлення лінгвістичної складової мовної інформації фонемно-формантним методом для задач оцінювання рівня її захищеності. Проведено комплексний аналіз історичного розвитку артикуляційного та формантного підходів — від класичних робіт Bell Laboratories і досліджень Гарві Флетчера до сучасних цифрових методів оцінювання розбірливості мовлення, закріплених у стандартах ANSI S3.5-1997 та ANSI S3.5-2007. Показано, що еволюція методів аналізу мовного сигналу безпосередньо вплинула на формування систем технічного захисту інформації, особливо у частині контролю каналів акустичного витоку. Запропоновано десятикроковий алгоритм аналізу та реконструкції мовного сигналу, який об'єднує спектральні, статистичні й контекстні підходи. Алгоритм включає етапи детекції мовної активності (Voice Activity Detection, далі VAD), сегментації сигналу, виділення формант за допомогою перетворення Фур'є (Fast Fourier Transform, далі FFT) та лінійного передбачення (Linear Predictive Coding ,далі LPC), а також побудову ймовірнісної трифонної моделі. Урахування фазових зсувів, коартикуляційних ефектів і міжформантних інтервалів забезпечує стійкість і точність відновлення навіть за умов низького співвідношення сигнал/завада (Signal-to-Noise Ratio, далі SNR). Особливу увагу приділено оцінюванню впливу спотворень спектрального складу формант на розпізнавання фонем і визначенню статистичних закономірностей появи дифонів і трифонів української мови. Це дозволило створити адаптивну модель, що поєднує акустичні та мовні ознаки для більш достовірної реконструкції лінгвістичної складової повідомлення. Розроблений алгоритм може бути використаний у практиці технічного захисту інформації для об'єктивного оцінювання рівня захищеності мовних каналів, а також для відновлення пошкоджених або частково зашумлених аудіозаписів. Його структура відповідає вимогам національних нормативів НД ТЗІ 1.6-005-2013 та НД ТЗІ 3.7-003-2023, що визначають методики аналізу каналів витоку мовної інформації та ефективності засобів захисту. Отримані результати мають подвійне прикладне значення: по-перше, вони можуть бути використані для створення експертних систем контролю витоку мовної інформації, а по-друге — у системах автоматичного розпізнавання, синтезу та відновлення мовлення, де необхідна висока точність відтворення фонемної структури. Запропонований фонемно-формантний підхід демонструє універсальність і здатність інтегруватися у сучасні цифрові засоби аудіоаналізу, що підвищує рівень інформаційної безпеки та якість технічних засобів акустичного контролю.

Ключові слова: фонемно-формантний аналіз; спектральна реконструкція; трифонна модель; коартикуляція; захист мовної інформації.

ВСТУП

Перші дослідження мовного сигналу, який передається комунікаційними лініями, з'явилися у 1920-х роках XX століття. Найбільший внесок у цю сферу зробили науковці лабораторії AT&T Bell Labs (Bell System), серед яких – І. Б. Крандалл, Гарві Флетчер, Х.

Д. Арнольд, Біддулф, Данн, Френч, Фрай, Гальт, Гарднер, Грем, Хартлі, Кінгсбері, Кеніг, Кранц, Лейн, Маккензі, Мунсон, Рієз, Сасія, Шауер, Сівіан, Стейнберг, Дж. К. Стюарт, Вегель та Венте. Вони вперше виміряли слуховий поріг людини, встановили залежність між гучністю та її зростанням (нині відому як закон Стівенса), ввели поняття адитивності гучності та адитивності артикуляції (що пізніше стало основою індексу артикуляції (Articulation Index, далі AI). Гарві Флетчер також сприяв створенню вакуумного слухового апарата, комерційного аудіометра, штучної гортані та системи стереофонічного звучання [1].

Подальший розвиток артикуляційного підходу здійснив Джон Коллард у 1929 році (публікація [2] та патент «Interference determining»), уперше застосувавши формантний підхід до аналізу мовної інформації.

У 1947 році Н. Р. Френч, Дж. К. Стейнберг [3] і Лео Л. Беранек [4] удосконалили формантний метод, що стало відомим як третя редакція артикуляційного методу.

Через кілька років, у 1950 році, опубліковані праці Гарві Флетчера [5], [6] фактично завершили формування методологічного базису цього методу, започаткувавши четверту хвилю розвитку артикуляційного аналізу. У межах цього етапу було сформульовано поняття артикуляційного методу визначення розбірливості мовної інформації, яка передається комунікаційними каналами (телефонними та радіолініями). Також розроблено математичні засоби розрахунку розбірливості мовлення диктора, запропоновано алгоритми експериментальних досліджень і визначено організаційні вимоги до проведення таких вимірювань.

Подальші модифікації методу стосувалися деталізації спектрального аналізу: частотний діапазон було поділено спочатку на низькі та високі частоти, а з середини 1970-х років – на октавні, а згодом – третинооктавні смуги.

Розвиток артикуляційних досліджень сприяв появі нового поняття – індекс зрозумілості мовлення (Speech Intelligibility Index, далі SII), тобто індексу зрозумілості мовлення. Перехід від традиційного AI до SII було офіційно закріплено у 1997 році стандартом ANSI S3.5-1997 “Methods for the Calculation of the Speech Intelligibility Index” [7], оновленим у 2007 році (ANSI S3.5-2007 [8]).

Бурхливий розвиток комп’ютерної техніки наприкінці 1970-х – на початку 1980-х років істотно вплинув на стан захисту мовної інформації в об’єктах інформаційної діяльності. Насамперед це було зумовлено появою вокодерів, здатних частково відновлювати пошкоджені фонограми, хоча їх ефективність обмежувалася рівнем і типом завад.

Такий технологічний прогрес стимулював фахівців із безпеки шукати нові засоби запобігання перехопленню мовної інформації. Тоді вважалося, що використання «білого шуму» в системах активних завад є цілком достатнім. Цьому сприяли як позитивний досвід попередніх експериментів, так і прагнення знизити вартість розробок. У той період нормативна та методична база не передбачала фільтрацію сигналу завади, створеного на основі «білого шуму». Його випадковий спектр і аналогові методи обробки забезпечували високу стійкість такого захисту до тодішніх способів очищення сигналу.

Згодом це призвело до появи інструментальних методів контролю рівня захищеності мовної інформації, у яких білий шум використовувався одночасно як модель небезпечного сигналу й як активна завада. Хоча ці сигнали не були корельованими (через використання різних генераторів), такий підхід було закріплено в документах СРСР із середини 1970-х років [9]. Рівень захищеності визначали за

коефіцієнтом AI, а після 1997 року – за індексом SII. У СРСР, а згодом і в Україні, також застосовувалося відношення сигнал/завада SNR.

Перехід до цифрових технологій мав подвійний ефект: він покращив ефективність систем захисту (наприклад, підвищення тактової частоти генератора «білого шуму» ускладнює реверс-інжиніринг сигналу), але водночас підвищив їх вразливість до цифрових методів фільтрації [10] – [12].

Метою роботи є: удосконалення алгоритму відновлення лінгвістичної складової мовної інформації фонемно-формантним методом для задач оцінки рівня її захищеності

Для досягнення поставленої мети в роботі вирішуються задачі:

– Аналіз впливу спектрального складу формант та його викривлення на розпізнавання диктором фонем;
– Удосконалення алгоритму відновлення лінгвістичної складової мовної інформації фонемно-формантним методом.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Перші дослідження спектру мови диктора та розподіл її на фонеми з виділенням формант відноситься до першої половини XX-ого століття. Однією з найвідоміших робіт є [13]. Робота, фактично, є звітом Bell Laboratories з досліджень мови диктора та створення пристрою для штучного її відтворення. Основними результатами роботи є:

– людська мова може бути відтворена через управління спектральними компонентами без прямої артикуляції;
– форманти відіграють ключову роль у розпізнаванні голосних;
– приголосні потребують додаткового шумового сигналу для точного відтворення;
– запропонований пристрій Voder продемонстрував технічну можливість синтезу мови.

Подальший розвиток теорія отримала в роботах [14] – [17].

Особливістю [14] є дослідження особливостей японської мови та порівняння її з іншими. Ключовими результатами роботи є:

– встановлено залежність F1, F2, F3 від артикуляції голосних;
– показано, що приголосні формують короткочасні шумові спектри, які можна аналізувати як акустичні маркери;
– запропоновано спектральну класифікацію приголосних на основі шумових характеристик та поєднання з формантами голосних;
– визначено, що форманти є ключовими акустичними характеристиками фонем (використовується до сьогодні у синтезі та розпізнаванні мови);
– надала математично основу для LPC та FFT-аналізу (LPC – кодування з лінійним предиктором; FFT – швидке перетворення Фур'є).

В [14] визначено набір параметрів, які забезпечують просте та точне описання артикуляції голосних звуків. Використовуючи електричну аналогію голосового тракту, автори дослідили зв'язок між артикуляційними параметрами та формантними частотами, що стало основою для подальших досліджень у галузі акустичної фонетики.

[15] є фундаментальним дослідженням у галузі акустичної фонетики, яке поєднує фізичні, акустичні та артикуляційні аспекти мовного процесу. Ця праця стала основою для формулювання моделі джерела-фільтра мовного виробництва (включає переклад, редагування, локалізацію, а також розробку мовних технологій, таких як програмне забезпечення для обробки природної мови, машинного перекладу та голосових помічників).

Робота [16] розглядає форманти як локальні максимуми спектральної обгортки звукового сигналу, що виникають через акустичні резонанси в повітряному стовпі голосового тракту. Автор підкреслює, що не всі піки в спектрі можна віднести до резонансів голосового тракту; деякі з них є спуріусними (помилковими) або латентними (прихованими). В [16] також обговорює складність точного визначення формантів, особливо у високочастотних ділянках спектра, де вплив навколишнього середовища та акустичних властивостей запису може спотворювати результати. Він зазначає, що найбільш надійне виявлення формантів спостерігається у чоловічому мовленні з низьким основним тоном, де гармонічні обертони співпадають з резонансами голосового тракту.

В [16] та врахувавши [17] формуємо модель голосового тракту як трубку зі змінним перерізом, що резонує і формує форманти – піки енергії в спектрі голосних

.

АНАЛІЗ ВПЛИВУ СПЕКТРАЛЬНОГО СКЛАДУ ФОРМАНТ ТА ЙОГО ВИКРИВЛЕННЯ НА РОЗПІЗНАВАННЯ ФОНЕМ

У частотній області спектр сигналу *S(f)* можна представити як:

$S(f) = G(f) \cdot H(f)$,

де *G(f)* – спектр джерела (вихід від голосових зв'язок), *H(f)* – частотна характеристика фільтра, що моделює резонанси голосового тракту.

Для спрощення, розглянемо голосовий тракт як трубку довжини L із закритим одним кінцем (глотка) і відкритим іншим (рот) [17].

Резонансні частоти (форманти) *Fn* визначаються за формулою стоячих хвиль [17]:

$$Fn = \frac{(2n-1)\cdot c}{4L}, \; n = 1, 2, 3, \ldots$$

де c – швидкість звуку в повітрі (≈ 343 м/с); L – ефективна довжина голосового тракту; n – порядковий номер форманту.

Особливості формування формант [17]:

– F_1 залежить від відкритості рота та висоти язика;

– F_2 визначається переднім/заднім положенням язика;

– F_3 і F_4 корелюють з іншими артикуляційними характеристиками мовного тракту (губи, глотка).

– Форманти відповідають за піки спектра голосного. Позиція та амплітуда формантів визначають вокальну якість (тембр) голосного. Наприклад, [i] має високий F_2 через переднє положення язика, а [u] – низький F_2 через заднє положення.

Для складнішого голосового тракту з *N* резонансами частотна характеристика *H(f)* може бути подана як добуток резонансних фільтрів другого порядку [17]:

$$H(f) = \prod_{n=1}^{N} \frac{f_n^2}{f_n^2 - f^2 + j2\beta_n f} \quad (1)$$

де *fn* – частота n-го форманту; *βn* – ширина резонансу (визначає «гостроту» піка); $j = \sqrt{-1}$ –уявна одиниця.

Значення частот (усереднене) для перших чотирьох формант, які впливають на семантичну складову інформації, наведено в табл. 1 (для англійської голосних фонем) та табл. 2 (для українських голосних фонем). Необхідно відмітити, що значення усереднені

та можуть варіюватися залежно від голосу та конкретного мовця, в тому числі від стану його здоров'я, втомленості та інше).

Як видно з табл. 1 та 2, шість з дев'яти фонем для англійської та української мов є ідентичними за спектральним складом. Однак по три фонеми є оригінальними (виділені затемненням) і визначають особливості звучання мови диктора.

Таблиця 1

Розподіл частот по формантам для англійських голосних фонем

Голосний	F_1	F_2	F_3	F_4
/i/	270	2290	3010	4000
/ɪ/	390	1990	2550	3600
/e/	530	1840	2480	3500
/æ/	660	1720	2410	3300
/ɑ/	730	1090	2440	3300
/ɔ/	570	840	2410	3300
/o/	570	840	2410	3300
/u/	300	870	2240	3400
/ʊ/	440	1020	2240	3400

Необхідно відмітити, що в таблицях наведені данні про голосні звуки які впливають на мовленнєву характеристику мови диктора (тобто на її «звучання»). В [18] наведено результати досліджень по заміні голосних звуків в словосполученнях та показано що, наприклад, при заміні звуку /a/ на звук /i/ при збереженні фонетичної функції звукосполучення слухач не помічає заміни і розпізнає звукосполучення. Також можливі заміни /e/ на /i/ або /o/ на /y/. Це підтверджується спілкуванням з дикторами, що мають місцевий діалект або пройшли процес асиміляції – «окання», «акання», «гикання» та інше. Такий ефект широко розповсюджений вокодерах та дає можливість підтримувати необхідний рівень зв'язку при допустимій кількості помилок при відновленні фонів в мовному повідомленні [18], [19].

Таблиця 2

Розподіл частот по формантам для українських голосних фонем

Голосний	F_1	F_2	F_3	F_4
і	270	2290	3010	4000
и	390	1990	2550	3600
е	530	1840	2480	3500
є	500	1800	2500	3500
а	730	1090	2440	3300
о	570	840	2410	3300
у	300	870	2240	3400
ю	350	1020	2240	3400
я	500	1800	2500	3500

В [20] – [22] більш детально досліджено вплив голосних звуків/фонів на сприйняття слухачами фонограм українською мовою та розглянуто їх формантно-спектральні характеристики.

Однак, в [18] – [22] також вказується на складності аналізу приголосних звуків, що пов'язано зі складністю виділення формант, особливо при аналізі дифонів та алофонів/трифонів. З іншого боку, як показує аналіз (1), якраз приголосні фони й відповідають за лінгвістичну складову мовної інформації. Так як (1) містить множник *j*

який є уявною одиницею, то частотна характеристика $H(f)$ може бути подана як сума дійсної та уявної частини

$$H(f)=\prod_{n=1}^{N}\left(\Re[Hn(f)]+j\Im[Hn(f)]\right)=\prod_{n=1}^{N}\left(\frac{f_n^2\left(f_n^2-f^2\right)}{\left(f_n^2-f^2\right)^2+\left(2\beta_n f\right)^2}-j\frac{2f_n^2\beta_n f}{\left(f_n^2-f^2\right)^2+\left(2\beta_n f\right)^2}\right) \quad (2)$$

де $\Re[Hn(f)]=\frac{f_n^2\left(f_n^2-f^2\right)}{\left(f_n^2-f^2\right)^2+\left(2\beta_n f\right)^2}$ – дійсна частина частотної характеристики, яка відповідає за форму та амплітуду сигналу, визначає розміщення формант в фонемі; $\Im[Hn(f)]=\frac{2f_n^2\beta_n f}{\left(f_n^2-f^2\right)^2+\left(2\beta_n f\right)^2}$ – уявна частина, відповідає за фазовий зсув сигналу та поєднання фонем в дифони та трифони.

Для одного множника $H_n(f)$ залежності для модуля і фази мають вигляд:

– модуль $|H_n(f)|=\frac{f_n^2}{\sqrt{\left(f_n^2-f^2\right)^2+\left(2\beta_n f\right)^2}}$;

– фаза $\varphi_n(f)=-\arctan\left(\frac{2\beta_n f}{f_n^2-f^2}\right)$.

Таким чином, для повної системи $H_n(f)$ отримуємо:

– модуль $|H(f)|=\prod_{n=1}^{N}|H_n(f)|$ (3)

– фаза $\varphi(f)=\sum_{n=1}^{N}\varphi_n(f)$ (4)

Приклади графіків модуля та фази для фонеми /я/ за залежностями (3) та (4) наведено на рис. 1 та 2.

Для приголосних звуків, так як в їх утворення приймає участь більше частин голосового тракту (за рахунок використання бокових порожнин або носового тракту) – додаються антиформанти, передавальна функція набуває вигляду [23] – [25]

$$H(f)=R_{rad}(f)\cdot S(f)\frac{\prod_{m=1}^{M_i}\left[\left(f_{a_{i,m}}^2-f^2\right)+j2\beta_{a_{i,m}}f\right]}{\prod_{n=1}^{N_i}\left[\left(f_{r_{i,n}}^2-f^2\right)+j2\beta_{r_{i,n}}f\right]} \quad (5)$$

де $R_{rad}(f)=C_{rad}\,j2\pi f$ – характеристика випромінювання губ/носа; $S(f)$ – джерело звуку (залежить від типу приголосного, див табл. 3); $f_{a_{i,m}}$ – частоти антиформантів (нулів) (500 – 2500 Гц (для носових , латеральних)); $f_{r_{i,n}}$ – частоти формант (резонансів); $\beta_{a_{i,m}}$ – коефіцієнти затухання антиформантів (50-300 Гц); $\beta_{r_{i,n}}$ – коефіцієнт затухання резонансів (30 – 200 Гц).

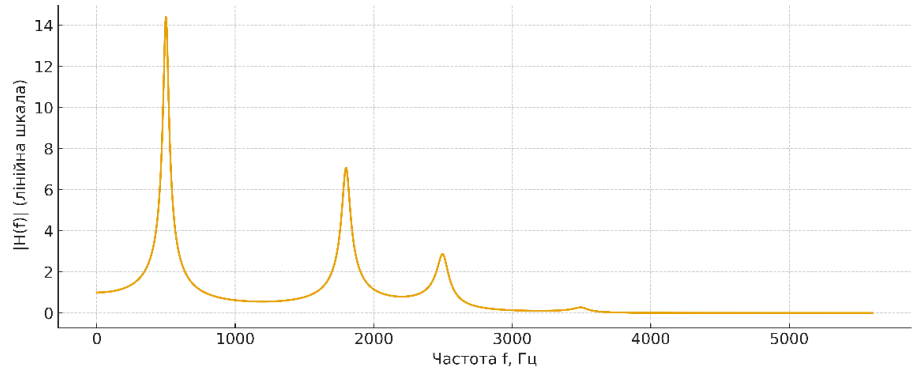

Рис. 1. Спектрограма модулю частотної характеристика H(f) для фонеми /я/

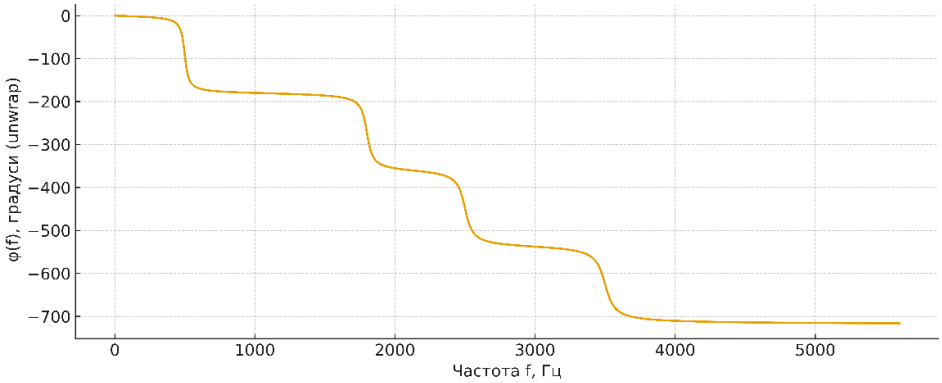

Рис. 2. Спектрограма фази частотної характеристика φ(f) для фонеми /я/

Після спрощення (5), отримуємо

$$H(f) = C_{rad}(j2\pi f)\cdot A_i f^{-p_i}\frac{\prod_{m=1}^{M_i}\left[\left(f_{a_{i,m}}^2 - f^2\right)+ j2\beta_{a_{i,m}} f\right]}{\prod_{n=1}^{N_i}\left[\left(f_{r_{i,n}}^2 - f^2\right)+ j2\beta_{r_{i,n}} f\right]} \tag{6}$$

де A_i – амплітудний коефіцієнт для типу приголосного; $p_i = \begin{cases} 0 & \text{для імпульсів (плозивні)} \\ 1-1{,}5 & \text{для фрикативних} \\ 2 & \text{для } voiced \end{cases}$; $M_i, N_i, f_{r_{i,n}}, f_{a_{i,m}}$ – множники, задаються в табл. 3.

За методикою (1) – (4) для залежності (6) знаходимо модуль та фазу для повної системи $H_n(f)$:

– модуль

$$|H(f)| = |C_{rad}|(2\pi f)\cdot|A_i|\, f^{-p_i}\frac{\prod_{m=1}^{M_i}\sqrt{\left(f_{a_{i,m}}^2 - f^2\right)+\left(2\beta_{a_{i,m}} f\right)^2}}{\prod_{n=1}^{N_i}\sqrt{\left(f_{r_{i,n}}^2 - f^2\right)+\left(2\beta_{r_{i,n}} f\right)^2}} \tag{7}$$

– фаза

$$\angle H(f) = \arg(C_{rad}) + \arg(j2\pi f) + \arg(A_i) - p_i \arg(f) + \sum_{m=1}^{M_i}\varphi\left(f; f_{a_{s,m}}, \beta_{a_{s,m}}\right) - \sum_{n=1}^{N_i}\varphi\left(f; f_{r_{s,n}}, \beta_{r_{i,n}}\right) \tag{8}$$

Приклади графіків модуля та фази для приголосної фонеми /н/ за залежностями (6) та (7) наведено на рис. 3 та рис. 4.

Таблиця 3

Параметри для типів приголосних української мови

Тип *i*	**Приклади**	**Джерело $S_{type}(f)$**	**Типові f_r (Гц)**	**Типові f_a (Гц)**	**Примітка**
Носові	/м, н/	*voiced* гармонічне джерело	[300, 1100, 2400]	[600, 2200]	Антиформанти виражені
Латеральні	/л/	*voiced*	[350, 1300, 2600]	[1000]	слабкий антиформант
Тремтячі	/р/	*voiced* (модуляції 20–30 Гц)	[400, 1300, 2500]	–	як сонорна, без нулів
Приблизні	/й, в/	*voiced*	[300, 1200, 2500]	–	гладка спектральна огина
Плозивні	/п, б, т, д, к, г/	короткий імпульс $S(f) = const$	[500, 1500, 3000]	–	широкосмугові *burst*-и
Фрикативні	/с, з, ш, ж, ф, в, х/	шум $S(f) \propto f^{-p}$ з максимумом у ВЧ	[1000, 2500, 4000]	–	$p = 1-1{,}5$; сибілянти мають підсилення 3–8 кГц
Африкатні	/ц, ч/	змішане: $S(f) = S_{burst} + S_{fric}$	[800, 2300, 3800]	–	комбінація *burst + noise*

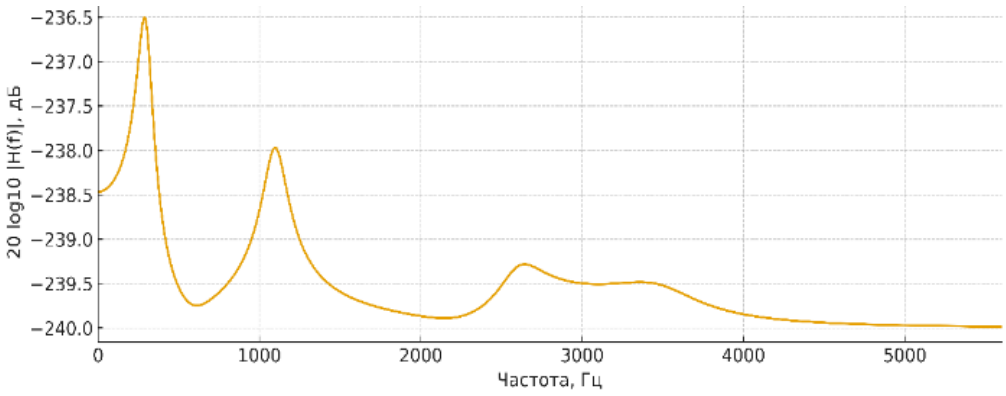

Рис. 3. Спектрограма модулю частотної характеристика H(f) для фонеми /н/

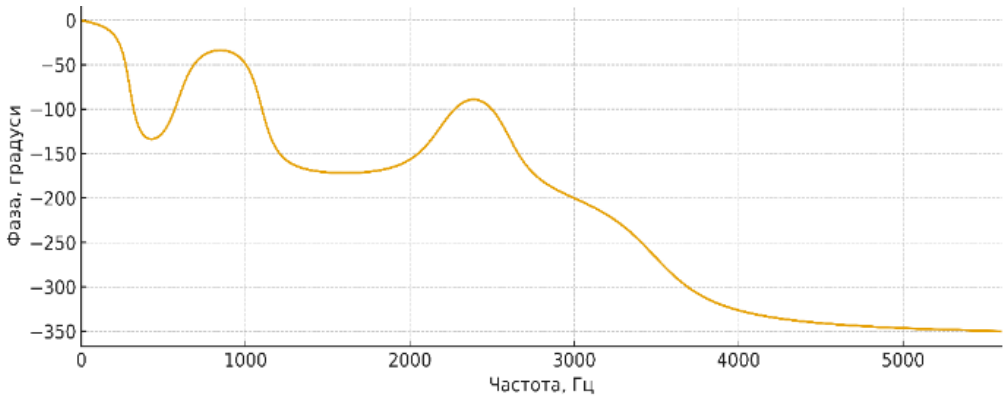

Рис. 4. Спектрограма фази частотної характеристика φ(f) для фонеми /н/

АЛГОРИТМ АНАЛІЗУ ТА ПРОГНОЗУ ФОНЕМ УКРАЇНСЬКОЇ МОВИ

КРОК 1. Первинний акустичний аналіз фонограми

1.1. Детекція мовної активності (VAD)

Варіант А (для чистих записів (SNR > 10 дБ)) – енергетичний метод [26]:

$$E[n] = \sum_{m=0}^{M-1} x^2[n-m],$$

$$VAD[n] = \begin{cases} 1, & \text{якщо } E[n] > \theta_E \text{ та } ZCR[n] < \theta_{ZCR} \\ 0, & \text{інакше} \end{cases}$$

де $E[n]$ – короткочасна енергія сигналу у кадрі n; $x[n]$ – миттєве значення сигналу; M – довжина вікна (у відліках); θ_E – поріг енергії; $ZCR[n]$ – частота нульових переходів; θ_{ZCR} – поріг частоти нульових переходів.

Варіант В (для шумових умов, з плаваючим порогом θ_H) – спектральна ентропія [26]:

$$P_k = \frac{|x_k|^2}{\sum_j |X_j|^2},$$

$$H = -\sum_k P_k \log(P_k),$$

$$VAD = \begin{cases} 1, & \text{якщо } H < \theta_H \text{ або } SNR > \theta_{SNR}, \\ 0, & \text{інакше.} \end{cases}$$

де P_k – нормована потужність спектральної компоненти k; X_k – комплексна амплітуда спектра; H – спектральна ентропія; θ_H – поріг ентропії; SNR – співвідношення сигнал/шум; θ_{SNR} – поріг SNR.

1.2. Сегментація мовлення та визначення пауз

Пауза фіксується, якщо VAD = 0 та тривалість мовчання $\Delta t > \theta_{pause}$. Типові налаштування: кадр 10–20 мс, мінімальна пауза 150–250 мс.

1.3. Формування тривимірної матриці сигналу

$$M(X,Y,Z), X = t, \Delta t = 0{,}5мс; Y = f, \Delta f = 1Гц; Z = A, A \in [0,0,8]$$

$$A_{norm}(t,f) = 0.8 \frac{A(t,f)}{max(A)}$$

де X – вісь часу; Y – вісь частоти; Z – амплітуда; A_{norm} – нормована амплітуда; Δt – часовий крок; Δf – частотний крок.

КРОК 2. Оцінка кількості фонів у слові

2.1. Початкова оцінка (використовувати для ініціалізації сегментації)

$$N_f \approx \frac{T_{слова}}{T_{фон}}, T_{фон} \in [50,100]мс$$

де N_f – очікувана кількість фонів; $T_{слова}$ – тривалість сигналу слова; $T_{фон}$ – середня тривалість одного фону.

2.2. Уточнення меж

$$\Delta F_i = F_{i+1} - F_i, \Delta E = E_{i+1} - E_i$$

де ΔF_i – зміна частоти форманти; ΔE – зміна енергії кадру; F_i – частота і-ї форманти.
Рекомендація: комбінувати з кроком 3.
КРОК 3. Спектральний аналіз першого фону
3.1. Обчислення спектра (FFT) [27], [28]

$$X[k] = \sum_{n=0}^{N-1} x[n] w[n] e^{\frac{-j2\pi kn}{N}}, f_k = k\frac{F_a}{N}$$

де x[n] – сигнал у часі; w[n] – віконна функція; N – довжина FFT; F_s – частота дискретизації; f_k – частота компоненти спектра.
Рекомендовані параметри: N = 2048–4096; F_s = 11025–22050 Гц.
3.2. Спектральна огинаюча (LPC/Cepstrum) [29]

$$A(z) = 1 + \sum_{i=1}^{p} a_i z^{-i}, H(z) = \frac{G}{A(z)};$$
$$c[n] = FFT^{-1}\{log|X[k]|\},$$

де a_i – коефіцієнти моделі; p – порядок LPC; G – нормувальний коефіцієнт.; c[n] – кепстральні коефіцієнти.
Рекомендація: $p \approx (F_s/1000) \times 3$.
3.3. Пошук максимумів і нормування

$$\frac{dS}{df} = 0, \quad \frac{d^2S}{df^2} < 0 \quad S_N(f) = 0.8\frac{S(f)}{max(S)}$$

де S(f) – спектральна щільність; $S_N(f)$ – нормована амплітуда; f – частота.
Результат: частоти формант F_1–F_4 (до F_7 для шумних фонів).
КРОК 4. Визначення можливих фонем за відомими максимумами
4.1. Байєсівська оцінка відповідності максимумів (для 1–2 максимумів (голосні, сонорні)) [30]

$$L_{k,i} = S(f_m) e^{-\frac{(f_m - F_{k,j})^2}{2\sigma_{k,j}^2}} \quad P(k|f_m) = \frac{\sum_i \pi_{k,i} L_{k,i}}{\sum_j \sum_i \pi_{j,i} L_{j,i}}$$

де $L_{k,i}$ – правдоподібність; $S(f_m)$ – амплітуда спектра; $F_{k,i}$ – еталонна частота; $\sigma_{k,i}$ – дисперсія; $\pi_{k,i}$ – апріорна вага.
4.2. Багатовимірна узгодженість формант (Махаланобіс), (при ≥ 2 піках.):

$$P(k|F_{obs}) \propto e^{-\frac{1}{2}(F_{obs} - \mu_k)^T \sum_k^{-1} (F_{obs} - \mu_k)}$$

де F_{obs} – вектор формант [F1,F2,F3,F4]; μ_k – середній вектор; Σ_k – коваріаційна матриця.

КРОК 5. Прогнозування наступних максимумів

5.1. Модель міжформантних інтервалів (при коартикуляційних переходах) [31], [32]:

$$f^{(k)}_{i\pm m} \approx F_{k,i} \pm \sum_{r=1}^{m} \Delta f_{i+r-1,i+r}, \Delta f \in [600,1300]\,Гц$$

де $F_{k,i}$ – середня частота форманти; Δf – відстань між сусідніми формантами; m – крок прогнозу.

Рекомендація: при неповному наборі піків.

5.2. Прогноз контекстних максимумів:

$$f_{pred} = f_{ods} \pm \Delta f_{mean} \cdot C_{ctx}$$

де Δf_{mean} – середнє зміщення формант; C_{ctx} – коефіцієнт контекстної взаємодії.

КРОК 6. Побудова трифонної моделі

6.1. Розрахунок умовної ймовірності P(Z | X, Y) (використовуйте корпус (словник) ≥ 100 000 слів з лемами української мови)

$$P(Z|X,Y) = \frac{C(X,Z,Y)}{\sum_{Z'} C(X,Z',Y)}$$

де C(X,Z,Y) – кількість спостережень трифона у корпусі; $\Sigma_{Z'}$ – сума частот усіх трифонів з тими ж *X* та *Y*; *P(Z|X,Y)* – умовна ймовірність центрального фону *Z* за контексту *X*,*Y*.

6.2. Нормування та фільтрація

Видаляються рідкі трифони (*P* < 0.0001), після чого залишені ймовірності нормуються так, щоб ∑*Z* = 1.

КРОК 7. Прямий прогноз (уперед)

7.1. Прогноз наступного фону Y (для прогнозування продовження фонемного ряду в словах)

$$P(Y|Z,X) = \frac{\sum_X P(Z|X,Y)P(X)}{\sum_{Y'}\sum_X P(Z|X,Y')P(X)}$$

де *P(Y|Z,X)*–ймовірність появи наступного фону *Y*; *P(Z|X,Y)* – умовна ймовірність трифона; *P(X)* – апріорна ймовірність попереднього фону.

7.2. Комбінований акустико-статистичний індекс (використовувати при наявності реальних спектральних даних)

$$K = exp\left[-\frac{(f_{real} - f_Y)^2}{2\sigma^2}\right], P^*(Y) = P(Y|Z,X)\cdot K$$

де f_{real} – спостережувана частота максимуму; f_Y – очікувана частота для фонеми Y; σ – стандартне відхилення; K – коефіцієнт збігу спектра; $P^*(Y)$ – скоригована ймовірність.

КРОК 8. Зворотний аналіз попереднього фону

8.1. Зворотне обчислення ймовірності (для реконструкції лівого контексту після ідентифікації центрального фону)

$$P(X|Z,Y) = \frac{P(Z|X,Y)\cdot P(X)}{\sum_{X'}\left[P(Z|X',Y)\cdot P(X')\right]}$$

де P(X|Z,Y) – ймовірність попереднього фону X за відомих Z і Y; P(Z|X,Y) – умовна ймовірність трифона; P(X) – апріорна ймовірність попереднього фону; ΣX′ – усі можливі попередні фони.

8.2. Очікувані частоти переходу X → Z (для оцінки коартикуляційних ефектів)

$$f_{exp} \approx \frac{f_x + f_Z}{2} \pm \sigma_f$$

де f_{exp} – очікувана частота максимуму на межі фонів *X* та *Z*; f_X, f_Z – середні формантні частоти; σ_f – стандартне відхилення.

КРОК 9. Перевірка максимумів для всіх фонів трифона

9.1. Гаусова оцінка збігу максимумів (оцінювання окремо для голосних, сонорних та шумних фонів)

$$w_j = exp\left[-\frac{(f_{obs,j}) - F_i)^2}{2\sigma}\right], K_{XYZ} = \frac{1}{N}\sum_j w_j,$$

$$R_{XYZ} = K_{XYZ}\cdot P(Z|X,Y)$$

де $f_{obs,j}$ – спостережувані частоти піків; F_i – еталонні формантні частоти; σ – середньоквадратичне відхилення; w_j – вага відповідності; N – кількість піків; K_{xyz} – коефіцієнт узгодженості; R_{xyz} – індекс ймовірнісно-акустичної відповідності.

9.2. Інтерпретація

– $R \approx 1$ – повна узгодженість;
– $R \approx 0.5$ – часткова, потрібна додаткова перевірка;
– $R < 0.2$ – ймовірно помилковий трифон.

КРОК 10. Перевірка кінця слова та ітерація

10.1. Умови завершення

– якщо $Y =$ «␣» → кінець слова;
– якщо $Y \neq$ «␣» → перехід до наступного фону.

10.2. Оновлення контексту

$$X' = Z, Z' = Y, Y' = Y_{next}$$

де *X′*, *Z′*, *Y′* – нові фони для наступного трифона; Y_{next} – прогнозований фон із Кроку 7.

10.3. Завершення циклу

Результати:

– послідовність фонем із ймовірностями R_{XYZ};
– часова структура слів;

– карта спектральних максимумів із ідентифікованими фонемами.

ВИСНОВКИ

У результаті проведеного дослідження вдосконалено алгоритм відновлення лінгвістичної складової мовної інформації фонемно-формантним методом, призначений для оцінювання рівня її захищеності від витоку. На основі поєднання спектрального, статистичного та контекстного аналізу запропоновано десятикрокову послідовність процедур, що забезпечує повний цикл — від виявлення фонем до прогнозування їхніх імовірностей у трифонних структурах.

У рамках **першого завдання** виконано комплексний аналіз впливу спектральної структури формант на розпізнавання фонем та встановлено, що зміщення максимумів, викривлення міжформантних інтервалів і зміни фазової складової істотно знижують достовірність акустичної ідентифікації, особливо у випадку дифонів і трифонів. Узагальнення моделей голосового тракту для голосних і приголосних (зокрема тих, що мають антиформанти) дозволило визначити критичні діапазони спектральної чутливості та обґрунтувати необхідність контекстно залежної реконструкції.

У межах **другого завдання** удосконалено алгоритм фонемно-формантного аналізу шляхом інтеграції процедур VAD, сегментації, FFT- і LPC-аналізу, багатовимірного оцінювання формант, прогнозування контекстних максимумів та побудови трифонної статистичної моделі. Уведено індекс ймовірнісно-акустичної відповідності R_xyz, який кількісно оцінює збіжність між спектральними максимумами та еталонними параметрами трифонів з урахуванням коартикуляцій та фазових зсувів, що забезпечує узгодженість реконструкції навіть за умов часткової втрати спектральної інформації.

Практичне значення отриманих результатів полягає у формуванні науково-технічного підґрунтя для побудови сучасних систем контролю витоку мовної інформації та інструментального оцінювання рівня захищеності мовних каналів. Розроблений алгоритм може бути інтегрований у засоби технічного захисту, використовуваний для експертної реконструкції зашумлених фонограм, а також застосований у задачах автоматичного аналізу, розпізнавання та синтезу мовлення, де потрібна підвищена стійкість до завад.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Allen, J. B. (1996). *Harvey Fletcher's role in the creation of communication acoustics. Journal of the Acoustical Society of America*, 99(4), Part 1.
2. Collard, J. (1929). *A theoretical study of the articulation and intelligibility of a telephone circuit. Electrical Communication: A Journal of Progress in the Telephone, Telegraph and Radio Art*, 7(3), 168–186.
3. French, N. R., & Steinberg, J. C. (1947). *Factors governing the intelligibility of speech sounds. Journal of the Acoustical Society of America*, 19(1), 90–119. https://doi.org/10.1121/1.1916407
4. Beranek, L. L. (1947). *Some notes on the measurement of acoustic impedance. Journal of the Acoustical Society of America*, 19(3), 420–427. https://doi.org/10.1121/1.1916499
5. Fletcher, H., & Galt, R. H. (1950). *The perception of speech and its relation to telephony. Journal of the Acoustical Society of America*, 22(2), 89–151.
6. Fletcher, H. (1950). *A method of calculating hearing loss for speech from an audiogram. Journal of the Acoustical Society of America*, 22(1), 1–5.
7. American National Standards Institute. (1997). *ANSI S3.5-1997: Methods for the calculation of the speech intelligibility index.* New York: Acoustical Society of America.
8. American National Standards Institute. (2007). *ANSI S3.5-2007: American National Standard — Methods for calculating the speech intelligibility index.* New York: Acoustical Society of America.

9. Anonymous. (1955). *Handbook of acoustic noise control* (CIA Report No. CIA-RDP81-01043R004000070005-1). Washington, D.C.: Central Intelligence Agency.
10. Blintsov, V., Nuzhniy, S., Parkhuts, L., & Kasianov, Y. (2018). *The objectified procedure and a technology for assessing the state of complex noise speech information protection. Eastern-European Journal of Enterprise Technologies*, 5(9 [95]), 26–34. https://doi.org/10.15587/1729-4061.2018.144146
11. Blintsov, V., & Nuzhniy, S. (2019). *Udoskonalennia metodu otsiniuvannia stupenia zakhystu movnoi informatsii [Improvement of the method for assessing the level of speech information protection]*. *Eastern-European Journal of Enterprise Technologies*, 6(9 [102]), 28–38. https://doi.org/10.15587/1729-4061.2019.185585
12. Kasianov, Yu. I., & Nuzhnyi, S. M. (2016). *Otsiniuvannia efektyvnosti heneratora realnoi movopodibnoi zavady za kryteriiem rozbirlyvosti movy [Evaluation of speech-like noise generator efficiency by speech intelligibility criterion]*. *Visnyk Natsionalnoho universytetu "Lvivska politekhnika": Avtomatyka, vymiriuvannia ta keruvannia*, 852, 105–110. https://ena.lpnu.ua/server/api/core/bitstreams/82888e90-c7c7-4af0-b6a5-e2def041dda7/content
13. Dudley, H. (1939). *The Voder: An electronic speech synthesizer.* Bell Laboratories.
14. Chiba, T., & Kajiyama, M. (1941). *The vowel: Its nature and structure.* Tokyo: Kaiseikan.
15. Stevens, K. N., & House, A. S. (1955). *Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America*, 27(4), 484–493.
16. Fant, G. (1960). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations.* The Hague: Mouton & Co.
17. Malinen, J. (2015). *Formants.* Aalto University. https://math.aalto.fi/~jmalinen/MyPSFilesInWeb/formants_OEL.pdf
18. Pirogov, A. A. (2001). *Osnovy foneticheskoi teorii rechi [Fundamentals of the phonetic theory of speech]*. *Zhurnal Russkogo fizicheskogo obshchestva (ZhRFM)*, 1–12, 15–28.
19. Lienard, J. S., et al. (1977). *Diphone synthesis of French: Vocal response unit and automatic prosody from the text. Proceedings of the International Congress on Acoustics, Speech and Signal Processing*, 560–563.
20. Ishchenko, O. S. (2008). *Akustychni kharakterystyky holosnykh zvukiv suchasnoi ukrainskoi literaturnoi movy [Acoustic characteristics of vowel sounds in modern Ukrainian literary language]*. *Ukrainska mova*, 4, 102–111.
21. Vakulenko, M. O. (2024). *Positional variations of Ukrainian back vowel formants. Proceedings of the International Conference on Modern Research in Social Sciences*, 1(1), 1–12. https://doi.org/10.33422/icmrss.v1i1.301
22. Vakulenko, M. O. (2018). *Ukrainski holosni zvuky v konteksti MFA [Ukrainian vowels in the context of IPA]*. *Govor*, 35(2), 189–214.
23. Fant, G. (1971). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations.* Hague – Berlin: Mouton / Walter de Gruyter.
24. Kent, R. D. (1993). *Vocal tract acoustics. Journal of the Acoustical Society of America*, 94(5), 2603. https://doi.org/10.1121/1.408664
25. Coleman, J. (2006). *Acoustic structure of consonants.* Oxford: University of Oxford.
26. Rabiner, L. R., & Shafer, R. W. (1978). *Tsifrova obrobka movnykh syhnaliv [Digital processing of speech signals]*. Kyiv: Prentis-Hol. https://doi.org/10.5555/5404
27. Fant, G. (1960). *Akustychna teoriia utvorennia movlennia [Acoustic theory of speech production]*. Stockholm: Mouton. https://archive.org/details/acoustictheoryofspeechproduction
28. Stevens, K. N. (1998). *Acoustic phonetics.* Cambridge: MIT Press. https://mitpress.mit.edu/9780262193980
29. Markel, D. D., & Gray, A. H. (1976). *Linear prediction of speech.* New York: Springer. https://link.springer.com/book/10.1007/978-3-642-66242-4
30. Ishchenko, O. S. (2003). *Akustychni kharakterystyky holosnykh zvukiv suchasnoi ukrainskoi movy [Acoustic characteristics of Ukrainian vowels]*. Kyiv: Instytut movoznavstva NAN Ukrainy.
31. ND TZI 1.6-005-2013. (2013). *Metodyka vyznachennia kharakterystyk kanaliv vytoku movnoi informatsii [Methodology for determining characteristics of speech information leakage channels]*. Kyiv: ADIT. https://cip.gov.ua/ndtzi
32. ND TZI 3.7-003-2023. (2023). *Metodyka otsiniuvannia efektyvnosti zasobiv zakhystu vid vytoku movnoi informatsii [Methodology for evaluating the effectiveness of protection means against speech information leakage]*. Kyiv: DSSZZI.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (27), 2025
ISSN 2663 – 4023

Serhii Nuzhnyi
Candidate of Technical Sciences, Associate Professor,
Head of the Department of Computer Technologies and Information Security
Admiral Makarov National University of Shipbuilding, Mykolaiv, Ukraine
ORCID: 0000-0002-7706-0453
s.nuzhniy@gmail/com

IMPROVEMENT OF THE ALGORITHM FOR RESTORING THE LINGUISTIC COMPONENT OF SPEECH INFORMATION USING THE PHONEME–FORMANT METHOD FOR ASSESSING ITS SECURITY LEVEL

Abstract. The article presents an improved algorithm for restoring the linguistic component of speech information using the phoneme–formant method in tasks of assessing its protection level. A comprehensive analysis of the historical development of articulation and formant approaches is conducted — from the classical studies of Bell Laboratories and Harvey Fletcher to modern digital techniques of speech intelligibility evaluation established in ANSI S3.5-1997 and ANSI S3.5-2007 standards. It is shown that the evolution of speech-signal analysis methods has directly influenced the formation of information protection systems, particularly in monitoring channels of acoustic leakage. A ten-step algorithm for speech signal analysis and reconstruction is proposed, integrating spectral, statistical, and contextual approaches. The algorithm includes speech activity detection (VAD), signal segmentation, formant extraction using Fourier and linear predictive analysis (FFT, LPC), and construction of a probabilistic triphone model. The consideration of phase shifts, coarticulation effects, and inter-formant intervals ensures high reconstruction accuracy and robustness even under low signal-to-noise ratio (SNR) conditions. Special attention is given to analyzing the influence of spectral distortions of formants on phoneme recognition and determining statistical regularities of diphone and triphone occurrence in Ukrainian speech. This enabled the creation of an adaptive model combining acoustic and linguistic features for more reliable restoration of the linguistic component of a message. The developed algorithm can be applied in technical information protection practice for objective evaluation of speech-channel security and for recovering damaged or noisy recordings. Its structure fully complies with Ukrainian national standards ND TZI 1.6-005-2013 and ND TZI 3.7-003-2023, which regulate the methodology for analyzing speech-information leakage channels and assessing protection efficiency. The obtained results have dual practical significance: they can serve as a foundation for expert systems of speech information leakage control and for systems of automatic speech recognition, synthesis, and restoration requiring high phonemic precision. The proposed phoneme–formant approach demonstrates universality and the ability to integrate with modern digital audio-analysis tools, thus improving both information security and the performance of acoustic control systems.

Keywords: phoneme–formant analysis; spectral reconstruction; triphone model; coarticulation; speech information security

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Allen, J. B. (1996). *Harvey Fletcher's role in the creation of communication acoustics. Journal of the Acoustical Society of America*, 99(4), Part 1.
2. Collard, J. (1929). *A theoretical study of the articulation and intelligibility of a telephone circuit. Electrical Communication: A Journal of Progress in the Telephone, Telegraph and Radio Art*, 7(3), 168–186.
3. French, N. R., & Steinberg, J. C. (1947). *Factors governing the intelligibility of speech sounds. Journal of the Acoustical Society of America*, 19(1), 90–119. https://doi.org/10.1121/1.1916407
4. Beranek, L. L. (1947). *Some notes on the measurement of acoustic impedance. Journal of the Acoustical Society of America*, 19(3), 420–427. https://doi.org/10.1121/1.1916499
5. Fletcher, H., & Galt, R. H. (1950). *The perception of speech and its relation to telephony. Journal of the Acoustical Society of America*, 22(2), 89–151.
6. Fletcher, H. (1950). *A method of calculating hearing loss for speech from an audiogram. Journal of the Acoustical Society of America*, 22(1), 1–5.

7. American National Standards Institute. (1997). *ANSI S3.5-1997: Methods for the calculation of the speech intelligibility index.* New York: Acoustical Society of America.
8. American National Standards Institute. (2007). *ANSI S3.5-2007: American National Standard — Methods for calculating the speech intelligibility index.* New York: Acoustical Society of America.
9. Anonymous. (1955). *Handbook of acoustic noise control* (CIA Report No. CIA-RDP81-01043R004000070005-1). Washington, D.C.: Central Intelligence Agency.
10. Blintsov, V., Nuzhniy, S., Parkhuts, L., & Kasianov, Y. (2018). *The objectified procedure and a technology for assessing the state of complex noise speech information protection. Eastern-European Journal of Enterprise Technologies*, 5(9 [95]), 26–34. https://doi.org/10.15587/1729-4061.2018.144146
11. Blintsov, V., & Nuzhniy, S. (2019). *Improvement of the method for assessing the level of speech information protection. Eastern-European Journal of Enterprise Technologies*, 6(9 [102]), 28–38. https://doi.org/10.15587/1729-4061.2019.185585
12. Kasianov, Yu. I., & Nuzhnyi, S. M. (2016). *Evaluation of speech-like noise generator efficiency by speech intelligibility criterion. Visnyk Natsionalnoho universytetu "Lvivska politekhnika": Avtomatyka, vymiriuvannia ta keruvannia*, 852, 105–110.
13. Dudley, H. (1939). *The Voder: An electronic speech synthesizer.* Bell Laboratories.
14. Chiba, T., & Kajiyama, M. (1941). *The vowel: Its nature and structure.* Tokyo: Kaiseikan.
15. Stevens, K. N., & House, A. S. (1955). *Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America*, 27(4), 484–493.
16. Fant, G. (1960). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations.* The Hague: Mouton & Co.
17. Malinen, J. (2015). *Formants.* Aalto University. https://math.aalto.fi/~jmalinen/MyPSFilesInWeb/formants_OEL.pdf
18. Pirogov, A. A. (2001). *Osnovy foneticheskoi teorii rechi [Fundamentals of the phonetic theory of speech]. Zhurnal Russkogo fizicheskogo obshchestva (ZhRFM)*, 1–12, 15–28.
19. Lienard, J. S., et al. (1977). *Diphone synthesis of French: Vocal response unit and automatic prosody from the text. Proceedings of the International Congress on Acoustics, Speech and Signal Processing*, 560–563.
20. Ishchenko, O. S. (2008). *Acoustic characteristics of vowel sounds in modern Ukrainian literary language. Ukrainska mova*, 4, 102–111.
21. Vakulenko, M. O. (2024). *Positional variations of Ukrainian back vowel formants. Proceedings of the International Conference on Modern Research in Social Sciences*, 1(1), 1–12. https://doi.org/10.33422/icmrss.v1i1.301
22. Vakulenko, M. O. (2018). *Ukrainski holosni zvuky v konteksti MFA [Ukrainian vowels in the context of IPA]. Govor*, 35(2), 189–214.
23. Fant, G. (1971). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations.* Hague – Berlin: Mouton / Walter de Gruyter.
24. Kent, R. D. (1993). *Vocal tract acoustics. Journal of the Acoustical Society of America*, 94(5), 2603. https://doi.org/10.1121/1.408664
25. Coleman, J. (2006). *Acoustic structure of consonants.* Oxford: University of Oxford.
26. Rabiner, L. R., & Shafer, R. W. (1978). *Tsifrova obrobka movnykh syhnaliv [Digital processing of speech signals]*. Kyiv: Prentis-Hol. https://doi.org/10.5555/5404
27. Fant, G. (1960). *Akustychna teoriia utvorennia movlennia [Acoustic theory of speech production]*. Stockholm: Mouton. https://archive.org/details/acoustictheoryofspeechproduction
28. Stevens, K. N. (1998). *Acoustic phonetics.* Cambridge: MIT Press. https://mitpress.mit.edu/9780262193980
29. Markel, D. D., & Gray, A. H. (1976). *Linear prediction of speech.* New York: Springer. https://link.springer.com/book/10.1007/978-3-642-66242-4
30. Ishchenko, O. S. (2003). *Acoustic characteristics of Ukrainian vowels.* Kyiv: Instytut movoznavstva NAN Ukrainy.
31. ND TZI 1.6-005-2013. (2013). */-Methodology for determining characteristics of speech information leakage channels.* Kyiv: ADIT. https://cip.gov.ua/ndtzi
32. ND TZI 3.7-003-2023. (2023). *Methodology for evaluating the effectiveness of protection means against speech information leakage.* Kyiv: DSSZZI.