

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. І. СІКОРСЬКОГО»

Факультет електроніки

Кафедра акустики та акустoeлектроніки

«На правах рукопису»  
УДК \_531.134\_\_\_\_\_

«До захисту допущено»  
Завідувач кафедри  
\_\_\_\_\_ Дідковський В.С.

\_\_\_\_\_

(підпис) (ініціали, прізвище)

“\_9\_”\_12\_\_\_\_\_2019\_\_р.

**Магістерська дисертація**

зі спеціальності (спеціалізації) 171 "Електроніка"

(код та назва спеціальності)

на тему: Дослідження акустичних особливостей інтегральних характеристик української мови.

Виконав: студент 2-ого курсу магістратури, групи ДГ-81мп

Глюк Ілля Романович \_\_\_\_\_

(прізвище, ім'я, по батькові)

(підпис)

Керівник доцент, к.ф-м.н., Луньова С. А. \_\_\_\_\_

(посада, науковий ступінь, вчене звання, прізвище та ін..ініціали)

(підпис)

Рецензент доцент каф. ЗТіРІ, к.т.н. Швайченко В.Б. \_\_\_\_\_

(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Засвідчую, що у цій дипломній роботі немає запозичень з праць інших авторів без відповідних посилань.

Студент \_\_\_\_\_

(підпис)

**Національний технічний університет України**  
**«Київський політехнічний інститут імені Ігоря Сікорського»**

Інститут/факультет Факультет електроніки

Кафедра акустики та акустоелектроніки

Рівень вищої освіти – другий (магістерський) за освітньо-професійною програмою

Спеціальність (спеціалізація) 171 «Електроніка»

ЗАТВЕРДЖУЮ  
Завідувач кафедри  
\_\_\_\_\_ Дідковський В.С.

\_\_\_\_\_

(підпис)

(ініціали, прізвище)

«\_9\_»\_12\_\_\_\_\_2019 р.

## ЗАВДАННЯ

### на магістерську дисертацію студенту

Глюку Іллі Романовичу

(прізвище, ім'я, по батькові)

1. Тема дисертації: Дослідження акустичних особливостей інтегральних характеристик української мови.

науковий керівник дисертації: Луньова Світлана Андріївна, к.ф-м.н., доцент ,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «07» 11. 2019р. №3859– С

2. Строк подання студентом дисертації: 9.12.2019.

3. Об'єкт дослідження: усереднені спектри західногерманської, романської, західнослов'янської і східнослов'янської мовних груп.

4. Вихідні дані для магістерської дисертації: мовні записи текстів художнього та публіцистичного стилю мов західногерманської, романської,

західнослов'янської і східнослов'янської груп; частотний діапазон 100 – 20 000Гц.

5. Перелік завдань, які потрібно розробити: фонологічні відмінності української мови, об'єктивні та інтегральні характеристики мови, експериментальне дослідження спектрального складу мов західногерманської, романської, західнослов'янської і східнослов'янської груп.

6. Перелік графічного (ілюстративного) матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо): презентація в Microsoft Power Point обсягом не менше 10 слайдів.

7. Орієнтовний перелік публікацій: А. Гарасюк, І. Глюк, С. Луньова «Акустичні особливості усереднених характеристик української мови» (знаходиться в редакції н.-т. журналу «Мікросистеми, електроніка та акустика»).

8. Дата видачі завдання 8.04.2019 року.

#### Календарний план

№ з/п	Назва етапів виконання дипломного проекту (роботи)	Строк виконання етапів дисертації	Примітка
1	Пошук літератури	15.04.2019-29.04.2019	
2	Аналіз літературних джерел	29.04.2019-20.05.2019	
3	Аналіз фонологічних особливостей української мови	20.05.2019-3.05.2019	
4	Аналіз способів обробки мовних сигналів	3.05.2019-17.06.2019	
5	Пошук і обробка мовних записів найпоширеніших мов Європи	9.09.2019-11.11.2019	
7	Оформлення дипломної роботи та підготовка презентації	11.11.2019-9.12.2019	

Студент

Глюк І.Р.

(підпис)

\_\_\_\_\_ (ініціали, прізвище)

Науковий керівник  
дисертації

Луньова С.А.

(підпис)

\_\_\_\_\_ (ініціали, прізвище)

## АНОТАЦІЯ

Дипломна робота складається зі вступу, 4 розділів, висновків, списку використаної літератури. Загальний обсяг дипломної роботи складає 91 сторінок. Робота містить рисунків 27, таблиць 18, кількість додатків 3.

Метою даної роботи є дослідження усереднених спектрів мов індоєвропейської родини, яка найбільш поширена в Європі, а саме: західногерманської, романської, західнослов'янської і східнослов'янської груп.

У сучасному світі все більше поширюються системи керування голосом різного роду процесами, починаючи від голосового відповідача, «розумного дому» та до управління безпілотними літальними апаратами тощо. Нині більшість таких систем створенні для англійської мови та декілька для російської мови, тому важливим є питання визначення головних особливостей української мови у порівнянні з іншими європейськими мовами, які потрібно врахувати при створенні даних систем.

Проведено аналіз інтегральних характеристик мовного сигналу, розглянутий короткочасний частотний аналіз, миттєвий та середній спектри потужності мови.

**Ключові слова:** *артикуляція, фонема, спектр формант, спектр мови, спектральна інтенсивність, середній спектр потужності.*

## SUMMARY

The graduation work consists of introduction, 4 chapters, conclusions and the list of used literature. Overall volume of the graduate work is 91 pages. The work has 27 figures, 18 spreadsheet, number of applications 3.

The purpose of this paper is to investigate the averaged spectrum of languages of the Indo-European family, which is most common in Europe, namely: West Germanic, Romance, West Slavic and East Slavic groups.

In the modern world, voice control systems of all kinds are increasingly distributed, ranging from the voice responder, the "smart home", and to the control of drones etcetera. Currently, most of these systems are designed for the English language and some for the Russian language, so it is important to determine the main features of the Ukrainian language compared to other European languages, which should be taken into account when creating these systems.

The analysis of the integral characteristics of the speech signal, the short-term frequency analysis, the instantaneous and average power spectra of the speech are considered.

**Key words:** *articulation, phoneme, formant specter, language specter, spectral intensity, average capacity factor.*

## ЗМІСТ

ВСТУП .....	7
РОЗДІЛ 1. ФОНОЛОГІЯ УКРАЇНСЬКОЇ МОВИ .....	8
1.1. Акустична природа мови.....	8
1.2. Утворення голосних і приголосних звуків .....	12
1.3. Фонемний склад української мови.....	18
РОЗДІЛ 2. УСЕРЕДНЕНІ ХАРАКТЕРИСТИКИ УКРАЇНСЬКОЇ МОВИ .....	20
2.1. Об’єктивні параметри мовного сигналу .....	20
2.2. Інтегральні характеристики мовного сигналу .....	21
2.3. Спектральний аналіз мови .....	26
РОЗДІЛ 3. ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ УСЕРЕДНЕНИХ СПЕКТРІВ ЗАХІДНОГЕРМАНСЬКОЇ, РОМАНСЬКОЇ, ЗАХІДНОСЛОВ’ЯНСЬКОЇ, СХІДНОСЛОВ’ЯНСЬКОЇ ГРУП МОВ ....	47
3.1. Умови проведення дослідження.....	47
3.2. Алгоритм обробки сигналів .....	49
3.3. Аналіз усереднених спектрів мови.....	49
РОЗДІЛ 4. РОЗРОБКА СТАРТАП-ПРОЕКТУ.....	59
ВИСНОВКИ .....	72
ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ .....	75
ДОДАТКИ.....	77

## ВСТУП

Питання мовної комунікації є надзвичайно нагальним у сучасному світі. Це те безпосереднє спілкування громадян, в тому числі різних країн.

На сьогодні все більшого поширення набувають системи управління голосом різного роду процесами, починаючи від голосового відповідача, керування різноманітними господарськими функціями, так званого, «розумного дому», і до управління безпілотними літальними апаратами та більш складними комплексами.

Більшість таких систем розроблена на англійській мові. Частково, із найбільш наближених до нас мов, існують програми, виконані російською або польською мовами.

Існує проблема з'ясування основних акустичних відмінностей української мови від споріднених мов східнослов'янської групи та інших західноєвропейських мов.

Найбільш об'єктивною акустичною характеристикою мови є її інтегральна характеристика, а саме усереднена спектральна густина потужності даної мови. Перша оцінка такої характеристики для української мови виконана в роботі [1].

За схемою проведення досліджень [1] аналізуються спектральні характеристики чотирьох груп індоєвропейської родини мов, до якої належать європейські мови, а саме: західногерманської, романської, західнослов'янської та східнослов'янської.

Порівняння спектральних характеристик європейських мов з відповідними характеристиками української мови дозволить встановити

закономірності та розбіжності, й таким чином полегшити задачу розробникам систем голосового управління.

## **РОЗДІЛ 1. ФОНОЛОГІЯ УКРАЇНСЬКОЇ МОВИ**

### **1.1. Акустична природа мови**

Мова – це система поєднання звуків, які несуть в собі певне значення і зміст, вона є засобом спілкування людей у суспільстві. За допомогою мови формується мислення у процесі пізнання людьми об’єктивної дійсності, здійснюється обмін думками, життєвим досвідом, закріплюються, зберігаються й передаються наступним поколінням набуті суспільством знання, виражаються найрізноманітніші людські почуття.

У мові можна виділити дві основні особливості. Перша полягає в тому, що це акустичний сигнал який може бути представлений у вигляді звукових хвиль. Тобто незалежно від того чи рідна якась конкретна мова для людини, яка її слухає чи ні, мова сприймається як і будь-яка інша акустична інформація - шум, свист тощо; тому аналіз мови проводиться як аналіз випадкового процесу. Друга особливість враховує те, що знайома (рідна) для людини мова спочатку сприймається як фонетична, а потім як семантична інформація. У цьому випадку важливим є дослідження фонем (звук мови, що типізується; найменша акустична одиниця мовного сигналу) і знаходження ознак їхньої ідентифікації.

Розглянемо процеси утворення звуків в українській мові.

### **1.2. Утворення голосних і приголосних звуків**



Формування значної кількості звуків відбувається за участі голосу, тобто за участю голосових зв'язок.

*Голосоутворення (фонація)* – процес, пов'язаний з вібраціями голосових зв'язок, коливання яких призводять до зміни площі голосової щілини і, відповідно, до модуляції потоку повітря, що видихається.

*Фонетика* – це розділ мовознавства, що вивчає звукову будову мови: утворення та акустичні властивості звуків мовлення і закономірності їх поєднання.

За умовами творення та особливостями звучання всі звуки поділяються на голосні та приголосні. В українській мові 36 фонем (6 голосних, 32 приголосні), в англійській – 44 (20 голосних: 8 чистих, 12 дифтонгів; 24 приголосних).

#### • Артикуляційно-акустичні умови творення голосних

*Голосними* називаються звуки, що складаються тільки з голосу. При їхньому утворенні повітря повільно проходить через ротову порожнину і не зустрічає жодних перешкод (*a, o, y, u, i, e*). При утворенні голосних звуків завжди виникає модуляція потоку повітря у результаті коливань голосових зв'язок – фонації. Тому ця ознака не може використовуватись для їх класифікації.

Місце *артикуляції* голосних визначається позицією спинки язика в горизонтальному напрямі щодо піднебіння. Якщо при вимові звука язик більше чи менше піднімається до твердого піднебіння, лишаючи порівняно невеликий прохід для повітря, яке видихається, а в задній частині ротової порожнини виникає великий простір, що служить резонатором, — утворюються голосні переднього ряду (*[i], [u], [e]*). Якщо ж при вимові звука основна маса язика зосереджена в його задній частині й піднімається до заднього піднебіння, утворюються голосні заднього ряду (*[a], [o], [y]*). При цьому в задній частині ротової порожнини залишається тільки невеликий

прохід, а великий резонатор утворюється в передній частині. Середнє положення між голосними переднього і заднього ряду займають голосні середнього ряду.

*Ступінь підняття язика* залежить від того, як спинка язика піднімається до піднебіння. Вона може бути максимально піднята до піднебіння, залишаючи невеликий прохід для струменя повітря, що видихається; може бути піднята менше і, нарешті, може підніматися незначною мірою або й не підніматися зовсім. Залежно від цього в українській мові розрізняють голосні високого (*[i], [y]*), високо-середнього (*[u]*), середнього (*[e], [o]*) і низького (*[a]*) підняття.

Рухи язика по вертикалі звичайно пов'язані з рухами щелепи: при опусканні щелепи опускається разом з під'язиковою кісткою спинка язика; при підніманні щелепи піднімається і спинка язика. Отже, відкритість і закритість голосних завжди пов'язані із взаємодією цих двох активних органів.

За *ступенем відкритості ротової порожнини* голосні поділяються на широкі й вузькі, або відкриті й закриті. До широких (відкритих) належать голосні низького й середнього підняття, до вузьких (закритих) – голосні високо-середнього та високого підняття. У межах чотирьох ступенів підняття кожен голосний також може утворюватися при більш або менш відкритій ротовій порожнині [2].

Розрізняють також голосні *напружені* та *ненапружені*. В українській мові більш напружені всі наголошені голосні і менш напружені – ненаголошені. Закриті голосні (*[i], [y], [u]*) – більш напружені, відкриті (*[a], [o], [ε]*) – менш напружені.

Губи легко змінюють форму й розмір вихідного отвору в ротовій порожнині, а також подовжують або скорочують резонуючу ротову порожнину. Залежно від їх участі в артикуляції голосні поділяються на неогублені, або нелабіалізовані, і огублені, або лабіалізовані.

До *лабіалізованих* належать [o], [y], до *нелабіалізованих* —[a], [e], [u],[i]. Якщо в утворенні звука беруть участь лише порожнини рота й глотки, то утворюються ротові (або неносові, неназалізовані) голосні. При цьому м'яке піднебіння підняте вгору й закриває вхід у носову порожнину. Однак при вимові голосного звука воно може бути й вільно опущене. У такому разі частина повітря виходить через носову порожнину й утворює носовий призвук. Так утворюються носові (або назалізовані) голосні. В українській мові носових голосних немає [2].

У багатьох мовах використовуються складні голосні, в яких відбувається плавний перехід від одного типу артикуляції до іншого. Якщо беруть участь дві голосні, звук називається *дифтонгом*. Наприклад, «я» [йа], «ю» [йу]; в англійській мові date [ei], file [ai], now [au]. Існують також *трифтонги*: hour [aue] [4].

#### • Артикуляційно-акустичні умови творення приголосних

*Приголосними* називаються звуки, що мають у своєму складі голос і шум або тільки шум. При їх утворенні повітря зустрічає на своєму шляху перешкоди ([б, п, д/д', т/т', г, к, т, ф, ж, з/з', ш, с/с', х, дж, дз/дз', ц/ц', ч, в, м, й, н/н', л/л', р/р']).

Якість приголосних залежить від умов їх творення. При утворенні приголосних струмів повітря, що видихається (з участю або без участі голосу) наштовхується на своєму шляху на перепону, яка становить зближення або повне зімкнення активних і пасивних мовних органів. Унаслідок подолання відповідної перепони виникає характерний шум, що визначає якість того чи іншого приголосного. При утворенні приголосних загальне напруження мовних органів відсутнє, воно зосереджене лише в ділянках, що є фокусом творення звука.

В українській мові приголосні класифікуються за такими ознаками:

1) за участю голосу й шуму в їх творенні;

- 2) за місцем творення (за активним і за пасивним мовним органом);
- 3) за способом творення;
- 4) за наявністю або відсутністю пом'якшення (палаталізації);
- 5) за наявністю або відсутністю носового забарвлення (назалізації).

Акустичні особливості приголосних звуків, як правило, залежать від артикуляційних умов їх творення. За участю голосу й шуму приголосні поділяють на сонорні й шумні [3].

*Сонорними* (від лат. *sonorus*– звучний, голосний) називаються приголосні, що складаються з голосу й шуму з перевагою голосу, наприклад, [р],[в],[л], [н] тощо. Під час артикуляції сонорних звуків мовні органи хоча й зближуються в основному таким самим способом, як і при творенні шумних приголосних, проте у цьому разі виникає досить широка щілина і повітряний струмінь, проходячи крізь неї, не створює сильного шуму, властивого шумним приголосним.

*Шумними* називаються приголосні, що складаються з голосового тону (або голосу) й шуму з перевагою шуму або з одного тільки шуму, чітко локалізованого. До шумних в українській мові належать: [б], [п], [д], [т],[з], [ш] тощо. Шумні, в свою чергу, поділяються на дзвінкі й глухі [3].

Коли повітряний струмінь, утворивши голосовий тон, зустрічає в ротовій порожнині щілину або зімкнення, виникають дзвінкі приголосні. Коли ж повітря вільно проходить у надгортанні порожнини через достатньо розкриту голосову щілину й зустрічає в порожнині рота або глотки вузьку щілину чи повне зімкнення, де утворюються специфічні для кожного приголосного шуми, виникають глухі приголосні.

Класифікація за місцем творення звука ґрунтується на дії активних і пасивних мовних органів при утворенні фокуса його артикуляції. За участю активних мовних органів приголосні звуки поділяються передусім на губні, язикові та глоткові або фарингальні. Губні приголосні творяться в результаті зближення або зімкнення активного органа – нижньої губи з верхньою губою або верхніми зубами. Язикові приголосні залежно від того, яка частина язика

торкається піднебіння, поділяються на передньоязикові, середньоязикові і задньоязикові.

*Фарингальні* творяться в результаті зближення задньої стінки глотки з коренем язика. За участю пасивних мовних органів губні приголосні поділяють на дві підгрупи: губно-губні (наприклад [b], [m]) та губно-зубні (наприклад [f]). Передньоязикові приголосні за цією ознакою також поділяють на дві підгрупи: зубні, тобто передня спинка язика і його кінчик зближуються з верхніми різцями (наприклад при [n], [c]) та піднебінно-зубні, або альвеолярні (передній край язика зближується з альвеолами верхніх зубів, наприклад ([p], [ʒ])).

Класифікація приголосних за способом утворення шуму враховує характер перепони, що утворюється між активними й пасивними мовними органами.

За способом творення шумні приголосні поділяють на зімкнені, ротові й щілинні. Шумні зімкнені називаються ротовими тому, що при їхньому творенні повітря не потрапляє до носової порожнини. Вони, в свою чергу, поділяються на зімкнені чисті й африкати. За вимови чистих зімкнених (наприклад [b], [m], [k]) між активним і пасивним мовним органом утворюється тісне зімкнення. Під тиском повітряного струменя зімкнення з силою розривається, внаслідок чого виникає специфічний для відповідної артикуляції вибух. Чисті зімкнені називаються також проривними (вибуховими, моментальними, зімкнено-проривними).

Під час вимови африкат (наприклад [tʃ], [dʒ]) активний мовний орган, наближаючись до пасивного, спочатку утворює зімкнення, однак замість моментального прориву в африкат зімкнення поступово переходить у коротку щілину. Отже, *африкати* – це звуки складної артикуляції, першим компонентом якої є зімкнення, а другим — щілина. Африкати називають також зімкнено-щілинними звуками. Однак відповідні звуки не можна розглядати як звичайну суму зімкненого і щілинного. Це особливі злиті звуки [4]. Для артикуляції африкат необхідні дві умови: обидва компоненти мають

утворюватися одним і тим самим мовним органом; обидва компоненти мають бути або дзвінкими, або глухими. Під час вимови щілинних активний мовний орган, наближаючись до пасивного, утворює посередині вузьку щілину. Струмінь повітря, що видихається, проходячи крізь щілину, утворює специфічний для цих звуків шум. Шумні щілинні за місцем утворення щілини називаються серединними.

Сонорні приголосні за способом творення можуть бути зімкнені носові, щілинні і дрижачі. Сонорні зімкнені називаються носовими тому, що під час їхньої вимови в ротовій порожнині утворюється тісне зімкнення, як у [ɒ], і одночасно опускається м'яке піднебіння, лишаючи вільний прохід через носову порожнину. Так утворюються зімкнені носові [м], [я]. Сонорні щілинні різні за характером і місцем утворення щілини поділяються на серединні і бокові. За вимови серединних щілина утворюється між губами або посередині язика між його крилами, притисненими до піднебіння по боках [4]. Під час вимови бокових сонорних ([л], [л']) у ротовій порожнині утворюється зімкнення між кінчиком язика й передніми зубами, але одночасно бокові краї язика опущені, звільняючи прохід для струменя повітря, яке видихається. При вимові дрижачих, або вібрантів, ([р']), кінчик язика періодично або притискується до альвеол, або відходить від них, утворюючи своєрідне дрижання чи вібрацію.

Приголосні розрізняють за наявністю або відсутністю палаталізації (або за твердістю і м'якістю). Палаталізація звуків, або пом'якшення, виникає внаслідок додаткового підняття передньої і середньої спинки язика в напрямі до твердого піднебіння. Залежно від відсутності чи наявності роботи передньо-середньої спинки язика приголосні поділяються на тверді й м'які. У більшості тверді й м'які приголосні утворюють співвідносні пари.

М'які приголосні за ступенем вияву м'якості, в свою чергу, поділяють на три типи:

1) власне м'які приголосні, при вимові яких середня спинка язика торкається твердого піднебіння;

- 2) пом'якшені, або палаталізовані;
- 3) напівпом'якшені.

### 1.3. Фонемний склад української мови

Оскільки фонетика вивчає звукові одиниці мовлення у зв'язку з їхньою роллю в утворенні значущих одиниць мови, то її акустичною основою є аналіз звуків мовлення. Тобто фізична інтерпретація відповідних акустичних явищ і процесів. Без цього інші аспекти фонетики такі, як артикуляційний та перцептивний просто не мають належного сенсу.

У звуковому аналізі мовлення поєднується фізика, біологія, психологія, медицина, техніка, кібернетика та мовознавство; ця галузь науки розвивається дуже швидко. При цьому мовленнєві та мовні технології спираються на формальні моделі або представлення мовного знання на фонологічному, фонетичному, морфологічному, синтаксичному, семантичному, прагматичному та дискурсивному рівнях. Отже, фонетика потребує комплексного науково-технічного підходу.

Стосовно особливостей звукоаналізу української мови слід зазначити декілька моментів. По-перше, на відміну від поширених мов “етимологічного” спрямування, правопис яких базується на традиційному написанні – це передусім англійська і французька, – українська мова має виразну фонетичну засаду. А це вимагає особливо ретельного дослідження її звуків. По-друге, в українській мові, на відміну від поширеної англійської, граматичні форми слів дуже мінливі: про це свідчить, зокрема, наявність семи відмінків іменників. Для комп'ютера слова з різними закінченнями – це фактично різні слова. Така особливість значно ускладнює завдання моделювання та розпізнавання української мови. Екстенсивний підхід, заснований на нехитрому запам'ятовуванні цілих словесних форм, для української (як і для інших слов'янських мов) непридатний: це поставило б насамперед не виправдано великі вимоги до комп'ютерної пам'яті. Флективні мови змінливою граматичною структурою слів – такі, як

українська – потребують розпізнавання та моделювання на нижчих фонологічних чи фонетичних рівнях, ніж словоформи: фонемному або складовому. А для цього необхідно мати чітке уявлення про фундаментальні елементи мовлення.

Створення акустичної бази даних для комп'ютерного моделювання мовлення (генерації звукових сигналів) та його розпізнавання (ідентифікації звукових сигналів) потребує докладної інформації про фізичну суть звуків мовлення та їх нелінійну взаємодію [4]. Відомо, що інформація про звук міститься не лише в ньому самому, а розкидана по декількох сусідніх і навіть віддалених звуках, і що міжзвукові переходи містять важливу інформацію для розпізнавання мовлення. Цю інформацію використовують на стадії попереднього аналізу даних.

Глибоке розуміння процесів, які зумовлюють формування фонем, відкриває можливість створення пристрою, здатного розпізнати будь-який потік мовлення незалежно від віку, статі та емоційного стану мовця – такого собі «комп'ютерного секретаря». З іншого боку, належне розуміння впливу психологічних чинників на звукові параметри дозволило б створити надійні детектори брехні, визначники захворювань і звукові замки з вибірковою реакцією на певний емоційний стан. Цілком імовірно, колись ми станемо спроможними точно змоделювати вимову будь-якої особи та відтворити голоси відомих співаків минулого. До того ж, цікавою є можливість доповнення комп'ютерних словників звуковим супроводом [4]. Всі ці завдання

так чи інакше торкаються багатьох науково-технічних проблем, зокрема, створення штучного інтелекту.

Особливість української мови – її милозвучність у порівнянні з російською. Це вимагає свідомого уникання незграбності в поєднанні звуків.

Засоби милозвучності української мови:

- 1) спрощення у групах приголосних;
- 2) велика кількість голосних звуків;



- 3) дзвінки приголосні не озвучуються;
- 4) побічний наголос, як наслідок наближення до поезії;
- 5) голосні не втрачають якісних і кількісних показників.

Особливістю мови є *українська лексика*, тобто вживання слів, які не належать іншим мовам. Наприклад: *кваша, куняти, спочивати, повінь, нишпорити*. Є слова, які утворені від спільних з іншими мовами слів та запозичених коренів. Власне українськими є слова, що позначають збірні назви з подовженими приголосними основи: *гілля, волосся, насіння, бадилля*.

Специфічними для українського словотвору є складні прізвища і загальні назви з першим компонентом у формі наказового способу дієслів: *Котигорошко, Нетийвода, Горицвіт, Перекотиполо, пройдисвіт*.

Аналогічно, *експресивно-оцінні дієслова* на зразок: *«питоньки, спатоньки, купоньки»*, у структурі яких наявні іменникові суфікси. Специфічно українськими є дієприслівними з суфіксами: *-учи/ючи, -ачи/ячи, -чи, -єши*, а саме: *читаючи, стежачи, роблячи, тремтячи, ідучи, вивчивши, купивши, оцінивши*. А також суфікси *-інь, -оц, -енк* (*далечінь, солодоці, поштаренко*).

Особливості українського словотвору виявляються у використанні префіксів: *перед-, під-, попід-, зне-, по-*, наприклад: *передбачувати, підказувати, попідписувати, знецінювати*. Особливе повторення префікса *по-*: *попоїли, попождеш*. Показовим є те, що цих форм не підтримують ані школа, ані ЗМІ, однак вони поширені в Україні повсюдно і послідовно регулярні. Слід сказати і про такі префікси: *якнай-, щонай-* (*якнайкращий, щонайгірше*).

Така ознака як повноголосся: *-оро-, -оло-, -ере-, -еле-* (*ворог, сором, береза, пелена*). В українській мові немає редуції: наголошені і ненаголошені голосні звучать однаково напружено (як-от в італійській): *перемикачами, непередбачуваного*.

Власне українськими є:

- прислівники і модальні слова прислівникового походження (*звичайно, зазвичай, крадькома, відтак, либонь, ймовірно, безумовно*); також пестливі форми прислівників (*теперечки, тутечки, недалечко, поволеньки*);
- прийменники (*біля, від, задля, коло, з-під, з-поміж*);
- сполучники (*та, але, аби, попри, проте, якщо, хоч, мовби, ніби, позаяк*); сполучник *що* у синтаксичних структурах: *книжки, що по них послано; земля, що він її продав*;
- займенники з префіксоїдами *будь-, казна-* (*будь-хто, будь-коли, казна-куди*);
- вигуки (*агов! ой леле! а нумо! матінко! лишенько!*);
- складні слова: *густошерстний, грушеподібний, блакитноокий, жовто-блакитний, землевласник*.

Поряд зі звичайним аналітичним майбутнім (*буду робити, будеш писати*), унікальною серед слов'янських мов морфологічною рисою української є синтетичний майбутній час дієслів недоконаного виду (*робитиму, писатимеш*), за аналогією до романських мов.

Українській не властиве «акання», на відміну від російської і ще більшою мірою білоруської.

Однією із найяскравіших рис української мови у фонологічних приголосних є розвиток зі старого проривного «г» (*гурт, голова*) фарингального щілинного приголосного «г» (*грунт, гава, ганок*).

Збереження старого індоєвропейського *-p* на кінці складу (*матір*), також він переважно твердий (*гірко, буквар*). У нашій мові став розвиненим протетичний «в» (*вуха, вуса, вулиця*), який наявний також у білоруській, польській і серболужицькій; цікаві фонемі «дз», «дж» (*дзвін, джміль*). Здебільшого відсутнє цокання (з фінського субстрату): *чапля, чіпляти*. Якщо

порівнювати з російською мовою, то у родовому відмінку української односкладових слів *лоб, рот* збережено -о- : *лоба, рта* ( рос. *лба, рта*).

Відсутність редукції єднає українську з усіма слов'янськими, крім російської та словенської; відсутність цокання поєднує українську з усіма слов'янськими, крім російської, білоруської та польської (частково); відсутність акання – є спільною рисою української з усіма слов'янськими, крім білоруської та російської. Історичний збіг «і», «ы», «и» єднає українську з усіма західнослов'янськими мовами, ствердіння губних і «р»– з усіма південнослов'янськими, наявність фарингального «г»(h) – з центрально-слов'янськими.

Отже, розглянувши різні відмінності та збіги української мови з російською, можна зазначити, що остання літературна мова має аналогію з українською приблизно у 20 відсотків фонологічних рис. Решта 80 відсотків це відмінність від російської; з них 25 відсотків є унікальними, властивими виключно українській мові. При порівнянні української мови з англійською, французькою, італійською і словенською, згідно наведених даних, бачимо ще меншу кількість спільних фонологічних рис.

## РОЗДІЛ 2. УСЕРЕДНЕНІ ХАРАКТЕРИСТИКИ МОВИ

### 2.1. Об'єктивні параметри мовного сигналу

Доцільно розглядати мовний сигнал у двох варіаціях. Це звичайний *акустичний сигнал*, що є процесом розповсюдження звукової хвилі в середовищі. До об'єктивних параметрів, яким характеризується сигнал можна віднести:

- часову структуру звукової хвилі;
- тривалість звучання;
- спектральний склад;
- місце розташування джерела в просторі та інші параметри, які викликають слухові суб'єктивні відчуття (гучності, висоти, тембру, локалізації, маскування та ін.).

А також мова призначена для передачі змістовної інформації. Тому в її акустичних характеристиках містяться фонетичні та *семантичні ознаки*, які аналізуються у зонах Брока та Верніке кори головного мозку.

Зв'язок між *акустичними* характеристиками сигналу і *фонетичними* ознаками вивчається для розпізнавання і синтезу мови, тобто реакцію як мозок здобуває інформацію про змістовність за зміненням звукового тиску.

Для обробки мовного сигналу за рівнеграмами звукового тиску  $P(t)$ , дБ (де  $P(t)$  – функція зміни звукового тиску в часі) проводять *статистичний* аналіз, *кореляційний* та *спектральний* аналізи.

*Статистичними методами* можливо дослідити:

- розподіл у часі миттєвих значень і рівнів мовного сигналу;

- розподіл миттєвої і середньої потужності;
- динамічний діапазон і пік-фактор;
- розподілення основної фонаційної частини;
- спектральний розподіл формант.

Оскільки мовний сигнал є *випадковим* процесом, то величини, що його характеризують, неперервно змінюються у часі. Але вухо не реагує на окремі миттєві значення акустичних коливань. Слухове сприйняття зумовлене *енергією* впливу, усередненою за деякий проміжок часу.

*Кореляційний метод* є логічним продовженням таких методів як аналітичне групування, дисперсійний аналіз і зіставлення паралельних рядів. У поєднанні з цими методами він надає статистичному аналізу завершений характер. Кореляційний аналіз дає змогу виміряти ступінь впливу факторних ознак на результативні, встановити єдину міру тісноти зв'язку і роль досліджуваного фактора у загальній зміні результативної ознаки. Цей метод дозволяє одержати кількісні характеристики ступеня зв'язку між двома і більшим числом ознак, а тому дає більш широке уявлення про зв'язок між ними.

## 2.2. Інтегральні характеристики мовного сигналу

Велику роль в техніці мають так звані *інтегральні характеристики мови*, які визначають властивості мовного процесу загалом. До них відносяться:

- спектр мови;
- спектр формант;
- амплітудний склад мови;
- відносна наявність формант по частотному спектру.

Спектр мови:

Рівень інтенсивності мови весь час коливається. Але, якщо спостерігати мовний сигнал протягом тривалого часу, то можна визначити доволі стійке середнє значення цього рівня. Середні рівні інтенсивності мови різні в різних частотних смугах. Зазвичай інтенсивність мови в смузі частот, яка нас цікавить, характеризують так званим *спектральним рівнем мовлення*. Це рівень енергії, який припадає на смугу шириною в 1Гц. Наприклад, якщо середня інтенсивність звуку протягом тривалого часу в смузі частот шириною  $\Delta f$  рівна  $I_{cp} \left[ \frac{Вт}{см^2} \right]$ , то *спектральна інтенсивність* буде рівна

$I_1 = \frac{I_{cp}}{\Delta f}$ , а спектральний рівень визначається :

$$B_p = 10 \lg \frac{I_1}{I_0}, \text{дБ} \quad (2.1)$$

де  $I_0 = 10^{-16} \left[ \frac{Вт}{см^2} \right]$  – інтенсивність нульового рівня, відносно якого можна обчислити рівні всіх інших звуків.

Якщо відома не інтенсивність, а звуковий тиск  $p_1$ , віднесений до 1 Гц, то маючи на увазі, що  $\frac{I_1}{I_0} = \left( \frac{p_1}{p_0} \right)^2$ , вираз для спектрального рівня мови можна піднести до виду:

$$B_p = 20 \lg \frac{p_1}{p_0}, \quad (2.2)$$

де  $p_0 = 2 \cdot 10^{-4}$  бар.

*Спектром мови* називається залежність середнього, протягом тривалого часу, спектрального рівня мови  $B_p$  від частоти  $f$ . Вигляд спектру мови представлений на рис.3. Ця характеристика показує розподіл енергії мови за різними частотами.

Відмінність між піковим рівнем мови і її середнім, протягом тривалого періоду, рівнем називається *пiкфактором мови*.

### Спектр формант:

Мова містить в собі форманти, сприйняття яких визначає її розбірливість і неформантні складові, до яких відносять основні тони, області частот між формантами та складові, які залежать від індивідуальних властивостей дикторів.

Оскільки розбірливість мови визначається лише формантами, то варто виділити та побудувати їхній спектр окремо. Спектром формант будемо називати залежність найімовірнішого, протягом тривалого часу, спектрального рівня формант ( $B'_p$ ) від частоти  $f$ . Вид спектру формант представлений також на рис.2.1.

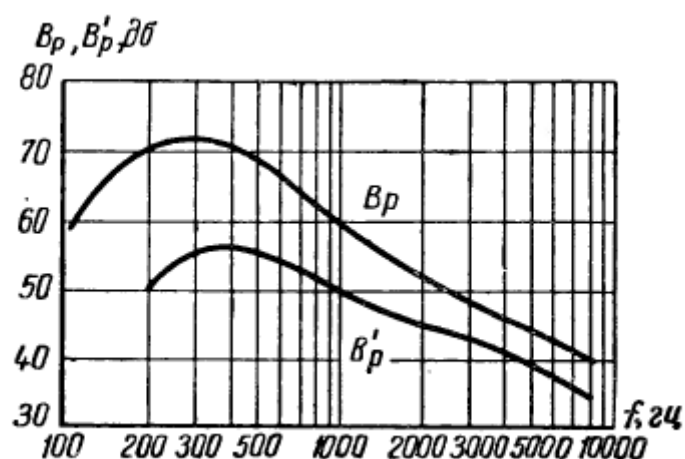


Рис.2.1. Спектр мови і спектр формант

Як бачимо, крива рівня формант лежить нижче кривої рівня мови. Варто мати на увазі відмінність між вище вказаними поняттями. В спектральному рівні мови визначається середній, протягом тривалого часу, енергетичний рівень мови  $B_p$ , а в спектральному рівні формант має на увазі найбільш імовірне значення рівня формант  $B_p'$ . Остання величина відрізняється від середнього рівня формантних складових, визначити який було б дуже не просто через те, що сам процес виокремлення формант з реального мовленнєвого потоку складний.

Оскільки значення  $B_p'$  відповідає моді розподілу енергії формант за рівнями, то ця величина може бути більш точно названа *середнім модальним спектральним рівнем формант*, протягом тривалого часу.

#### Амплітудний склад мови:

Якщо спостерігати появу різних енергетичних рівнів мови загалом чи в окремих частотних смугах, то можна встановити відносно часту появу того чи іншого рівня. Частіше всього складові мови будуть мати рівні близькі до середнього, а дуже потужні або, навпаки, більш тихі складові, з'являтимуться відносно рідко.

Характеристикою амплітудного складу мови прийнято називати залежність відносного числа випадків ( $z$ ), у яких збільшується той чи інший рівень, від величини цього рівня ( $B$ ). На рис.2.2. наведена така характеристика в загальному випадку.



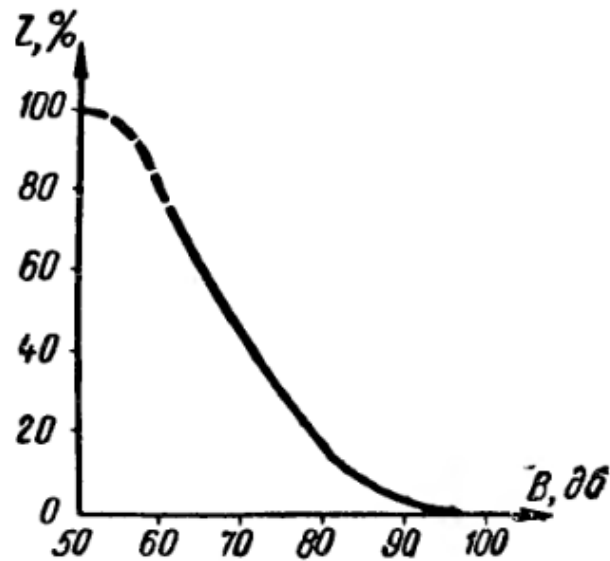


Рис.2.2. Характеристика амплітудного складу мови

Не складно помітити, що частота збільшення того чи іншого рівня з достатнім наближенням подібна нормальному закону розподілу імовірності (крива Гауса в інтегральній формі).

Характеристики були отримані шляхом амплітудного аналізу мови в обмежених смугах за допомогою електричних фільтрів. Амплітудний аналіз мови в смузі проводився шляхом розподілу мовного процесу на однакові часові інтервали, а потім підрахунком відносного числа ( $z$ ) тих інтервалів, в яких збільшувався той чи інший заданий рівень ( $B$ ).

Тривалість часових інтервалів вимірювання була вибрана з урахуванням середньої тривалості звуків мови та інтегрованої здатності вуха. Розгляд даних про тривалість звуку мовлення і властивостей вуха сумувати в деякому часовому інтервалі гучність звуку, приводить до вибору тривалості інтервалів вимірювання  $T = 0,125c$ .

У якості величини, яка визначає інтенсивність мови в кожному інтервалі та її слухове сприйняття, вимірювалось значення квадрату звукового тиску,

доданого в границях заданого проміжку, тобто величини  $\int_0^T p^2 dt$ . Результати

таких вимірювань дають можливість визначити рівень інтенсивності мови, який переважає, та середнє значення цього рівня за тривалий проміжок часу.

Відносна поява формант:

Форманти мови розподілені по частотному спектру нерівномірно. В низькочастотних областях, формант доволі мало, в деякій середній частині спектру, зосереджена їхня основна частина, а в області високих частот, знову ж таки, наявність зменшується. Якщо розбити весь спектр на смуги шириною 100 Гц і визначити відносне відсоткове число формант, яке знаходиться в кожній смугі, то отримаємо криву, яка відображена на рис.2.3.

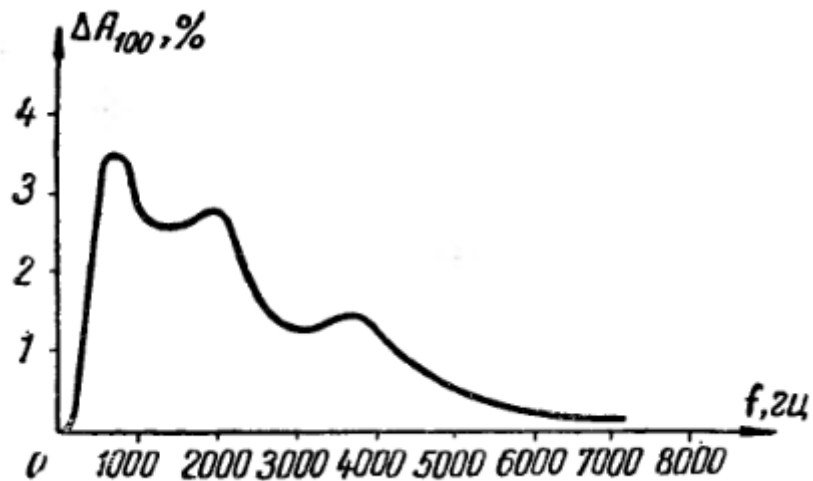


Рис.2.3. Відносна поява формант

Дана крива дозволяє судити про частку, яку та чи інша область частот вносить у загальний склад формант і про її важливість для забезпечення розбірливості мови.

Важливий момент, що звуковий склад національних мов дуже різноманітний, тому і фізичні характеристики відрізняються. Наприклад, російська мова, яка в свою чергу набагато більше досліджена, ніж українська, характеризується відносно великим коефіцієнтом високих частот порівняно з англійською. З фонетичної точки зору це пояснюється тим, що в російській мові велика кількість пом'якшених приголосних, які і дають в результаті

високочастотні складові в спектрі. Ще більш виражена роль високих частот в польській мові за рахунок чималої кількості шиплячих звуків.

### 2.3. Спектральний аналіз мови

#### • Короткочасний частотний аналіз

Вид мовної інформації в частотній області має ряд переваг. По-перше, акустичний аналіз механізму мовотворення показує, що розподіл нулів і полюсів коефіцієнта передачі дає достатньо чіткий опис звуків мовлення. По-друге, очевидно, що в початковій стадії процесу сприйняття, вухо виконує грубий частотний аналіз. Таким чином, характерні особливості, які з'являються в результаті частотного аналізу, відіграють важливу роль в процесах сприйняття і відтворення.

Встановлено також, що механізм створення мови є квазістаціонарним джерелом звуків зі змінним в часі характером збудження і частотною характеристикою. Тому будь-яка спектральна характеристика, яка підходить для мовного сигналу, має відображати не тільки спектральні особливості сприйняття значень, але й часові особливості.

Звичайний математичний зв'язок між аперіодичною функцією часу  $f(t)$  і її комплексним спектром щільності амплітуд  $F(\omega)$  виражається перетворенням Фур'є:

$$\left. \begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \\ f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega \end{aligned} \right\} \quad (2.3)$$

Аби перетворення відбувалось, вираз  $\int_{-\infty}^{\infty} |f(t)| dt$  повинен мати кінцеву величину. Неперервний мовний сигнал не задовольняє попередню умову та

умову визначеності на нескінченному інтервалі часу. З цього витікає, що сигнал має бути перетворений таким чином, аби його перетворення виконувалось для випадку, коли інтегрування відбувається в інтервалі відомих минулих величин. Крім цього, щоб відобразити важливі часові зміни, інтегрування слід виконувати тільки за такий відрізок часу, який відповідає тривалості квазівстановлених елементів мовного сигналу. Бажано мати спектр з реальним часом в якості незалежної змінної, де для визначення спектру використовуються вагові попередні значення сигналу.

Такого роду результат можна отримати за рахунок аналізу деякої частини сигналу, яку ми розглядатимемо через особливе часове вікно чи вагову функцію. Вагова функція обирається таким чином, щоб її взаємодія з сигналом давала функцію, яка буде забезпечувати умову дійсності перетворення Фур'є. Для практичного значення вагова функція  $h(t)$  являє собою імпульсний відгук фізично реалізованої лінійної системи. В такому випадку для  $t < 0, h(t) = 0$ . Зазвичай прагнуть щоб  $h(t)$  була уніполярною і являла собою відгук фільтру низьких частот. Отже, перетворення Фур'є можна видозмінити так, щоб перетворенню піддавалась та частина сигналу, яка в даний момент часу розглядається через вікно. Аналітично бажаний процес можна виразити так:

$$F(\omega, t) = \int_{-\infty}^{\infty} f(\lambda)h(t - \lambda)e^{-i\omega\lambda} d\lambda$$

або

$$F(\omega, t) = e^{-i\omega t} \int_0^{\infty} f(t - \lambda)h(\lambda)e^{i\omega\lambda} d\lambda \quad (2.4)$$

За визначенням, доцільний варіант також оберненого співвідношення:

$$[f(\lambda)h(t - \lambda)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega, t)e^{i\omega\lambda} d\omega$$

Сигнал з його попередніми значеннями, зважені за допомогою  $h(t)$ , для даного моменту  $t$  представлений на рис.2.4. Перетворення за короткий проміжок часу, визначене таким чином, є згорткою  $[f(t)e^{-i\omega t} \cdot h(t)]$  або, за вибором,  $e^{-i\omega t} [f(t) \cdot h(t)e^{-i\omega t}]$ . Якщо припустити, що вагова функція  $h(t)$  має розмірність  $[c^{-1}]$ , тобто вважати перетворення Фур'є від  $h(t)$  безрозмірним, тоді  $|F(\omega, t)|$  буде являти собою миттєвий спектр амплітуд з тою ж розмірністю, яку має сигнал. Так, як і звичайне перетворення Фур'є  $F(\omega, t)$  є, в загальному випадку, комплексною величиною з амплітудою і фазою, а саме  $|F(\omega, t)|e^{i\mathcal{G}(\omega, t)}$ , де  $\mathcal{G}(\omega, t)$  і є миттєвим спектром фаз.

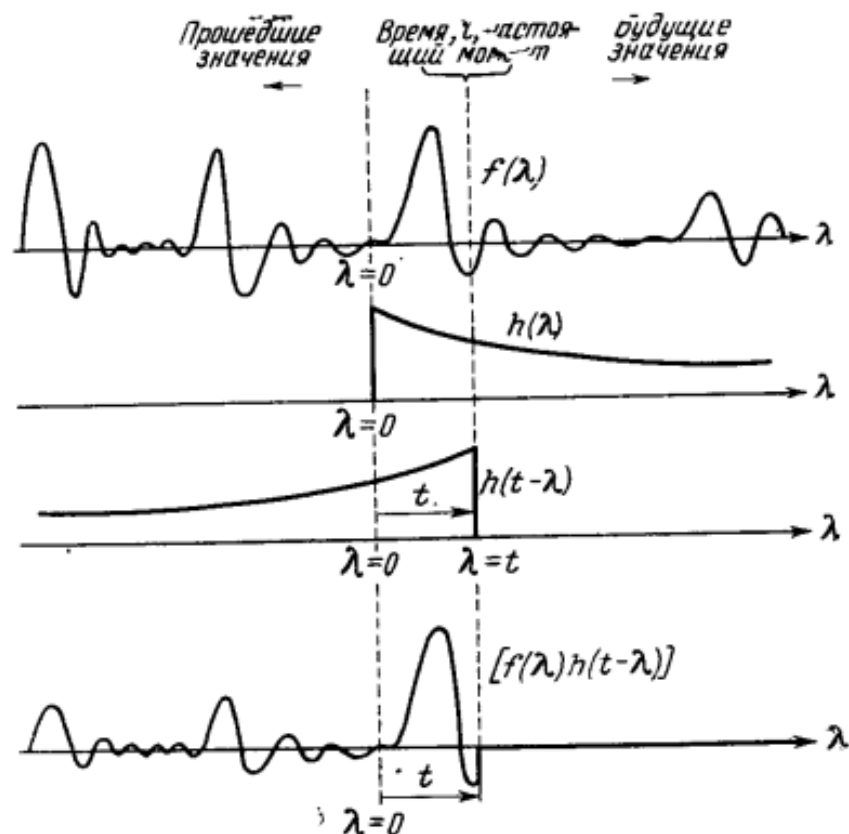


Рис.2.4. «Зважування» неперервного сигналу  $f(t)$  за допомогою фізично реалізованого часового вікна  $h(t)$ .  $\lambda$  – змінна інтегрування, яка

тимчасово введена для виконання перетворення Фур'є у довільний момент часу  $t$ .

- **Вимірювання миттєвого спектру**

Вираз (2.4) можна подати в такому вигляді:

$$F(\omega, t) = \int_{-\infty}^t f(\lambda) \cos \omega \lambda h(t - \lambda) d\lambda - i \int_{-\infty}^t f(\lambda) \sin \omega \lambda h(t - \lambda) d\lambda = [a(\omega, t) - ib(\omega, t)] \quad (2.5)$$

Крім цього

$$|F(\omega, t)| = [F(\omega, t)F^*(\omega, t)]^{\frac{1}{2}} = (a^2 + b^2)^{\frac{1}{2}} \quad (2.6)$$

$$\text{і } \vartheta(\omega, t) = \arctg \frac{b}{a},$$

де  $F^*(\omega, t)$  – величина комплексного спряжена з  $F(\omega, t)$ . Видно, що  $|F(\omega, t)|$  є скалярною величиною, тоді як  $F(\omega, t)F^*(\omega, t)$  – формально величина комплексна, а  $|F(\omega, t)|^2$  є миттєвим спектром потужності. З цього випливає, що вимірювання  $|F(\omega, t)|$  може бути виконане за допомогою функціональної схеми, яка представлена на рис.2.5.

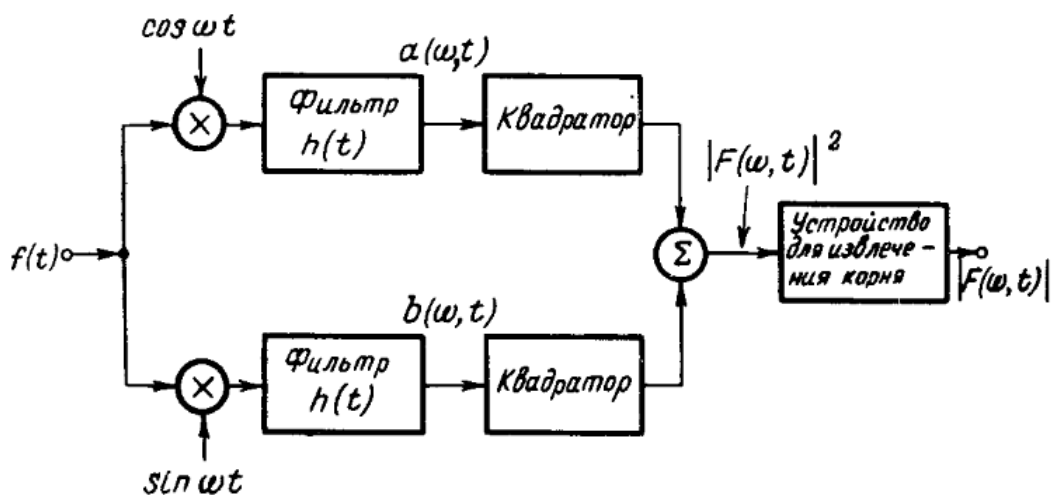


Рис.2.5. Вимірювання миттєвого амплітудного спектру  $|F(\omega, t)|$

Частотна інтерпретація цих операцій не викликає труднощів. Зміщений спектр функцій  $f(t)$  за допомогою множення на  $\cos \omega t$  і  $\sin \omega t$ , потрапляє у смугу прозорості фільтра  $h(t)$ , який, зазвичай, є фільтром нижніх частот. Частотні складові функції  $f(t)$ , які знаходяться у безпосередній близькості від частоти  $\omega$ , в результаті биття з цією частотою утворюють різні компоненти, які і вносять основний вклад у величину сигналу на виході фільтра  $h(t)$ . Обидва сигнали, які знаходяться у квадратному співвідношенні, підносяться до квадрату та додаються, у результаті чого утворюється миттєвий спектр потужності  $|F(\omega, t)|^2$ .

Вираз (2.4) можна також записати у вигляді

$$F(\omega, t) = e^{-i\omega t} \left\{ \int_0^{\infty} f(t - \lambda) h(\lambda) \cos \omega \lambda d\lambda + i \int_0^{\infty} f(t - \lambda) h(\lambda) \sin \omega \lambda d\lambda \right\} = \quad (2.7)$$

$$= [a'(\omega, t) + ib'(\omega, t)] e^{-i\omega t}$$

Відповідний для цього запису спосіб вимірювання  $|F(\omega, t)| = [a'^2 + b'^2]^{\frac{1}{2}}$  можна здійснити за допомогою функціональної схеми, представленої на рис.2.6.

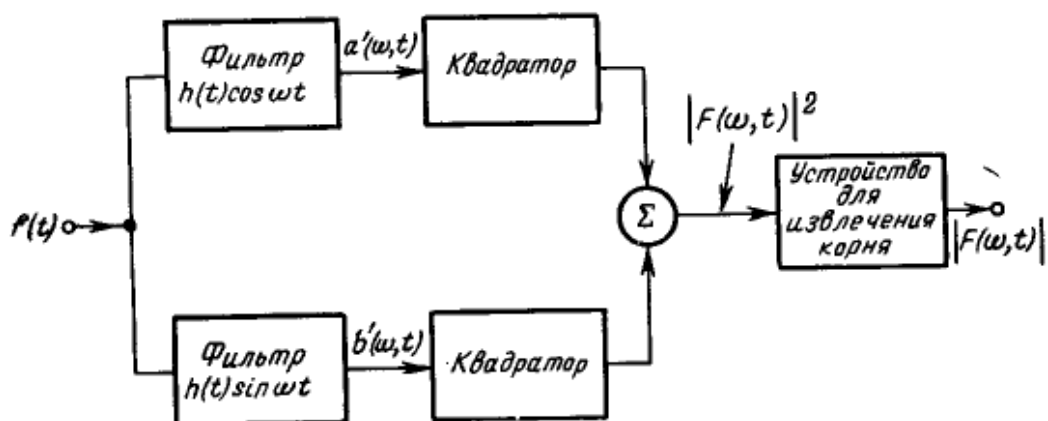


Рис.2.6. Другий спосіб вимірювання миттєвого амплітудного спектру

$$|F(\omega, t)|$$

У наведеному способі вимірювання застосовується фільтрація за допомогою смугових фільтрів зі взаємно зсунутими на  $\frac{\pi}{2}$  фазовими характеристиками та з симетричними відносно  $\omega$  частотними характеристиками, які мають вдвічі більшу смугу пропускання порівняно з характеристикою фільтра низьких частот  $h(t)$ .

Сигнали, які надходять з обох фільтрів, підносяться до квадрату та додаються, утворюючи миттєвий спектр потужності  $|F(\omega, t)|^2$ . Обидва фільтра мають імпульсні відгуки, обвідні яких являють собою часове вікно  $h(t)$ . Тобто потрібно стільки пар фільтрів, скільки відліків миттєвого спектру треба визначити. Варто відзначити, що для обох способів вимірювання, представлених на рисунках (2.5) і (2.6), миттєвий спектр одиничного імпульсу  $f(t)$  виражається просто функцією  $h(t)$ , тобто ваговою функцією.

На практиці за можливості спрощують апаратуру шляхом приблизного виконання вимірювань, які наведені на попередніх двох рисунках. Функція,

яку ми досліджуємо  $|F(\omega, t)| = \left[ a'^2(\omega, t) + b'^2(\omega, t) \right]^{\frac{1}{2}}$  є часовою обвідною функції  $a'(\omega, t)$  або  $b'(\omega, t)$ . Часова обвідна функції  $u(t)$ , для якої є

перетворення Фур'є, може бути визначена як  $e(t) = \left[ u^2(t) + \hat{u}^2(t) \right]^{\frac{1}{2}}$ , де

$\hat{u}(t) = u(t) \cdot \frac{1}{\pi t}$  є перетворенням Гільберта від  $u(t)$ .

Використовуючи ці співвідношення і змінюючи порядок інтегрування в згортці, отримаємо:

$$\left. \begin{aligned} a'(\omega, t) &= \left[ f(t) * h(t) \cos \omega t \right] \\ \hat{a}'(\omega, t) &= \left[ a'(\omega, t) \cdot \frac{1}{\pi t} \right] = f(t) \cdot \left[ h(t) \cos \omega t \cdot \frac{1}{\pi t} \right] = f(t) \cdot \left[ h(t) \sin \omega t \right] = b'(\omega, t) \end{aligned} \right\} (2.8)$$

При умові, що частота  $\omega$  не потрапляє у спектр  $h(t)$ . Таким чином, величина  $|F(\omega, t)|$  дійсно є часовою обвідною  $a'(\omega, t)$  або  $b'(\omega, t)$   $[a(\omega, t)]$ , або



$b(\omega, t)$ ]. Цю обвідну можна електрично апроксимувати формою обвідної на виході фільтра в будь якій ділянці фільтра, представленого на рис.2.6. Зазвичай вона отримується в результаті лінійного детектування з послідовною фільтрацією за допомогою фільтра низьких частот, як показано на рис.2.7. При відповідному виборі імпульсного відгуку фільтра низьких частот вихідна напруга  $|f(t) \cdot p(t)| \cdot q(t)$  наближено описує  $|F(\omega, t)|$ .

Спосіб вимірювання, який представлений на рисунку 2.7, є способом, який використовується у широко відомому звуковому спектрографі і в більшості аналізаторів спектру паралельного типу.



Рис.2.7. Блок-схема вимірювання миттєвого спектру  $|F(\omega, t)|$  з

використанням смугового фільтра, детектора та згладженого кола (фільтра низьких частот)

Даний спосіб зазвичай використовується для отримання миттєвого спектру у вокодерах та деяких пристроях для автоматичного формантного аналізу.

*Вокодер* – це пристрій, що дозволяє синтезувати мову на основі мінімальної інформації, деякого коду. В якості прикладу нижче на рис.2.8 представлені графіки миттєвого спектру мови, які відтворені гребінкою з 24 фільтрів.

Ширина смуги кожного фільтра приблизно 150Гц, які перекривають частотний діапазон від 150 до 4000Гц. За кожним фільтром знаходяться випрямляч і RC-ланка. Сигнали, які отримуються на виході гребінки фільтрів, по черзі комутуються з періодом 10мс. Дискретні відліки спектру з'єднуються прямими лініями. Відображені криві розгортки спектру були автоматично накреслені на обчислювальній машинці, в яку вводились дані з

виходів гребінки фільтрів. Лінії, які з'єднують пікові значення, являють собою формантні частоти мови, які були визначені в процесі машинної обробки миттєвого спектру.

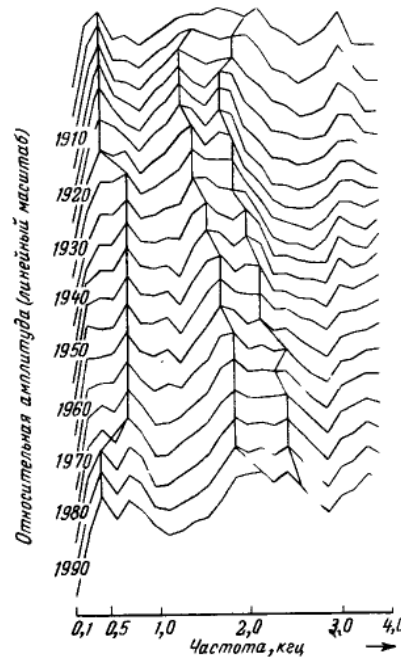


Рис.2.8. Миттєві амплітудні спектри мови

### • Вибір вагової функції

Для короткочасного аналізу мови доречно окремо розглядати такі властивості голосу, як вокалізовані та невокалізовані збудження, частоту основного тону та формантну структуру. Вибір вагової функції при аналізі визначається у результаті компромісу між допустимими можливостями за частотою і за часом. Коротка за тривалістю вагова функція відповідає фільтру з широкою смугою пропускання. За допомогою такого фільтру можна провести спектральний аналіз, у процесі якого виокремлюється часова структура одного періоду мовного сигналу. З іншого боку, вагова функція тривалістю в декілька періодів основного тону відповідає фільтру з більш

вузькою смугою пропускання. За допомогою цього фільтра можна провести аналіз по частоті кожної гармонічної спектральної складової.

Аби проілюструвати порядок практично використаних величин смуг пропускання і тривалості вагових функцій, введемо в розгляд ідеальні аналізуючі смугові фільтри (які, природно, не можуть бути реалізовані) з прямокутною амплітудно-частотною характеристикою і нульовою (чи строго лінійною) фазовою характеристикою. Припустимо, частотна характеристика фільтра має вигляд:

$$\left. \begin{aligned} P(\omega) &= 1; \\ (\omega_0 - \omega_1) \leq \omega \leq (\omega_0 + \omega_1) &= 1 \\ -(\omega_0 + \omega_1) \leq \omega \leq -(\omega_0 - \omega_1) &= 0 \end{aligned} \right\} \quad (2.9)$$

Відповідний для неї імпульсний відгук запишеться як

$$p(t) = \left( \frac{2\omega_1}{\pi} \right) \left( \frac{\sin \omega_1 t}{\omega_1 t} \right) \cos \omega_0 t = h(t) \cos \omega_0(t) \quad (2.10),$$

а вагова функція для цього ідеального фільтра є обвідною імпульсного відгуку виду  $\sin \frac{x}{x}$ . Якщо в якості ефективної тривалості  $D$  вагової функції

вибрати інтервал часу між першими нулями обвідної, то  $D = \frac{2\pi}{\omega_1} = \frac{4\pi}{\Delta\omega}$ , де

$\Delta\omega = 2\omega_1$  є смуга пропускання фільтра (іноді в якості ефективної тривалості вагової функції вибирають половину вказаної величини). У таблиці 1 в якості прикладу наведені величини  $D$ , які відповідають декільком значенням  $\Delta\omega$ .

Таблиця 1

Варіант	1	2	3
$\frac{\Delta\omega}{2\pi}$ [Гц]	50	100	250

D[мс]	40	20	8
-------	----	----	---

У першому варіанті фільтр, який аналізують, має таку смугу частот, яка дозволяє забезпечити необхідну точність аналізу при виділенні кожної гармонічної складової на вокалізованій ділянці мови. Цій смузі частот відповідає тривалість часового вікна, яка приблизно рівна чотирьом чи п'яти періодам основного тону чоловічого голосу.

З іншого боку, широкосмуговий фільтр у третьому варіанті характеризується ваговою функцією, яка порівнюється тривалістю лише з одним періодом основного тону чоловічого голосу. В цьому випадку точність аналізу в часі дозволяє враховувати амплітудні вимірювання, тривалість яких співвідносна з одним періодом основного тону. Фільтри, аналогічні тим, які наведені у варіантах 1 і 3, використовуються у звуковому спектрографі. Варіант №2 це свого роду компроміс між вимогами, які висуваються до аналізу мови з часовою і частотною точок зору. Фільтр у даному випадку має таку смугу пропускання, яка вважається найбільш підходящою при визначенні миттєвого спектру у приладах типу вокодер.

Доречно дати оцінку ефективній тривалості вагової функції для механічного миттєвого аналізу, який створюється базиллярною мембраною у вусі людини. Достатньо гарне наближення до закону зміщення базиллярної мембрани при впливі на неї одиничного імпульсу в точці максимальної чутливості до кругової частоти  $\beta$  створюється виразом

$$p(t) = (\beta t)^2 e^{-\beta t/2} \sin \beta t = h_{bm}(t) \sin \beta t \quad (2.11)$$

Вагова функція для базиллярної мембрани, побудована на основі вище згаданого виразу, показана на рис.2.9.

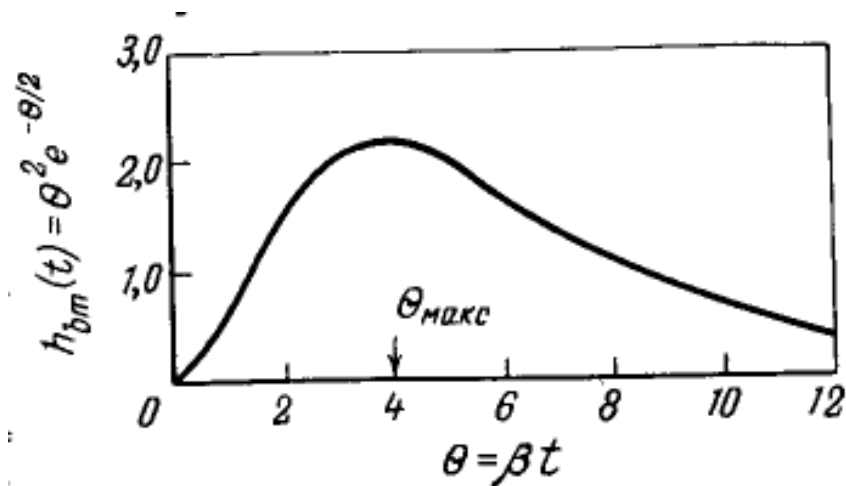


Рис.2.9.Протяжність вагової функції обернено пропорційна круговій частоті  $\beta$ .

Функція має максимум при  $t_{\text{макс}} = \frac{4}{\beta}$ . Якщо ефективну тривалість  $D$  першому наближенні вважати рівною  $2t_{\text{макс}}$ , то для різних точок мембрани отримаємо таблицю 2.

Таблиця 2

$\beta / 2\pi [\text{Гц}]$	100	1000	5000
$D = 2t_{\text{макс}}$	12,0	1,2	0,2

Отже, що для мовних сигналів механічний аналізатор вуха забезпечує кращу здатність по часу, ніж по частоті. Єдиною спектральною складовою, яка розрізняється цим механічним аналізатором є лише перша гармоніка основного тону.

• **Короткочасна функція кореляції і миттєвий спектр потужності**

Якщо  $x(t)$  стаціонарний випадковий процес, то його автокореляційна функція  $\varphi(\tau)$  і спектр потужності  $\Phi(\omega)$  зв'язані перетворенням Фур'є:

$$\varphi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} x(t)x(t+\tau)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\omega)e^{i\omega\tau} d\omega$$

і

$$\Phi(\omega) = \int_{-\infty}^{\infty} \varphi(\tau)e^{-i\omega\tau} d\tau \quad (2.12)$$

Для аперіодичного детермінованого сигналу  $y(t)$ , для якого існує перетворення Фур'є, зв'язок між автокореляційною функцією  $\psi(\tau)$  і енергетичним спектром  $\Psi(\omega)$  виражається аналогічними співвідношеннями:

$$\left. \begin{aligned} \psi(\tau) &= \int_{-\infty}^{\infty} y(t)y(t+\tau)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi(\omega)e^{i\omega\tau} d\omega \\ \Psi(\omega) &= \int_{-\infty}^{\infty} \psi(\tau)e^{-i\omega\tau} d\tau \end{aligned} \right\} \quad (2.13)$$

де  $\Psi(\omega) = Y(\omega)Y^*(\omega)$ , а  $Y(\omega) = \int_{-\infty}^{\infty} y(t)e^{-i\omega t} dt$ .

В обох випадках функція кореляції є дійсною і парною функцією параметра затримки  $\tau$ , а спектр – дійсною і парною функцією частоти  $\omega$ . Таким чином всі перетворення можна записати як косинус-перетворення[10]. Ці пари перетворення вказують на можливість визначення інформації про миттєвий спектр за допомогою кореляційних методів за умови, що останні розглядаються як випадок, який відповідає короткочасному аналізу.

При розгляді короткочасного спектрального аналізу для поставленої задачі підходили з точки зору необхідності аналізувати перетворені за допомогою Фур'є «шматочки» сигналу, отримані підходящим «зважуванням» його минулих значень. Кореляційні співвідношення для аперіодичних функцій можна також розповсюдити для опису мовного сигналу. Згідно зроблених попередніх висновків допускаємо, що в будь який момент часу  $t$  для мовного сигналу  $f(t)$  справедливі наступні перетворення:

$$\left. \begin{aligned} F(\omega, t) &= \int_{-\infty}^{\infty} f(\lambda)h(t-\lambda)e^{-i\omega\lambda} d\lambda \\ [f(\lambda)h(t-\lambda)] &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega, t)e^{i\omega\lambda} d\omega \end{aligned} \right\} \quad (2.14)$$

де  $h(t)$  вагова функція. Звідси формально маємо:

$$\left. \begin{aligned} \psi(\tau, t) &= \int_{-\infty}^{\infty} f(\lambda)h(t-\lambda)f(\lambda+\tau)h(t-\lambda-\tau)d\lambda \\ \psi(\tau, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi(\omega, t)e^{i\omega\tau} d\omega \\ \Psi(\omega, t) &= [F(\omega, t)F^*(\omega, t)] = \int_{-\infty}^{\infty} \psi(\tau, t)e^{-i\omega\tau} d\tau \end{aligned} \right\} \quad (2.15)$$

Практично при вимірюваннях у реальному масштабі часу  $t$  значення функції  $f(t+\tau)$  для  $\tau > 0$  невідоме. Але  $\psi(\tau, t)$  формально є парною функцією  $\tau$ . Тому вона може бути визначена тільки за допомогою від'ємних значень  $\tau$ , тоді:

$$\Psi(\omega, t) = \int_{-\infty}^{\infty} \psi(\tau, t)e^{-i\omega\tau} d\tau = 2 \int_{-\infty}^0 \psi(\tau, t)\cos\omega\tau d\tau \quad (2.16)$$

де  $\Psi(\omega, t)$  є також парною функцією  $\omega$ .

Таким чином може бути знайдена короткочасна автокореляційна міра, яка зв'язана з миттєвим спектром потужності  $|F(\omega, t)|^2$ . Вимірювання  $\psi(\tau, t)$  для від'ємних значень  $\tau$  може бути виконане за допомогою приладу, представленого на рис.2.10.

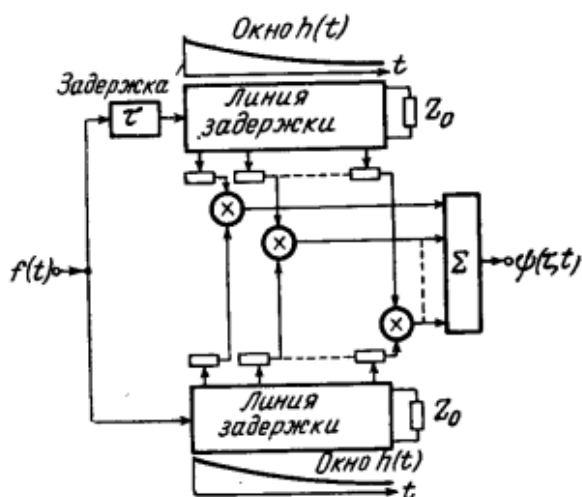


Рис.2.10. Спосіб вимірювання короткотривалої функції кореляції  $\psi(\tau, t)$

Сигнали, які знімаються з виводів лінії затримки, зважуються у відповідності зі значеннями функції  $h(t)$  у точках виводів. Відповідні відліки перемножуються, а операція інтегрування замінюється сумою кінцевого числа відліків. Таким чином,  $\psi(\tau, t)$  є поточною кореляцією, яка пов'язана з  $|F(\omega, t)|^2$  чи  $\Psi(\omega, t)$  перетворенням Фур'є [7].

Можна також визначити короткотривалу функцію кореляції як результат зваженого добутку вихідного та затриманого сигналів. Для цього випадку маємо

$$\varphi(\tau, t) = \int_{-\infty}^t f(\lambda) f(\lambda + \tau) k(t - \lambda) d\lambda \quad (2.17)$$

де  $k(t) = 0$  при  $t < 0$  є ваговою функцією. Вимірювання можна провести для  $\tau \leq 0$  за допомогою пристрою, схема якого представлена на рис.2.11.



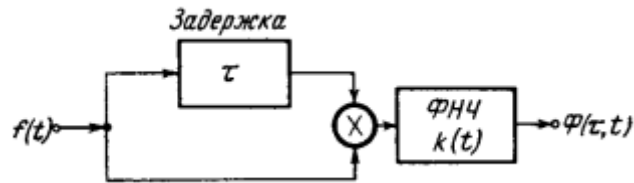


Рис.2.11. Схема для вимірювання поточних значень короткотривалої функції кореляції  $\varphi(\tau, t)$

У загальному випадку не існує простого співвідношення між  $\varphi(\tau, t)$  і миттєвим спектром потужності, отриманим на виході вимірювального пристрою. Однак, в особливому випадку  $k(t) = 2ae^{-2at} = [h(t)]^2$  функція  $\varphi(\tau, t)$  може бути пов'язана з  $\Psi(\omega, t) = |F(\omega, t)|^2$ :

$$\begin{aligned} \psi(\tau, t) &= \int_{-\infty}^t f(\lambda)h(t-\lambda)f(\lambda+\tau)h(t-\lambda-\tau)d\lambda = \\ &= e^{a\tau} \int_{-\infty}^t 2af(\lambda)f(\lambda+\tau)e^{-2a(t-\lambda)}d\lambda = e^{a\tau}\varphi(\tau, t); \tau \leq 0. \end{aligned} \quad (2.18)$$

Вище було показано, що  $\psi(\tau, t)$  є парною функцією  $\tau$  і якщо  $\varphi(\tau, t)$  також визначена як парна функція, то  $\psi(\tau, t) = e^{-a|\tau|}\varphi(\tau, t)$  для всіх  $\tau$ , чи :

$$\varphi(\tau, t) = e^{a|\tau|}\psi(\tau, t) = \frac{e^{a|\tau|}}{2\pi} \int_{-\infty}^{\infty} \Psi(\omega, t)e^{i\omega\tau} d\omega$$

та

$$\Psi(\omega, t) = \int_{-\infty}^{\infty} e^{-a|\tau|}\varphi(\tau, t)e^{-i\omega\tau} d\tau = \int_{-\infty}^{\infty} e^{-a|\tau|}\varphi(\tau, t)\cos\omega\tau d\tau \quad (2.19a)$$

Звідси також слідує, що

$$\begin{aligned}\Psi(\omega, t) &= \frac{1}{2\pi} \left[ \mathfrak{F}\{e^{-a|\tau|}\} \cdot \mathfrak{F}\{\varphi(\tau, t)\} \right] = \frac{1}{2\pi} \left[ \left( \frac{2a}{a^2 + \omega^2} \right) \cdot \Phi(\omega, t) \right] = \\ &= \frac{1}{2\pi} \left[ |H(\omega)|^2 \cdot \Phi(\omega, t) \right]\end{aligned}\quad (2.196)$$

де  $\mathfrak{F}$  означає перетворення Фур'є.

Таким чином, миттєвий спектр потужності  $\Psi(\omega, t)$  є результатом згортки спектру потужності  $\Phi(\omega, t)$  з квадратом коефіцієнта передачі фільтра низьких частот  $\left[ 2a / (a^2 + \omega^2) \right]$ . Тому  $\Psi(\omega, t)$  має більш низьку допустиму здатність за частотою, ніж перетворення Фур'є від  $\varphi(\tau, t)$ . Варто звернути увагу на те, що для  $\left[ (2a)^{1/2} e^{-at} \cos \omega t \right]$  по суті, є результатом вимірювання одиночних резонансних контурів з імпульсними відгуками виду  $\left[ (2a)^{1/2} e^{-at} \cos \omega t \right]$  і  $\left[ (2a)^{1/2} e^{-at} \sin \omega t \right]$ .

Вагові функції, які відрізняються від тільки що розглянутої експоненціальної, не приводять до простих співвідношень між  $\varphi(\tau, t)$  і спектром потужності. Однак для кореляційних функцій, отриманих на виході вимірювального пристрою і миттєвих спектрів потужності можна дати інші визначення, які дозволяють пов'язати ці функції перетвореннями особливого виду. Наприклад, миттєвий спектр можна визначити як:

$$\Omega(\omega, t) = \int_{-\infty}^{\infty} \varphi(\tau, t) m(|\tau|) \cos \omega \tau d\tau, \quad (2.20a)$$

в якому  $\varphi(\tau, t)$  як і у виразі (2.17), визначена як парна функція  $\tau$ , тому :

$$\varphi(\tau, t) = \int_{-\infty}^t f(\lambda) f(\lambda - |\tau|) n(t - \lambda) d\lambda \quad (2.20b)$$

де  $m(t)$  і  $n(t)$  суть фізично реалізована та рівна нулю для  $t < 0$  вагові функції. Таким чином,  $\Omega(\omega, t)$  і  $\varphi(\tau, t)$  виявляються зв'язаними за допомогою визначень (2.20). Вимірювання  $\Omega(\omega, t)$  також можна виконати безпосередньо, підставивши значення  $\varphi(\tau, t)$  з формул (2.20б) у (2.20а), отримаємо:

$$\begin{aligned} \Omega(\omega, t) &= 2 \int_{-\infty}^t f(\lambda) n(t - \lambda) d\lambda \int_0^{\infty} f(\lambda - \tau) m(\tau) \cos \omega \tau d\tau = \\ &= 2 \{ n(t) \cdot f(t) [ f(t) \cdot m(t) \cos \omega t ] \} \end{aligned} \quad (2.21)$$

Отже, як видно з виразу (2.21), для того щоб отримати миттєвий спектр  $\Omega(\omega, t)$ , необхідно профільтрувати сигнал  $f(t)$  смуговим фільтром, який має імпульсну реакцію виду  $[m(t) \cos \omega t]$ , помножити вихідну напругу цього фільтру на вихідний сигнал і результат профільтрувати фільтром нижніх частот, який має імпульсну реакцію  $n(t)$ . Пристрій для вимірювання сигналу  $\Omega(\omega, t)$  представлений на рис.2.12.

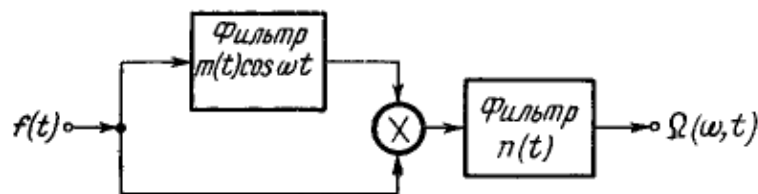


Рис.2.12. Пристрій для вимірювання короткотривалого спектру  $\Omega(\omega, t)$

При  $m(t) = n(t) = e^{-at} \Omega(\omega, t)$  перетворюється в  $\Psi(\omega, t)$ . Із визначення для  $\Omega(\omega, t)$  впливає зворотне співвідношення:

$$\varphi(\tau, t) = \frac{1}{2\pi m|\tau|} \int_{-\infty}^{\infty} \Omega(\omega, t) \cos \omega \tau d\omega \quad (2.22)$$

Співвідношення (2.20) також означає, що

$$\Omega(\omega, t) = M(\omega) \cdot \Phi(\omega, t) \quad (2.23)$$

де

$$M(\omega) = \int_{-\infty}^{\infty} m(|\tau|) e^{-i\omega\tau} d\tau; \quad \Phi(\omega, t) = \int_{-\infty}^{\infty} \varphi(\tau, t) e^{-i\omega\tau} d\tau$$

Цей результат може бути співставлений з виразом ( 2.19), де

$$|H(\omega)|^2 = \int_{-\infty}^{\infty} e^{-a|\tau|} e^{-i\omega\tau} d\tau$$

$$H(\omega) = \int_0^{\infty} (2a)^{1/2} e^{-a\tau} e^{-i\omega\tau} d\tau = \int_0^{\infty} h(\tau) e^{-i\omega\tau} d\tau$$

Оскільки  $\Omega(\omega, t)$  отримана з  $\Phi(\omega, t)$  за допомогою згортки з коефіцієнтом передачі  $M(\omega)$  (фільтру низьких частот), вона має більш низьку спектральну чіткість порівняно з  $\Phi(\omega, t)$ .

### • Середній спектр потужності

У розглянутих вище схемах для вимірювання спектрів, зважування минулих значень сигналу відбуваються за допомогою функцій відносно короткої тривалості. В результаті отримуються спектри, які відображають достатньо швидкі часові зміни. Якщо важливі не короткочасні зміни, а середній спектральний розподіл, цікавість викликають також значення спектру, яке усереднене за тривалий проміжок часу, виду  $\overline{|F(\omega, t)|^2}$ . Таке усереднення може бути записане у вигляді:

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T F(\omega, t) \cdot F^*(\omega, t) dt &= \overline{|F(\omega, t)|^2} = \overline{|\Psi(\omega, t)|} = \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-\infty}^t dt \int_{-\infty}^t f(\lambda) h(t - \lambda) e^{-i\omega\lambda} d\lambda \int_{-\infty}^t f(\eta) h(t - \eta) e^{i\omega\eta} d\eta \end{aligned} \quad (2.24)$$

Замінюючи змінні та переставляючи члени, маємо:

$$\overline{|F(\omega, t)|^2} = \int_0^{\infty} d\lambda h(\lambda) e^{i\omega\lambda} \int_0^{\infty} d\eta h(\eta) e^{-i\omega\eta} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t-\lambda) f(t-\eta) dt \quad (2.25)$$

У відповідності з (2.12) останній інтеграл є просто функцією  $\varphi(\lambda - \eta)$ , яка є перетворенням Фур'є від  $\Phi(\omega)$ . Таким чином

$$\varphi(\lambda - \eta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\delta) e^{i\delta(\lambda - \eta)} d\delta = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\delta) e^{-i\delta(\lambda - \eta)} d\delta.$$

Оскільки  $\Phi(\omega)$  – функція дійсна і парна, то

$$\begin{aligned} \overline{|F(\omega, t)|^2} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\delta) d\delta \int_0^{\infty} h(\lambda) e^{i\lambda(\omega - \delta)} d\lambda \int_0^{\infty} h(\eta) e^{-i\eta(\omega - \delta)} d\eta = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\delta) H(\omega - \delta) H^*(\omega - \delta) d\delta = \overline{|F(\omega, t)|^2} = \frac{1}{2\pi} \left[ \Phi(\omega) \cdot |H(\omega)|^2 \right] \end{aligned} \quad (2.26)$$

З цього випливає, що усереднене, за тривалий проміжок часу, значення спектру потужності  $|F(\omega, t)|^2$  є згорткою спектра потужності  $\Phi(\omega)$  з квадратом перетворення Фур'є від вагової функції  $h(t)$ . Чим вузьчий спектр  $|H(\omega)|^2$ , тим точніше наближається  $\overline{|F(\omega, t)|^2}$  до спектру потужності  $\Phi(\omega)$ . У проміжку  $H(\omega)$  перетворюється в єдиний імпульс при  $\omega = 0$ , вагова функція – в одиничну функцію, а  $\overline{|F(\omega, t)|^2}$  набуває ті ж спектральні характеристики, якими володіє  $\Phi(\omega)$ . Для будь якого  $\omega$   $\overline{|F(\omega, t)|^2}$  є інтегралом енергетичного спектру, «побачений» крізь розташовану в  $\omega$  апертуру  $|H(\omega)|^2$ . Таким чином,  $\overline{|F(\omega, t)|^2}$  є середньою потужністю сигналу в частотній смузі фільтру, наведеного на рис.2.7.

Вище було показано (вираз 2.196), що при  $h(t) = \left[ (2a)^{1/2} e^{-at} \right] \Psi(\omega, t) = \frac{1}{2\pi} \left[ |H(\omega)|^2 \cdot \Phi(\omega, t) \right]$ . У цьому випадку

усереднений за тривалістю проміжок часу буде мати вигляд:

$$\overline{\Psi(\omega, t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \int_{-\infty}^{\infty} e^{-a|\tau|} \varphi(\tau, t) \cos \omega \tau d\tau dt = \int_{-\infty}^{\infty} e^{-a|\tau|} \overline{\varphi(\tau, t)} \cos \omega \tau d\tau \quad (2.27)$$

Підставляємо замість  $\varphi(\tau, t)$  його значення з (2.17) і замінюючи змінні, отримаємо:

$$\overline{\Psi(\omega, t)} = \int_0^{\infty} e^{-a|\tau|} \varphi(\tau, t) \cos \omega \tau d\tau \int_0^{\infty} k(\beta) d(\beta) \quad (2.28)$$

Оскільки  $\int_0^{\infty} k(t) dt = \int_0^{\infty} h^2(t) dt = 1$ , тоді  $\overline{\Psi(\omega, t)} = \frac{1}{2\pi} \left[ |H(\omega)|^2 \cdot \Phi(\omega) \right]$ , що

відповідає результату, отриманому в (2.26) [9].

## РОЗДІЛ 3. ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ УСЕРЕДНЕНИХ СПЕКТРІВ ЗАХІДНОГЕРМАНСЬКОЇ, РОМАНСЬКОЇ, ЗАХІДНОСЛОВ'ЯНСЬКОЇ, СХІДНОСЛОВ'ЯНСЬКОЇ ГРУП МОВ

### 3.1. Умови проведення досліджень

У роботі було розглянуто основні групи індоєвропейської родини мов, яка домінує на території Європи згідно генеалогічної класифікації. Дана класифікація базується на існуванні регулярних системних фонетичних, лексичних і граматичних відповідностей, які закономірно простежуються у споріднених мовах.

Для дослідження були використанні записи аудіокниг з електронних ресурсів: <https://www.twirpx.com/>, <http://www.literatureaudio.com>, <https://www.lingq.com>, <https://tracker.cztorrent.net>. Використані уривки творів, належать до двох стилів мови:

- **художній стиль**

1. Західногерманська група.

Англійська мова – Стівен Кінг «11/22/63».

Німецька мова – Роберт Стівенсон «Острів скарбів».

2. Романська група.

Італійська мова – Жуль Верн «Двадцять тисяч льє під водою».

Французька мова – Артур Конан Дойл «Шість Наполеонів».

3. Західнослов'янська група.

Польська мова – Ернст Теодор Амадей Гофман «Лускунчик і Мишачий король».

Словацька мова – Дан Домінік «Вузол».

4. Східнослов'янська група.

Українська мова – Василь Манько «Жовтий князь».

Російська мова – Микола Гоголь «Вій».

- **Публіцистичний стиль**

- 1. Західногерманська група.

Англійська мова – Пеллі Скотт «Істина, гідна розповіді: репортер шукає сенс в історіях нашого часу».

Німецька мова – Журнал «Наука».

- 2. Романська група.

Італійська мова – МакКенна Пол «Зміни своє життя за 7 днів».

Французька мова – Чарльз Дарвін «Походження видів».

- 3. Західнослов'янська група.

Польська мова – Рафал Земкевіч «Polaństwo».

Словацька мова – Іво Томан «Позитивне мислення не для всіх».

- 4. Східнослов'янська група.

Українська мова – Еплбом Енн «Історія ГУЛАГу».

Російська мова – Сухотін Анатолій «Парадокси науки».

Тривалість досліджуваних записів становила 2хв. 30с.±5с. Похибка читання зумовлена тим, що файли формувалися таким чином, щоб слова були дочитані до кінця аби не порушувати інтонаційну картину.

Зазначимо, що декілька стилів мови було обрано, щоб дізнатись чи впливають особливості лексики, інтонація та емоційний стан диктора на спектральні характеристики мови.

Досліджуваний діапазон частот був обраний, спираючись на діапазон слуху людини. Нижня межа становила 100 Гц, що відповідає першій формантній частоті чоловічого голосу (перша формантна частота жіночого голосу є вищою); верхня межа - 20кГц, оскільки для української, італійської, польської мові характерні звуки, що потрапляють в дану частотну область.



### 3.2. Алгоритм обробки сигналів

Для знаходження усередненого спектру густини потужності мовленнєвих сигналів їхня обробка виконувалася у програмному середовищі Matlab.

Наведемо алгоритм обробки сигналу для отримання усередненого спектру:

- 1) нормування сигналу, записаного кожним диктором, шляхом ділення на стандартне відхилення цього сигналу, за допомогою команди **std**, завдяки чому отримані сигнали з однаковою (одиничною) дисперсією;
- 2) для кожного нормованого сигналу обчислюється спектр потужності за допомогою команди **pwelch**;
- 3) усереднення отриманих оцінок спектрів потужності, за допомогою команди **mean**;
- 4) нормування отриманої усередненої оцінки до максимуму командою **max(abs(.))**;
- 5) для отримання рівня спектральної густини сигналу, максимум якого буде прив'язаний до 0дБ, обчислюємо функцію **10\*log10** від усередненої та нормованої за максимумом оцінки.

Не менш важливим є те, що більш «плавну» криву графіка можливо одержати при усередненні результатів запису більшої кількості дикторів та при збільшенні тривалості звукового сигналу.

### 3.3. Аналіз усереднених спектрів мови

У результаті виконання алгоритму для знаходження усередненого спектру мови було отримано графіки частотної залежності рівня густини потужності записів для кожної мови різних стилів мови; для груп мов за

стилями мови та загалом, а також графіки порівняння кожної мови, групи мов з українською мовою (Додатки А-Б ).

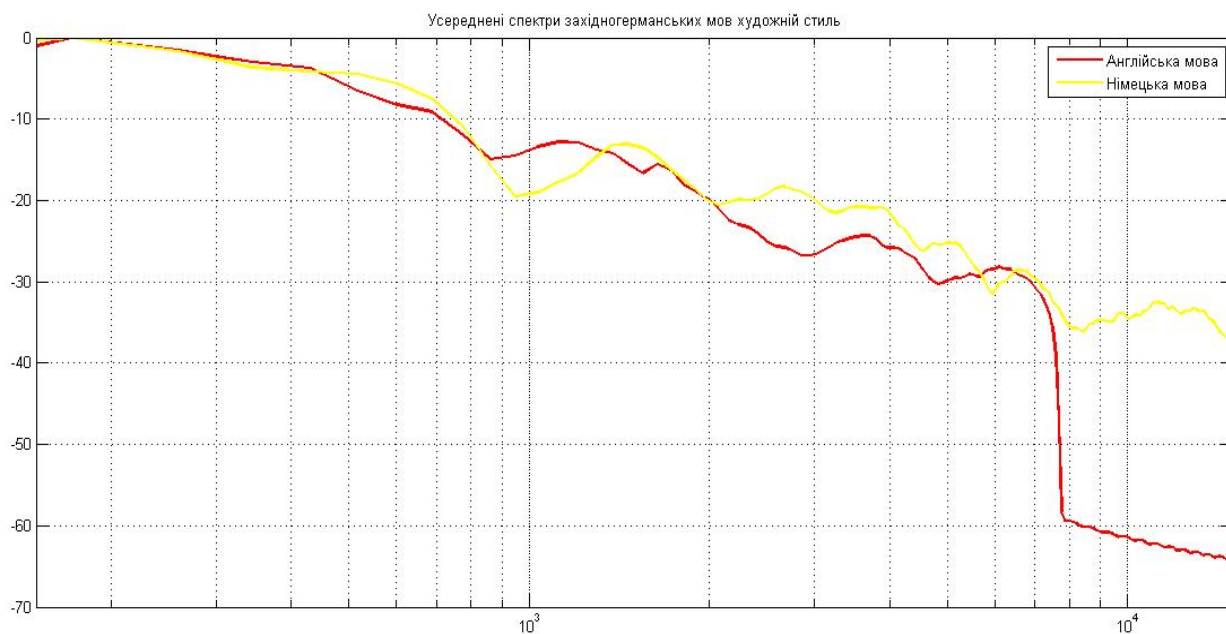


Рис. 3.1. Усереднений спектр західногерманської групи мов художній стиль

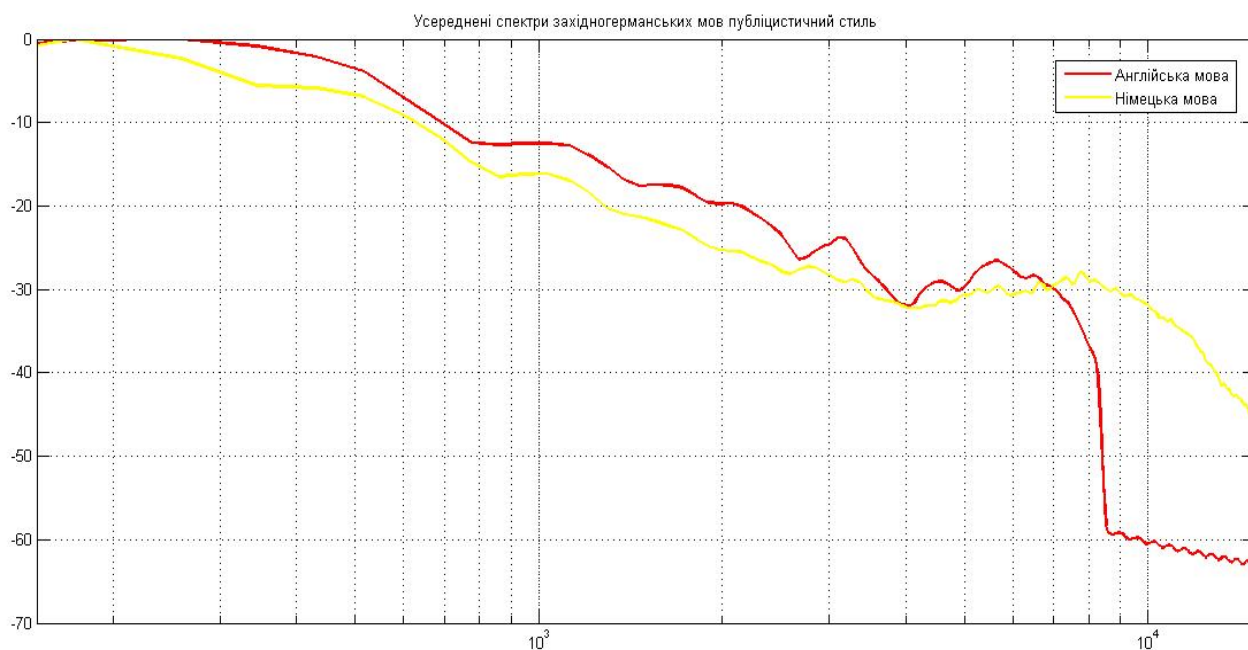


Рис. 3.2. Усереднений спектр західногерманської групи мов публіцистичний стиль

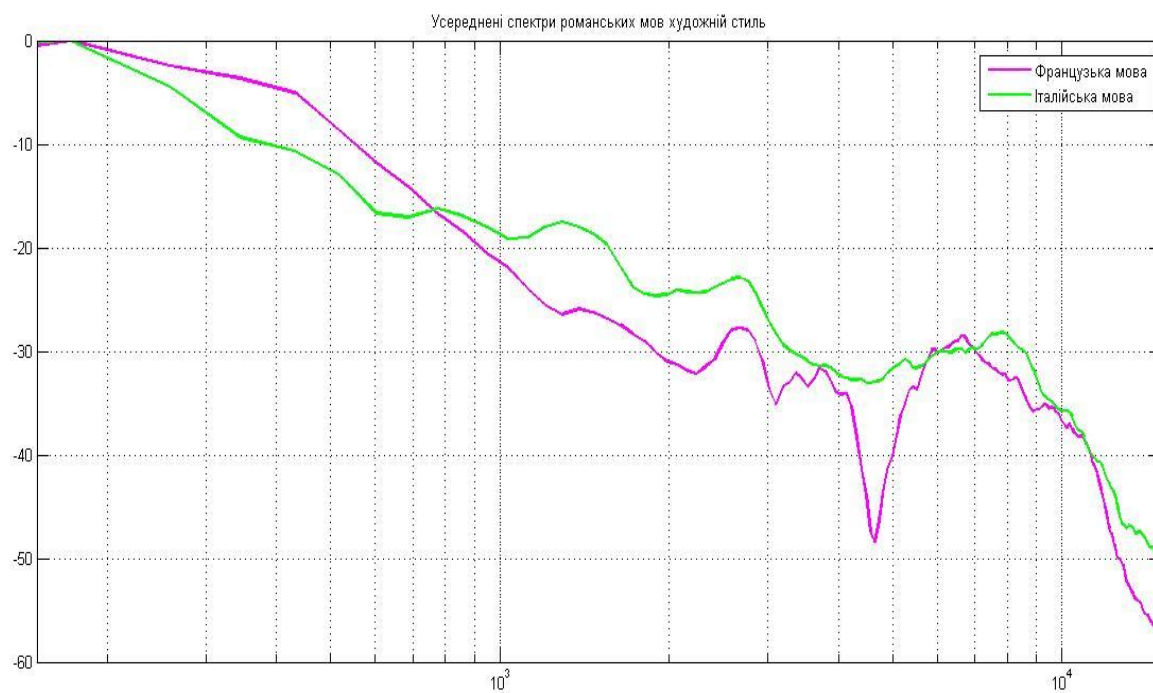


Рис. 3.3. Усереднений спектр романської групи мов художній стиль

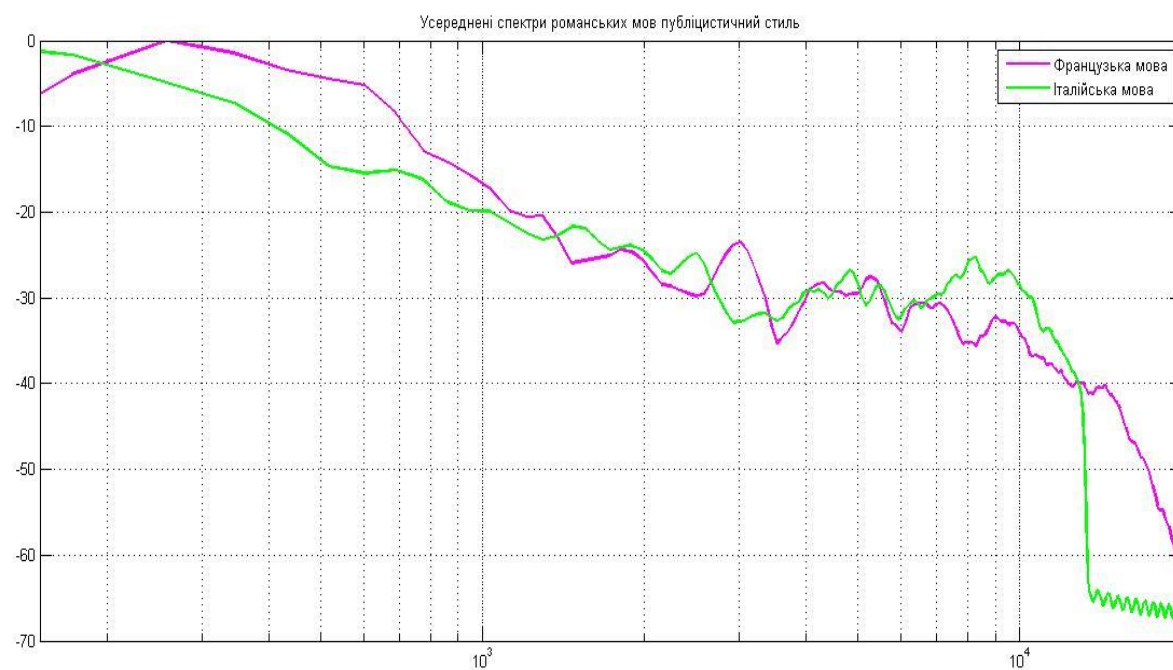


Рис. 3.4. Усереднений спектр романської групи мов публіцистичний  
СТИЛЬ

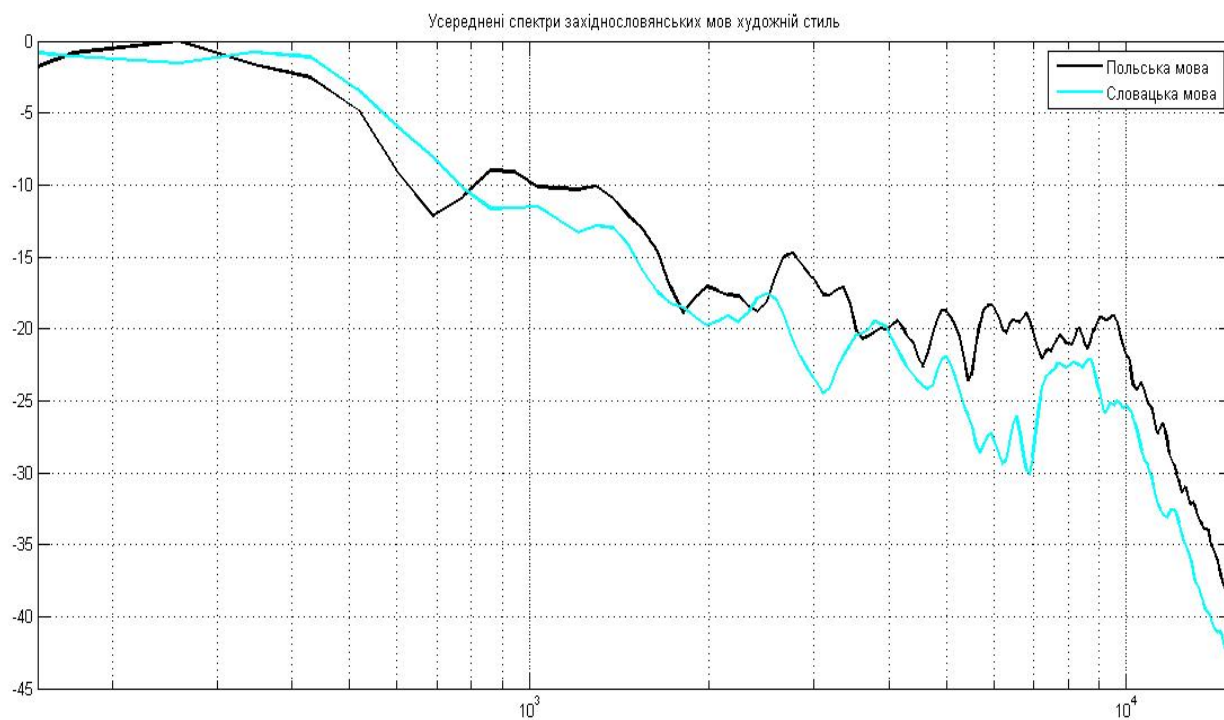


Рис. 3.5. Усереднений спектр західнослов'янської групи мов художній стиль

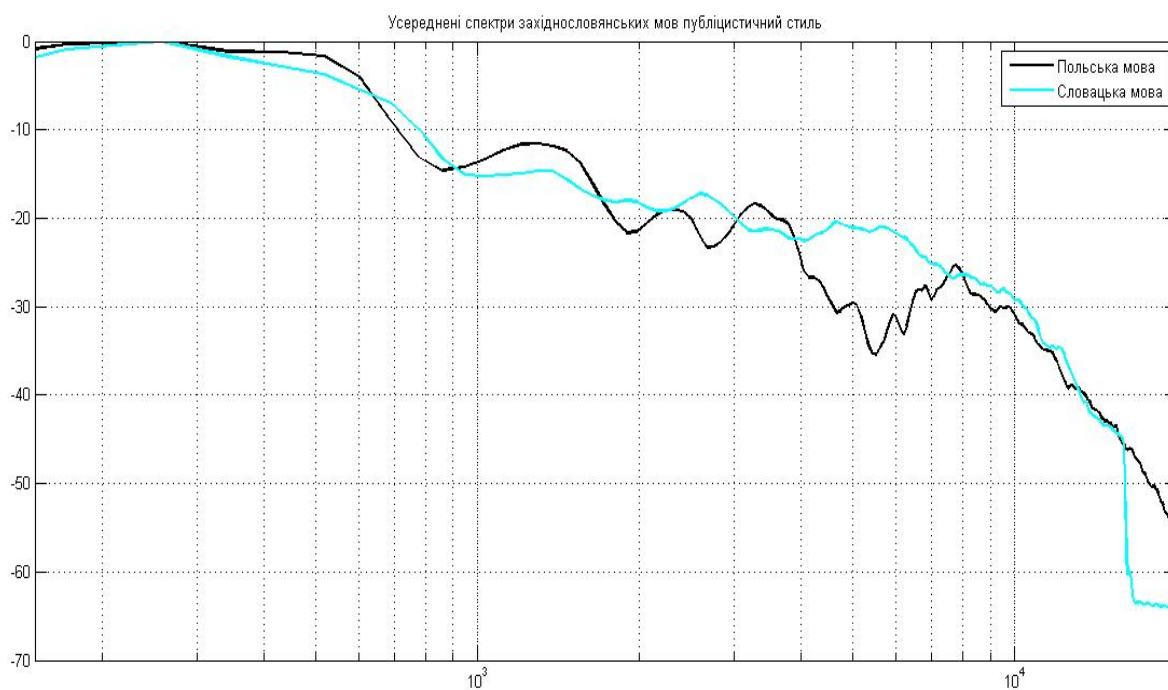


Рис. 3.6. Усереднений спектр західнослов'янської групи мов публіцистичний стиль

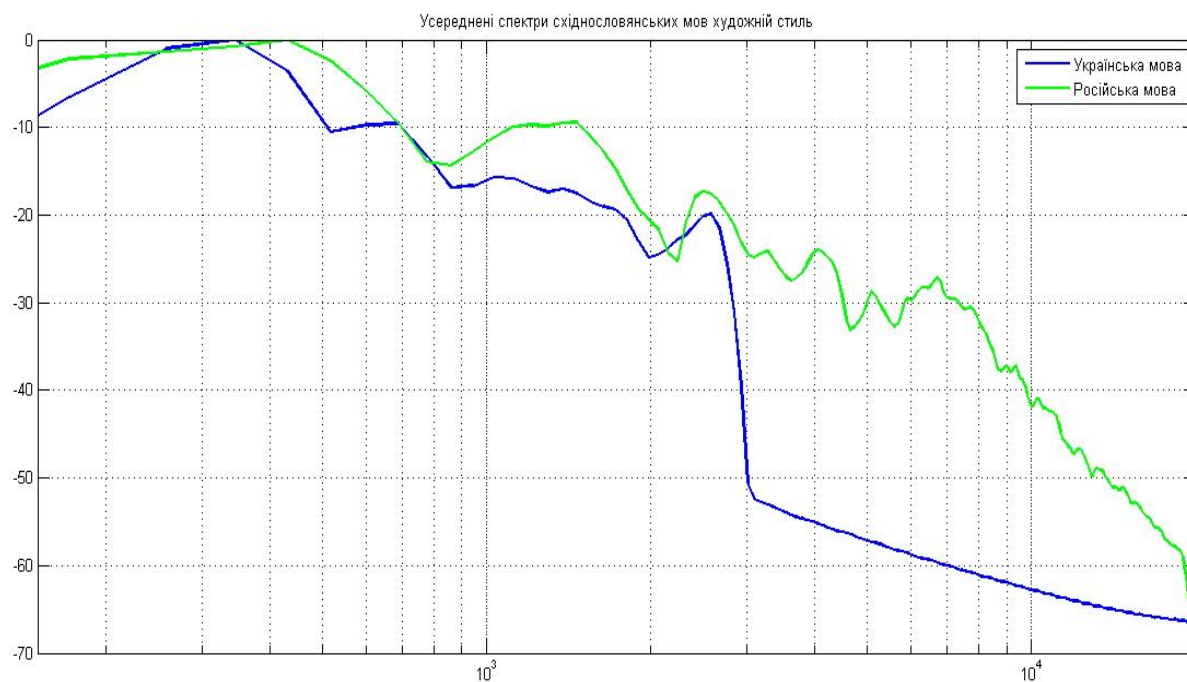


Рис. 3.7. Усереднений спектр східнослов'янської групи мов художній стиль

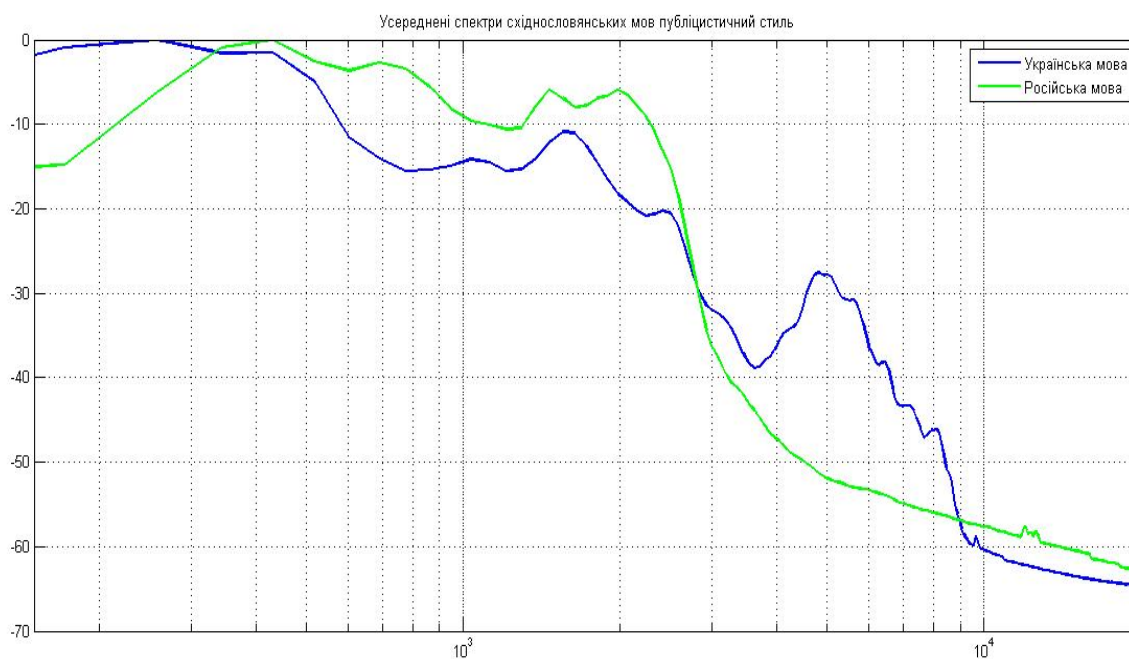


Рис. 3.8. Усереднений спектр східнослов'янської групи мов публіцистичний стиль

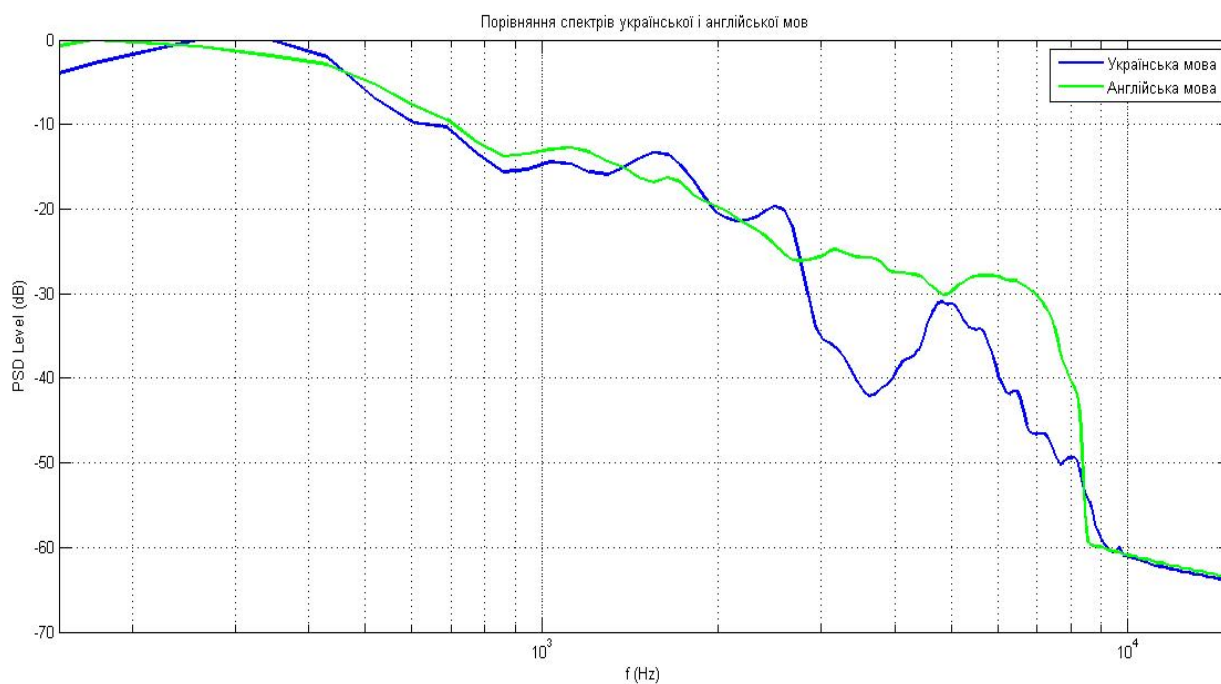


Рис. 3.9. Усереднений спектр англійської мови в порівнянні з українською мовою

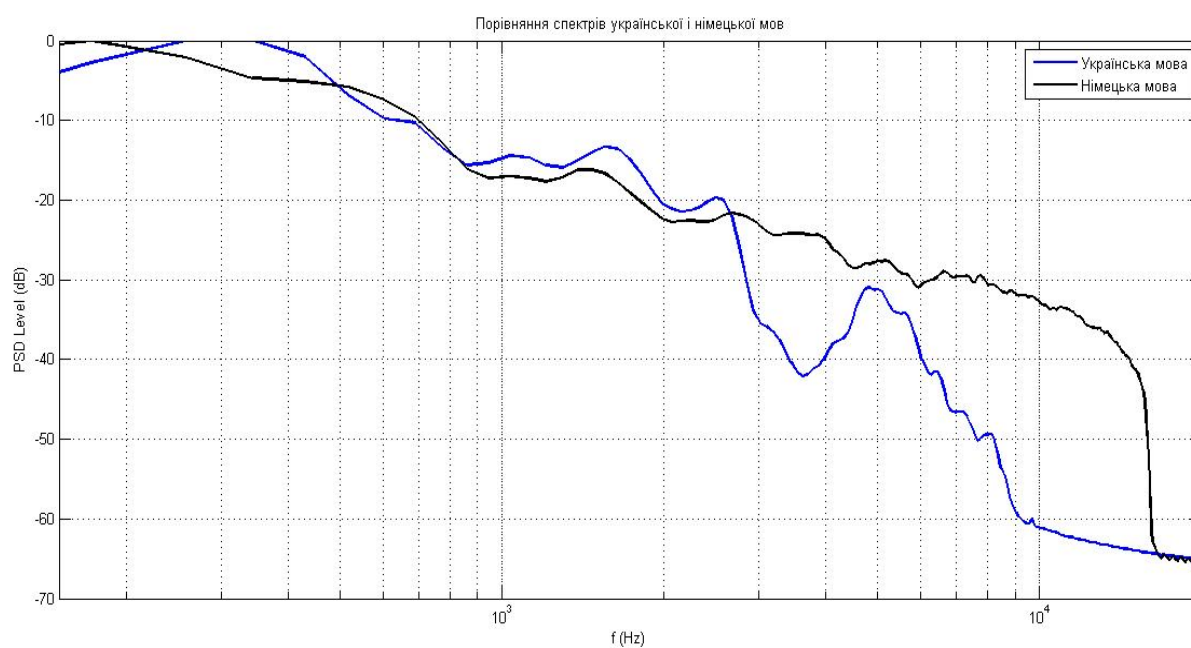


Рис. 3.10. Усереднений спектр німецької мови в порівнянні з українською мовою



Рис. 3.11. Усереднений спектр італійської мови в порівнянні з українською мовою

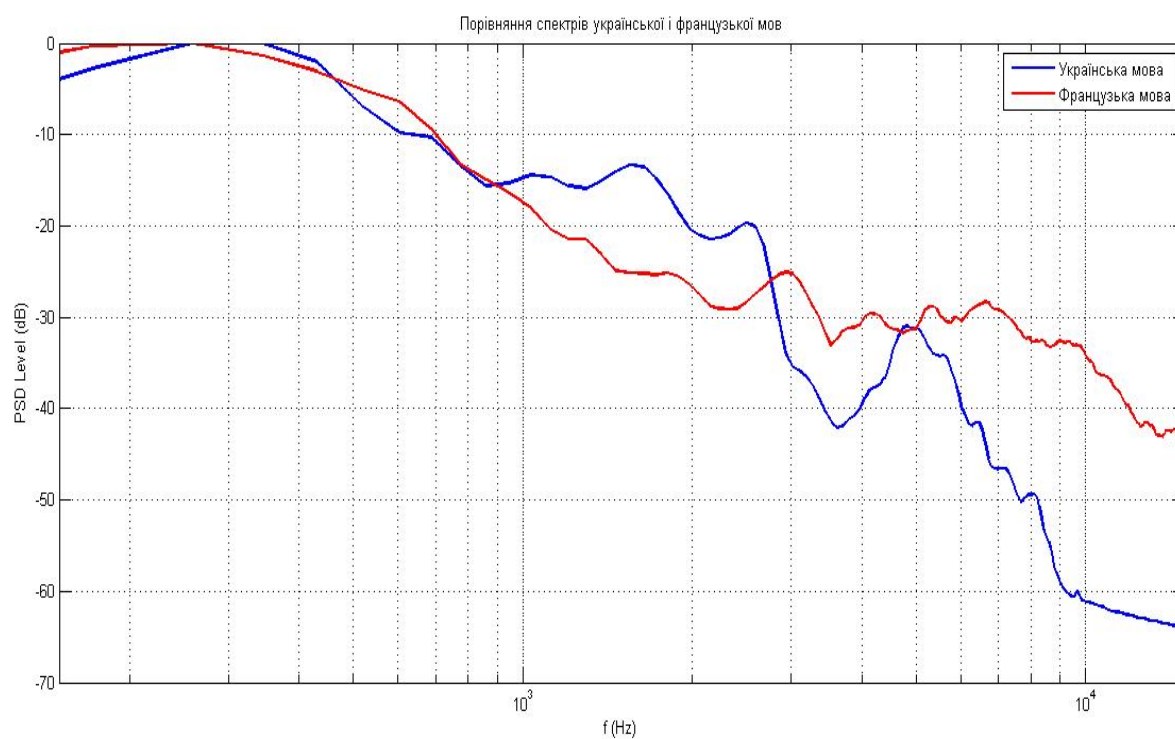


Рис. 3.12. Усереднений спектр французької мови в порівнянні з українською мовою

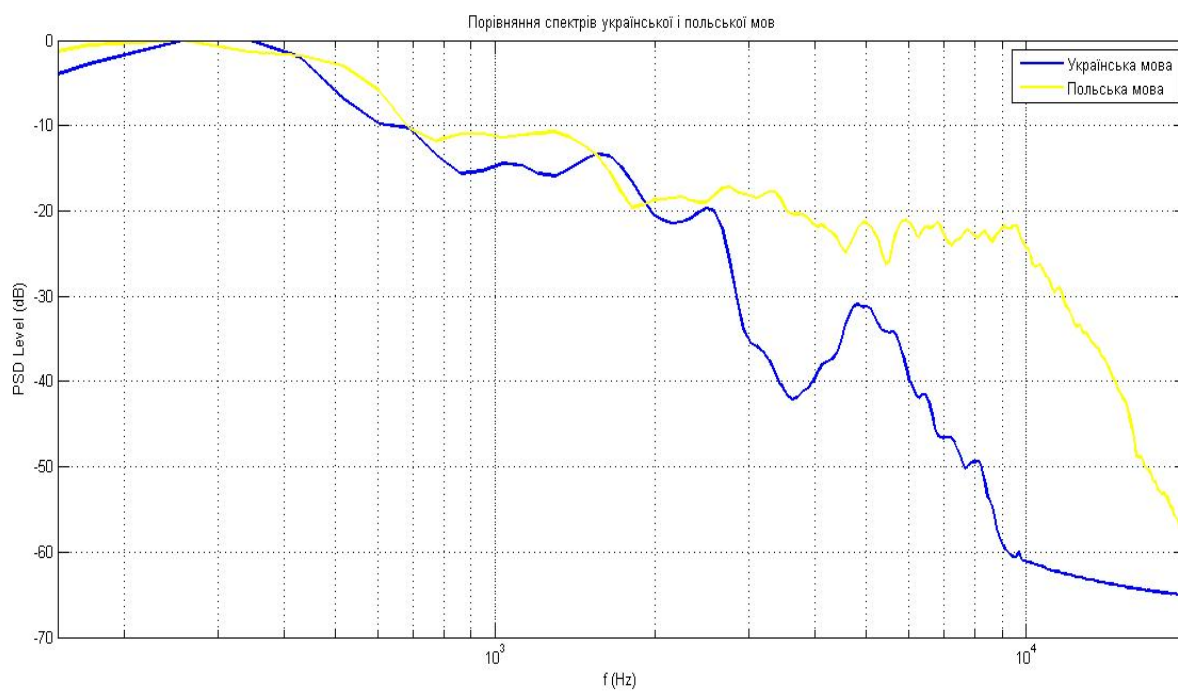


Рис. 3.13. Усереднений спектр польської мови в порівнянні з українською мовою

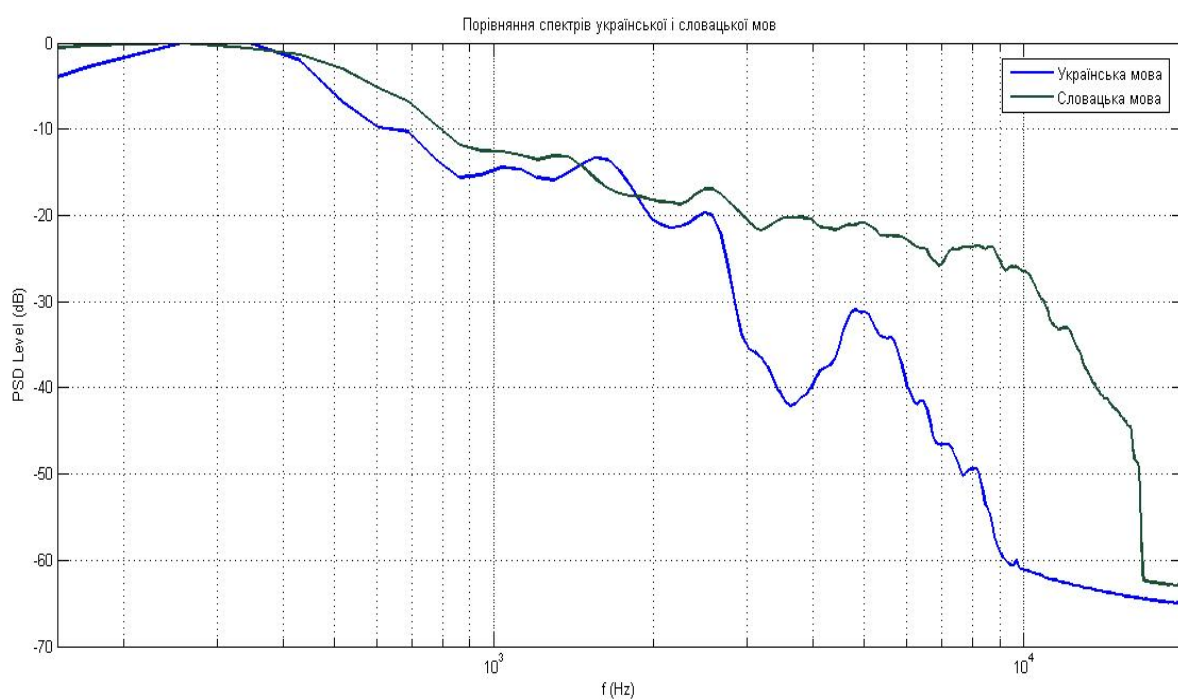


Рис. 3.14. Усереднений спектр словацької мови в порівнянні з українською мовою



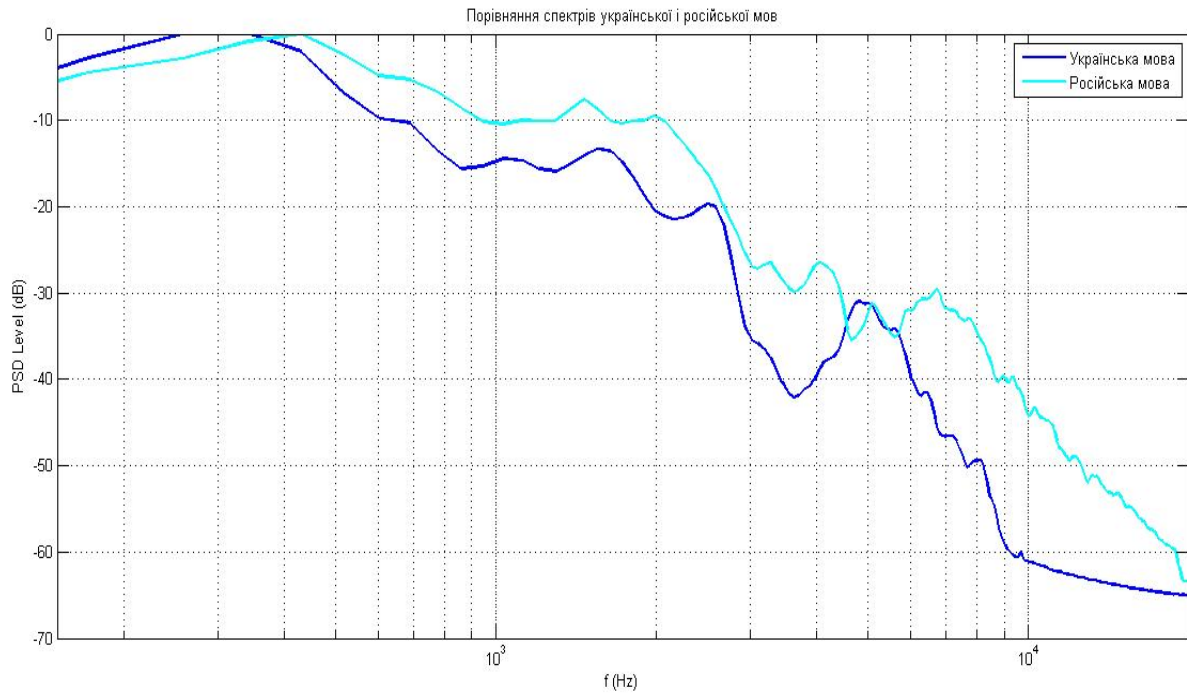


Рис. 3.15. Усереднений спектр російської мови в порівнянні з українською мовою

Мови західногерманської групи (німецька, англійська) мають дуже схожий спектр до частоти 7 кГц (рис. 3.1, 3.2). З підвищенням частоти спектральний рівень потужності німецької мови залишається доволі високим, в той час, як для англійської мови – різко знижується. Така закономірність просліджується і в художньому, і в публіцистичному стилях.

Про високу спорідненість мов романської групи (французької, італійської) свідчать спектри цих мов (рис.3.3, 3.4). Особливо це очевидно для публіцистичного стилю (рис.3.3). У художньому стилі (рис.3.4) спостерігаються суттєві відмінності в районі частот 4-6 кГц, де у французькій мові має місце різке зниження рівня спектральних складових. Загалом, французька мова має більш потужний низькочастотний спектр.

Мови західнослов'янської групи (рис.3.5, 3.6) дуже наближені одна до одної за винятком високочастотного діапазону в межах 3-8 кГц.

Що стосується порівняння спектрів української і російської мов, які належать до східнослов'янської групи, то для окремих стилів мов ця різниця

дуже суттєва (рис.3.7, 3.8). У зв'язку з цим, краще проаналізувати спектральні характеристики, усереднені за двома стилями: художнім і публіцистичним (рис.3.15).

В українській мові більше підкреслені низькочастотні складові ( до частоти 400 Гц). Також має місце яскраво виражена форманта в районі 4 кГц, що свідчить про відмінність фонетики української мови від російської.

Порівняння спектральних характеристик української мови з аналогічними характеристиками західноєвропейських мов ( рис.3.9-3.14) свідчать про значну різницю між ними в області високих частот. Відмінності мають місце в районі 2-3 кГц і, особливо, на частотах вищих за 5кГц, де спектральні складові української мови різко спадають. Отже, попри те, що досі чимало спеціалістів стверджують, що можна проводити дослідження на діапазоні, обмеженому частотою 5-6 кГц, оскільки цього діапазону цілком достатньо для передачі змістовної інформації, такий діапазон не дозволяє передати особливості національної мови.

Таким чином, на основі проведеного аналізу спектральних характеристик більшості європейських мов можна зробити висновок, що українська мова за своїм спектральним складом найбільше наближається до російської, а також частково до англійської мов.

## РОЗДІЛ 4

### РОЗРОБКА СТАРТАП-ПРОЕКТУ

Стартап-проекти призначені вирішувати проблеми та задачі, які з часом вдається втілити завдяки використанню технічного прогресу.

**Метою розділу** є формування інноваційного мислення, підприємницького духу та формування здатностей щодо оцінювання ринкових перспектив і можливостей комерціалізації основних науково-технічних розробок, сформованих у попередній частині магістерської дисертації у вигляді розроблення концепції стартап-проекту в умовах висококонкурентної ринкової економіки глобалізаційних процесів.

У даному розділі магістерської дисертації буде виконано перший етап розробки стартап-проекту, а саме маркетинговий аналіз. У межах цього етапу необхідно:

- розробити опис самої ідеї проекту та визначити загальні напрями використання потенційного товару чи послуги, а також їхні відмінності від конкурентів;
- проаналізувати ринкові можливості щодо його реалізації;
- на базі аналізу ринкового середовища розробити стратегію ринкового впровадження потенційного товару в межах проекту.

#### 4.1. Опис ідеї проекту

Таблиця 4.1. Опис ідеї стартап-проекту

Зміст ідеї	Напрямки застосування	Переваги для користувача
Мобільний додаток для	Архітектурна акустика	Простота у

проведення акустичної експертизи приміщення		використанні, мобільність, менша вартість проведення експертизи, швидкість вимірювання і оцінки приміщення
---	--	--

У табл. 4.1 наведені основні напрямки використання запропонованої ідеї. Споживачами можуть бути як компанії, які займаються проектуванням залів різного призначення, наприклад, концертних, театральних, конференц-залів тощо; так і акустиками при виконанні перевірки приміщення і, навіть, невідготовленими спеціалістами для швидкої оцінки розбірливості в приміщенні.

Даний стартап проект не має аналогій на ринку в тому прояві, що пропонується, тому проведення аналізу з конкурентами неможливе.

## 4.2. Технологічний аудит проекту

У межах даного підрозділу проводиться аудит технології, за допомогою якої можна реалізувати ідею проекту (технології створення товару).

Визначення технологічної здійсненності ідеї проекту передбачає аналіз таких складових (табл.4.2):

- за якою технологією буде виготовлено товар згідно ідеї проекту?
- чи існують такі технології, чи їх потрібно розробити/доробити?
- чи доступні такі технології авторам проекту?

Таблиця 4.2. Технологічна здійсненність ідеї проекту

№ п/п	Ідея проекту	Технології реалізації ідеї	Наявність технологій	Доступність технологій
1.	ПЗ, яке працює в онлайн-режимі через сайт компанії	Javascript, PHP, Java	Наявна	Доступні, але досить дорого коштують
2.	ПЗ, яке працює в онлайн-режимі через мобільний додаток на смартфоні	C#, Java	Наявна	Платна, доступна
Обрана технологія реалізації ідеї проекту: технологія 2, хоча вона може бути дорожчою у реалізації, проте зможе бути впровадженна в застосування у більшому обсязі та буде зручнішою у використанні, оскільки не потребуватиме наявності прив'язки до конкретної компанії.				

### 4.3. Аналіз ринкових можливостей запуску стартап-проекту

Визначення ринкових можливостей, які можна використати під час ринкового впровадження проекту, та ринкових загроз, які можуть перешкодити реалізації проекту, дозволяє спланувати напрями розвитку проекту із урахуванням стану ринкового середовища, потреб потенційних клієнтів та пропозицій проектів-конкурентів. Спочатку проводиться аналіз попиту: наявність попиту, обсяг, динаміка розвитку ринку.

Таблиця 4.3. Попередня характеристика потенційного ринку стартап-проекту

№ п/п	Показники стану ринку (найменування)	Характеристика
1.	Кількість головних гравців, од	1

2.	Загальний обсяг продаж, у.о.	0
3.	Динаміка ринку (якісна оцінка)	зростає
4.	Наявність обмежень для входу (вказати характер обмежень)	немає
5.	Специфічні вимоги до стандартизації та сертифікації	є
6.	Середня норма рентабельності в галузі (або по ринку), %	35%

З табл. 4.3 можна зазначити, що вихід на ринок є рентабельним, так як на є низька конкуренція у даній галузі та високий відсоток рентабельності, що дає змогу швидко покрити витрачені кошти на розробку пристрою.

Таблиця 4.4. Характеристика потенційних клієнтів стартап-проекту

Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних цільових груп клієнтів	Вимоги споживачів до товару
Зменшення часу та грошових витрат проведення акустичної експертизи приміщень	Компанії, що займаються архітектурною акустикою, фахівці акустики, звичайні користувачі	Співвідношення ціна/якість проведення діагностики	Простота у використанні, швидкість обрахунку показників оцінювання розбірливості мови в приміщенні, отримання результатів, що відповідають реальним

			експериментам, які застосовуються при звичних методах оцінки приміщення
--	--	--	---

Таблиця 4.5. Фактори загрози

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
1.	Незацікавленість аудиторії	Неготовність використання продукту	Збільшення клієнтів
2.	Якісний	Не належна якість моделі	Зміна технологічних процесів виробництва
3.	Несанкціоноване використання ПЗ	Взлом ПЗ, використання ПЗ без ліцензії	Перевірка ліцензії користувача при кожній можливості та блокування ПЗ у разі відсутності ліцензії

Таблиця 4.6. Фактори можливостей

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
1.	Збільшення попиту	Різка збільшення зацікавленості до продукту	Підвищення виробництва
2.	Зріст кількості споживачів	Розширення ринку споживачів	Розширення компанії, постійний випуск оновлень ПЗ
3.	Поява постійних	Можливість додавання	Введення програми

	клієнтів	нових систем до існуючої для пришвидшення розвитку компанії	лояльності, акцій
--	----------	---	-------------------

Таблиця 4.7. Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
Тип конкуренції – чиста	Мала кількість аналогів даної моделі	Розвивати систему збільшуючи її продуктивність та надійність
За рівнем конкурентної боротьби – міжнародний	Наявність замовників та виробників із інших країн	Вихід на міжнародний ринок
За галузевою ознакою – одногалузєва	Використання тільки в акустиці	Проведення потужної рекламної кампанії
Конкуренція за видами товарів – товарно-видова	Запропонований товар є одного виду	Орієнтація стратегії компанії на клієнта та адаптація до змін ринкових умов
За характеристиками конкурентних переваг – нецінова	Основною перевагою є якість і простота у використанні мобільного додатку	Проведення робіт щодо постійного покращення продукту
За інтенсивністю – марочна	Бренд грає велику роль у постачанні продукту	Проведення рекламної кампанії та доведення якості продукту



Ринок є конкурентним, проте вид конкуренції є чистим, так як окремі гравці мало впливають на ціну товару. Конкурентний ринок є міжнародним. Конкуренція за видами товарів – видова.

Таблиця 4.8. Обґрунтування факторів конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння)
1.	Собівартість	Низька собівартість – більша доступність кінцевого пристрою
2.	Продуктивність	Більша продуктивність у порівнянні з конкурентами
3.	Надійність	Збільшення продуктивності продукту збільшує її надійність

Таблиця 4.9. SWOT- аналіз стартап-проекту

<b>Сильні сторони</b>	<b>Слабкі сторони</b>
<ul style="list-style-type: none"> <li>• Здатність адаптуватися та вчитися новому, реагуючи на нові тенденції;</li> <li>• Наявність інтелектуальних ресурсів;</li> <li>• Ентузіазм, згуртований колектив;</li> </ul>	<ul style="list-style-type: none"> <li>• Нове ім'я</li> <li>• Відсутність досвіду</li> <li>• Мала кількість персоналу</li> </ul>

Можливості	Загрози
<ul style="list-style-type: none"> <li>• Розширення асортименту, додаючи нові можливості додатку;</li> <li>• Реклама додатку в ЗМІ;</li> <li>• Можливість задіяти в процес вдосконалення продукту користувачів продукту, використовуючи зворотній зв'язок з ними;</li> </ul>	<ul style="list-style-type: none"> <li>• Вихід на ринок конкурента;</li> <li>• Технічні негаразди;</li> <li>• Нерозуміння з боку користувачів.</li> </ul>

### Зіставлення сторін

#### **1. Стратегія розвитку – використання сильних сторін для реалізації можливостей:**

- ✓ Рекламу продукту можна здійснювати в соціальних мережах, які користуються зараз більшим ефектом довіри у споживачів.
- ✓ Завдяки наявності інтелектуальних ресурсів та згуртованого, готового до нових рішень колективу створити мобільний додаток для використання ресурсу у будь-який час та довільному місці.
- ✓ Враховуючи побажання користувачів вдосконалити мобільний додаток для його зручнішого використання.

#### **2. Стратегія внутрішніх перетворень – укріплення слабких сторін за допомогою використання можливостей:**

- ✓ За допомогою реклами стати ближчим до споживачів, розказавши про свою команду.
- ✓ Використовуючи зворотній зв'язок набрати за необхідністю нових людей у команду.
- ✓ На основі зворотного зв'язку зі споживачами додати в додаток нові функції, спростити його використання.

### 3. Обмеження стратегічного розвитку.

- ✓ Поява конкурента, що має хорошу репутацію в інших напрямках та його перевага над нами, оскільки наша команда є зовсім новою на ринку.
- ✓ Неможливість подолання технічних негараздів через недостатню кількість персоналу.
- ✓ Створення продукту, що буде не зовсім зрозумілий та простий для використання через відсутність досвіду або з цієї ж причини несприйняття громадянами.

### 4. Стратегія потенційних переваг – використання сильних сторін для подолання загроз:

- ✓ Завдяки зворотному зв'язку дізнатися, що не влаштовує користувачів та змінити це.
- ✓ При неможливості подолання технічних негараздів власними силами звернутися за консультацією до інших спеціалістів, оскільки ми завжди відкриті до нових знань.

Таблиця 4.10. Альтернативи ринкового впровадження стартап-проекту

№ п/п	Альтернатива (орієнтований комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
1.	Максимізація власного виграшу (індивідуалізм)	Середня	18 місяців
2.	Максимізація спільного виграшу (кооперація)	Висока	20 місяців
3.	Суперництво	Середня	24 місяці

Була обрана кооперація, як альтернативна ринкова поведінка, так як за відносно не високий термін існує велика ймовірність отримання ресурсів.

#### 4.5. Розроблення ринкової стратегії проекту

Таблиця 4.11. Вибір цільових груп потенційних споживачів

№ п/п	Опис профілю цільової групи потенційних клієнтів	Готовність споживачів сприйняти продукт	Орієнтовний попит в межах цільової групи (сегменту)	Інтенсивність конкуренції в сегменті	Простота входу у сегмент
1.	Державний сектор	-	+	висока	-
2.	Приватний сектор	+	+	висока	+

Які цільові групи обрано: основною характеристикою вибору цільової групи є готовність прийняти продукт. У даній області приватний сектор є більш готовим, адже державний сектор потребує більше дозволів та роз'яснень для введення нового продукту в системи.

Таблиця 4.12. Визначення базової стратегії розвитку

№ п/п	Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції до обраної альтернативи	Базова стратегія розвитку
1.	Індивідуалізм	Стратегія недиференці-	Адаптація до вимог ринку	Стратегія спеціалізації

		йованого маркетингу	використання новацій	
--	--	------------------------	-------------------------	--

Таблиця 4.13. Визначення стратегії позиціонування

№ п/п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкуренто- спроможні позиції власного стартап- проекту	Вибір асоціацій, які мають сформувані комплексну позицію власного проекту
1.	Продуктивність	Стратегія спеціалізації	Продуктивна	Висока швидкодія роботи
2.	Надійність	Стратегія спеціалізації	Якість	Висока надійність роботи

#### 4.6. Розроблення маркетингової програми стартап-проекту

Таблиця 4.14. Визначення ключових переваг концепції потенційного товару

№ п/п	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
1.	Введення швидкої та надійної системи комунікації	Висока продуктивність та надійність	Ціна, продуктивність, надійність

Визначившись з основними перевагами концепції товару, можливе створення відповідної рекламної кампанії для кінцевих клієнтів.

Таблиця 4.15. Формування системи збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
1.	Продаж	Повний супровід товару до замовника	Нульовий рівень	Безпосередній (прямий)

Основним каналом збуту є продаж товару. На старті компанії очікуються відносно невеликі об'єми виробництва, тому на даному етапі можливо обійтись без посередників, і продавати товар напряму клієнтам.

Таблиця 4.16. Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
1.	Консерватор	Реклама SMM Відео-інструкції по використанню товару на ресурсі youtube.com	Висока продуктивність та надійність. Простота користування	Показати переваги технології	Демонстрація можливостей даної системи та принцип її використання
2.	Нерішучий	Інтернет, телебачення	Ціна	Ознайомити клієнтів з існуючою технологією	Реклама повинна вразити клієнта новизною технології

#### 4.7. Висновки до розділу

Було розроблено перший етап створення стартап-проекту. Кожна наукова робота повинна знаходити своє місце в застосуванні у реальному житті, тому стартап-проект може бути практичним відображенням наукової праці.

Спочатку було висвітлено зміст ідеї проекту. Для цього було розглянуто потенційних зацікавлених осіб, які в майбутньому можуть стати клієнтами запропонованої продукції. Також були розглянуті ризики реалізації продукції. Аналіз сильних та слабких сторін дають можливість визначити аспекти, на які слід зробити ставку.

Далі було проведено технічний аудит проекту. Були визначені технології, які використовуватимуться. Запропоновані технології вже існують, проте їх використання є дороговартісним, що може створити певні труднощі.

Після цього було проведено аналіз усіх аспектів ринку. Даний аналіз показав, що імплементація проекту можлива в реальних умовах. Не дивлячись на бар'єри входження і наявність конкурентів, проект матиме успіх через свою високу конкурентоспроможність. На перших порах доцільно буде впровадити гнучку систему цін для нових клієнтів, систему знижок та лояльності для постійних клієнтів.

## Висновки

У роботі, поряд з процесами звукоутворення, особлива увага приділена фонемному складу української мови та її відмінностям від інших мов. Зокрема, українська мова має виразну фонетичну засаду, що спонукає до ретельного дослідження її звуків.

Граматичні форми українських слів дуже мінливі, про що свідчить наявність семи відмінків слів. Це значно відрізняє українську мову від англійської, але поєднує з російською. Така особливість ускладнює розпізнавання української мови, зокрема, технічними засобами.

У порівнянні з російською мовою українську мову відрізняє милозвучність, якій притаманні спрощення у групах приголосних, велика кількість голосних звуків, недозвученість дзвінких приголосних, побічний наголос, як наближення до поезії, відсутність скорочення голосних звуків.

Особливою є також українська лексика, тобто вживання слів, які належать виключно українській мові. Так, специфічними є прізвища і загальні назви, експресивно-оцінні дієслова, а також суфікси, префікси, сполучники, вигуки, складні слова.

Характерною ознакою української мови є повноголосся, а також схожість з італійською мовою, в яких наголошені і ненаголошені голосні звучать однаково.

Проаналізовані інтегральні характеристики мовного сигналу, однією з яких є спектральний склад. Розглянуті короткочасний частотний аналіз, функція кореляції, миттєвий спектр потужності, а також середній спектр потужності мови.



Виконане експериментальне дослідження усереднених спектрів мов індоєвропейської родини, а саме західногерманської, романської, західнослов'янської та східнослов'янської груп мов.

Проаналізовані закономірності спектральних характеристик мов у межах груп споріднених мов та в порівнянні з українською мовою.

Зроблені висновки про найбільшу фонетичну подібність української мови до російської та частково до англійської, що пояснюється впливами цих двох мов на українську мову.

## ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. O. Pedchenko, S. Lunova, I. Rudenko. Speech spectrum of the Ukrainian language// IEE: Conference on Electronics and Nanotechnology, №39, 2019. – p.444– 449.
2. Сучасна українська літературна мова: Підручник / А. П. Грищенко, Л. І. Мацько, М. Я. Плющ та ін.; За ред. А. П. Грищенка. – 3-тє вид., допов. – К.: Вища шк., 2002. –439с.
3. Українська термінологія: комплексний лінгвістичний аналіз: [монографія] / М.О. Вакуленко. – Івано-Франківськ: Фоліант, 2015. – 361с., іл.
4. В. А. Широков. Феноменологія лексикографічних систем / В. А. Широков. – К. : Наукова думка, 2004. – 328с.
5. Луньова С., Дідковський В., Педченко О. Акустика мовотворення/ Навч. посібник. - Рига: LAP LAMBERT Academic Publishing.-2018, 134 с.
6. Тоцька Н. І. Сучасна українська літературна мова: фонетика, орфоєпія, графіка, орфографія / Н. І. Тоцька. – К.: Вища школа, 1981. – 183с.
7. Речевая коммуникация: От звука к высказыванию. – М.: Языки славянских культур, 2012. – 464 с.
8. Покровский Н. Б. Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. – 390с.
9. Вахитов Я. Ш. Теоретические основы электроакустики и электроакустическая аппаратура. – М.: Изд-во «Искусство», 1982. – 415с.
10. Сапожков М. А. Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1963.
11. Харкевич А.А. Спектры и анализ. Физмат, 1962.
12. Gabor D. New possibilities in Speech transmission. Journ. Inst. Electr. Engineers, Vol. 94, Pt. III, №32, Nov. 1947.

13. Винцюк Т. К. Анализ, распознавание и интерпретация речевых сигналов. – Киев: Наук. думка, 1987. – 264с.
14. Базаров В.Г. Основы аудиологии и слухопротезирования. – М.: Медицина, 1984. – с.37– 43.
15. Voirs W.D. Diagnostic Acceptability Measure for Speech Communication System// Proc. 1977. – IEEE ICASSP, p.204–207.
16. Альтман Я.А. Руководство по аудиологии / Я. А. Альтман, Г. А. Таварткиладзе. – М.: ДМК Пресс, 2003. – 360 с.
17. Робота з артикуляції голосних і приголосних звуків української мови. – [Електронний ресурс] / Режим доступу: <http://journlib.univ.kiev.ua/index.php?act=article&article=329>

## ДОДАТКИ

### Додаток А. Усередненні спектри мов за стилями

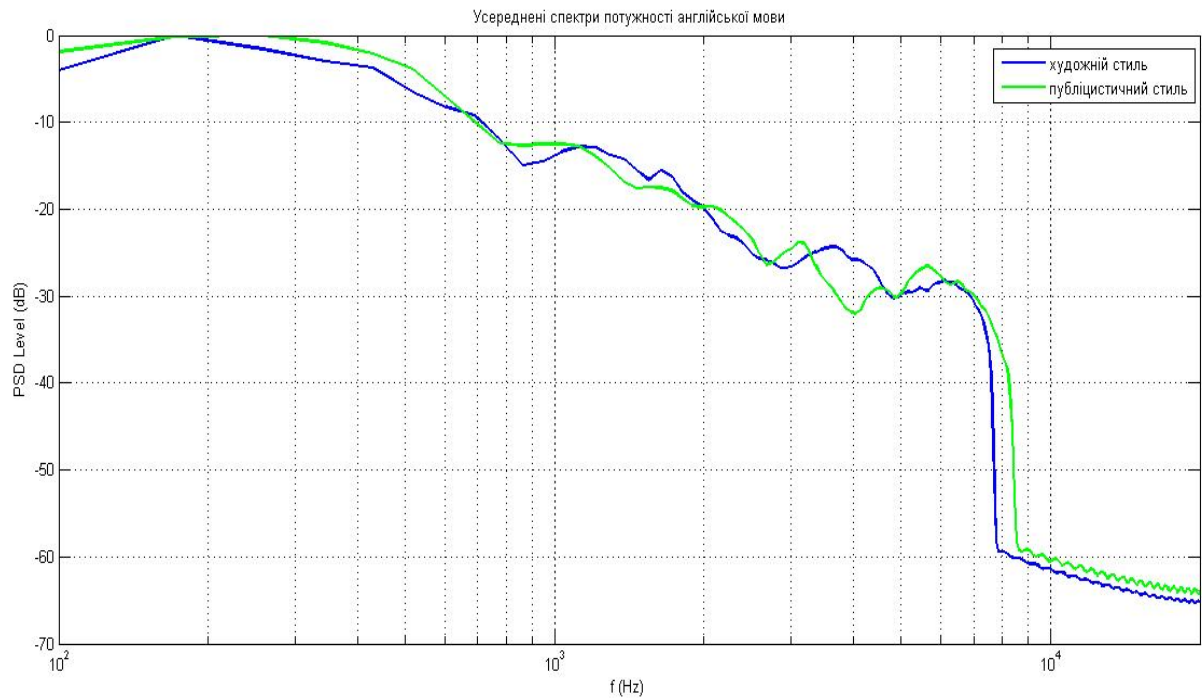


Рис.А.1. Усереднений спектр англійської мови

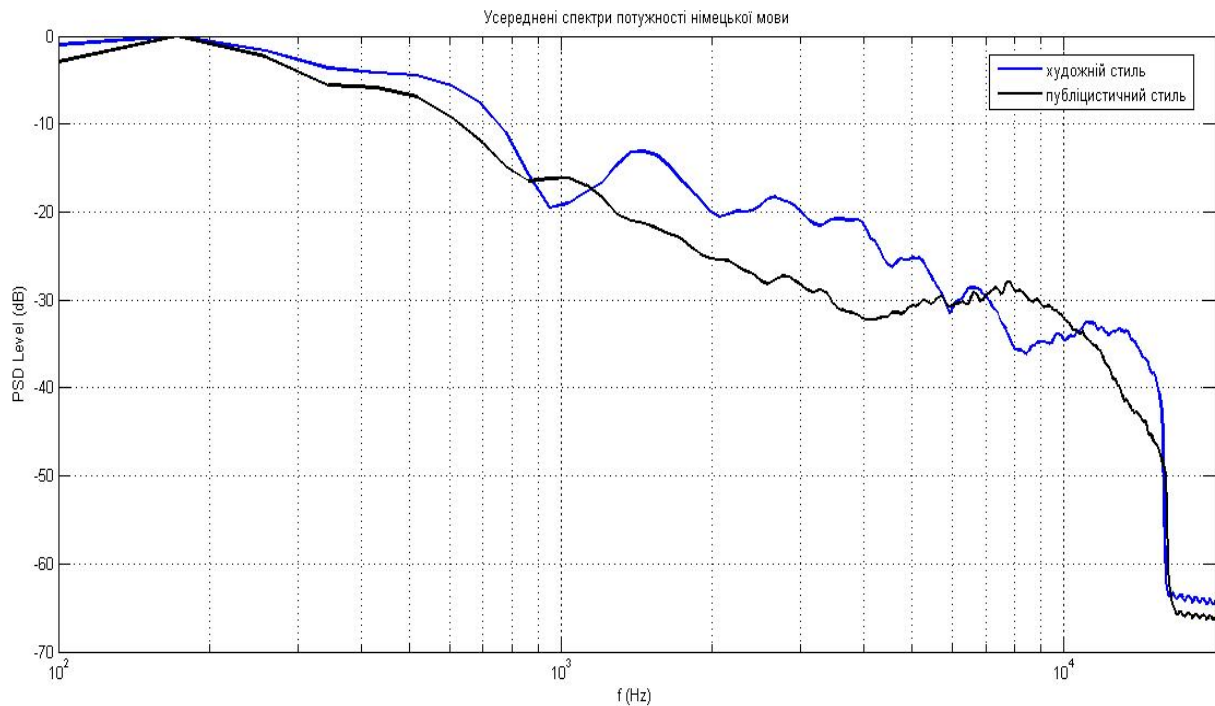


Рис.А.2. Усереднений спектр німецької мови

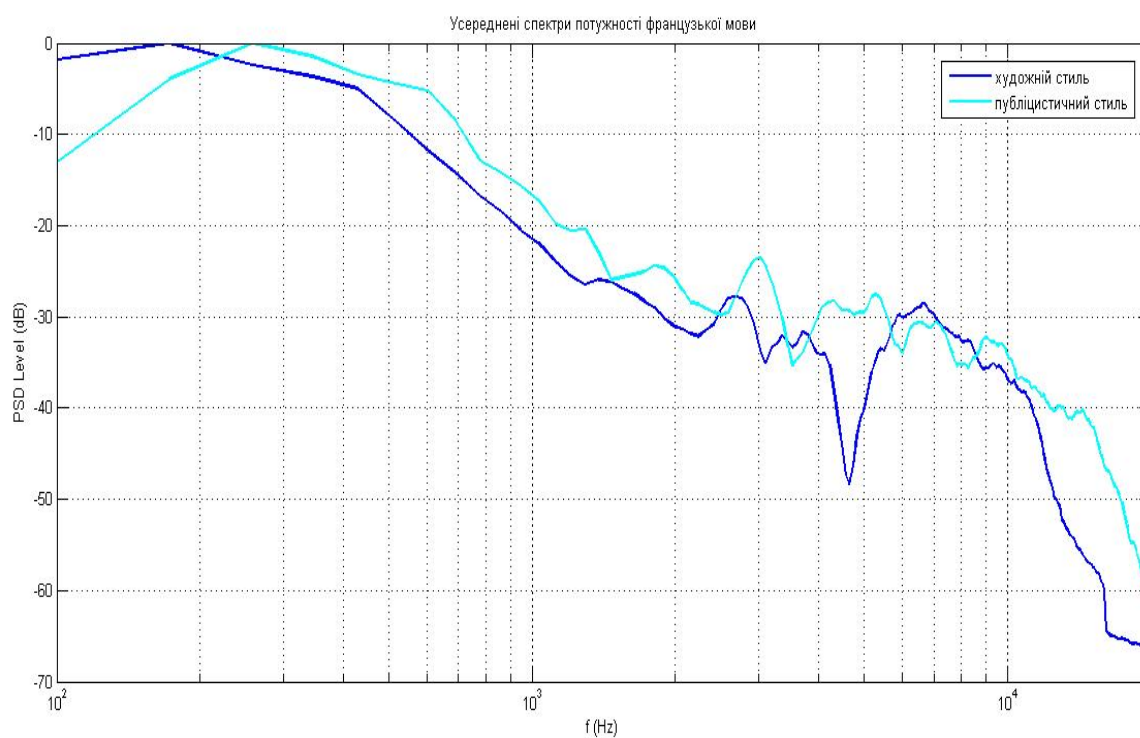


Рис.А.3. Усереднений спектр французької мови

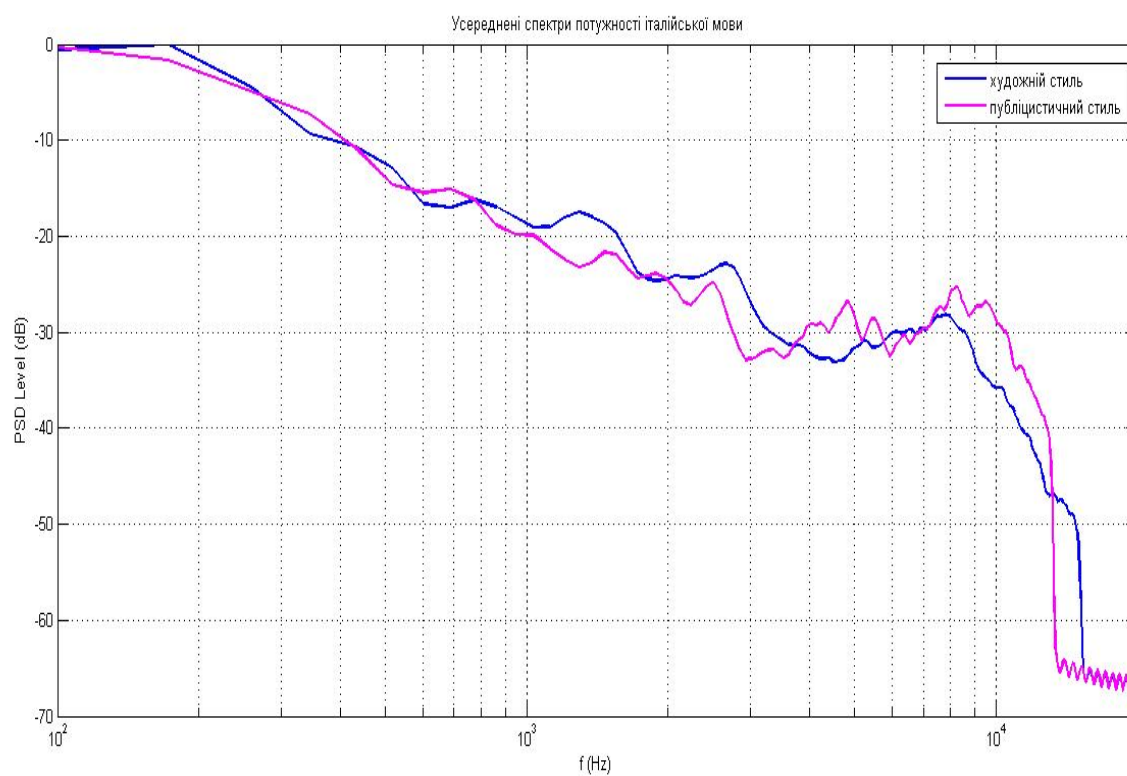


Рис.А.4. Усереднений спектр італійської мови

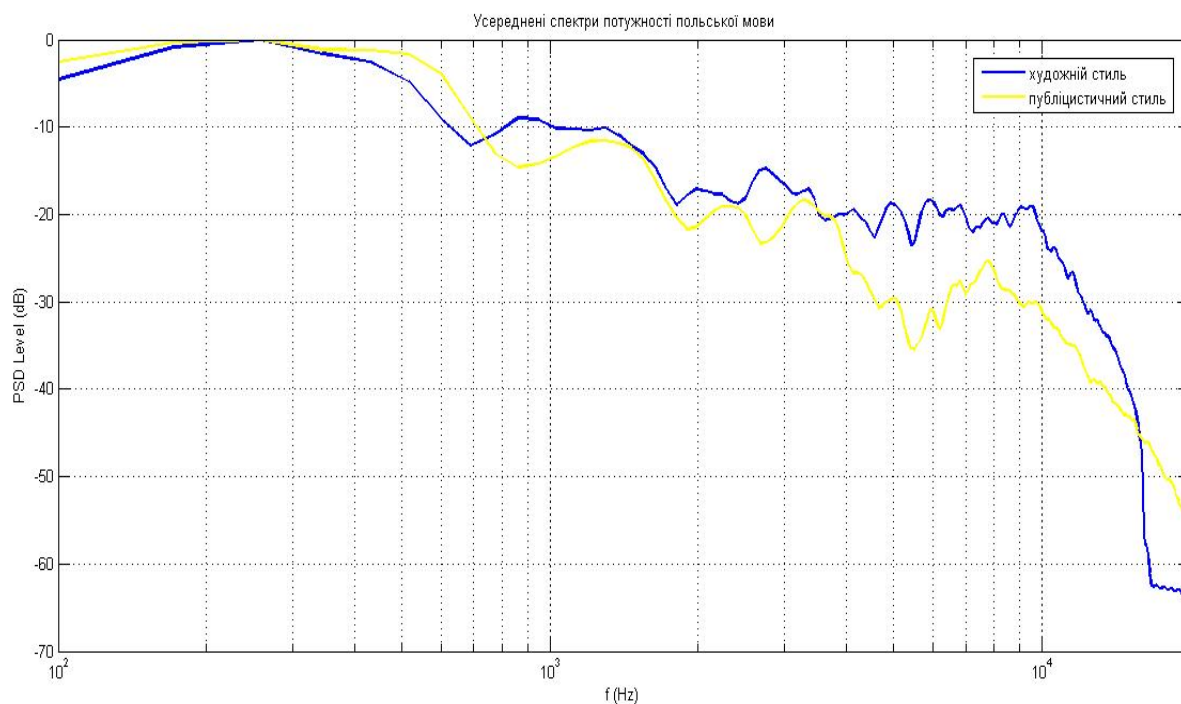


Рис.А.5. Усереднений спектр польської мови

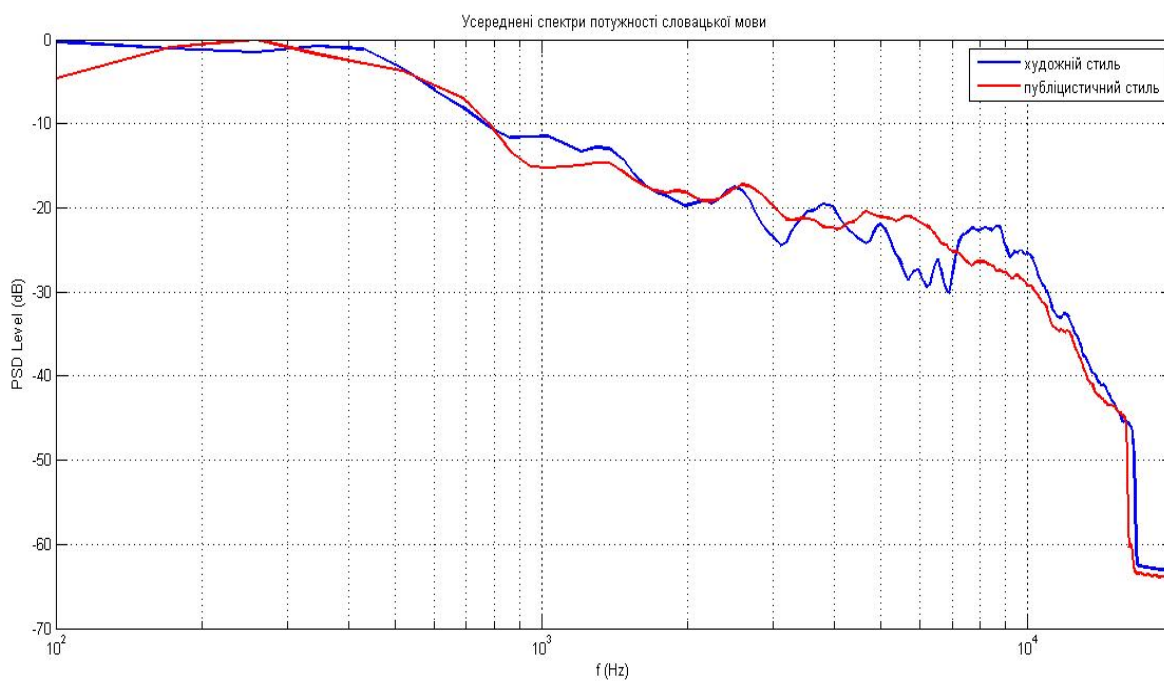


Рис.А.6. Усереднений спектр словацької мови

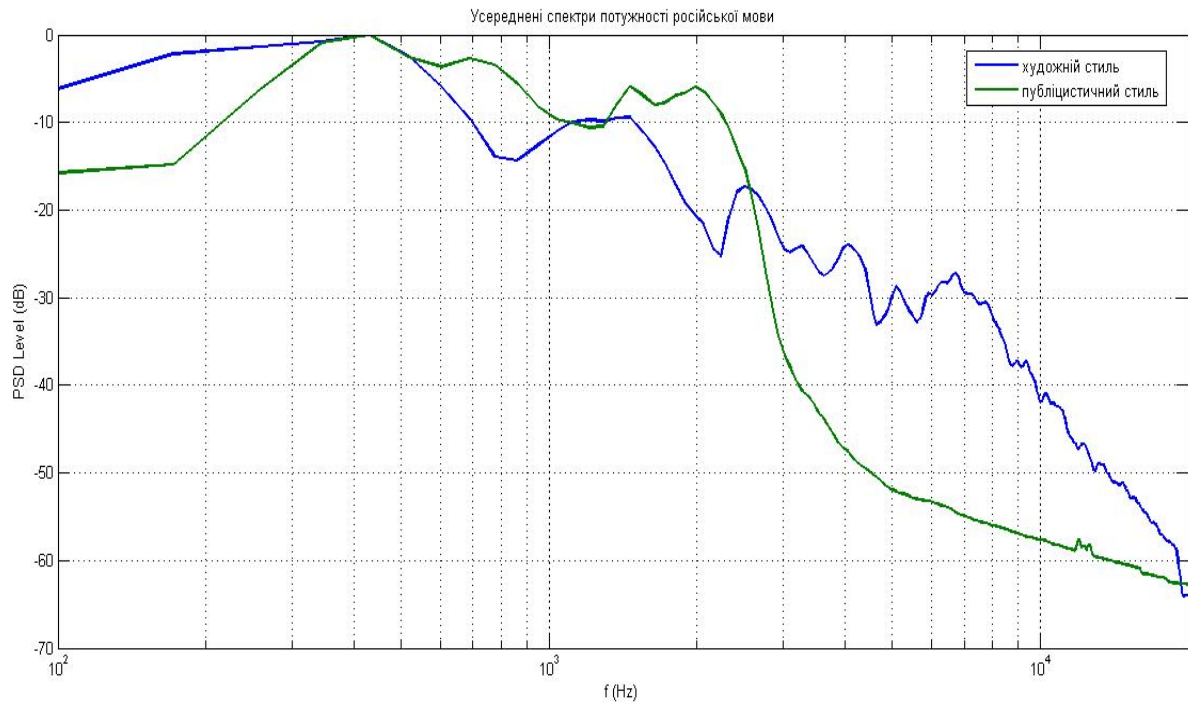


Рис.А.7. Усереднений спектр російської мови

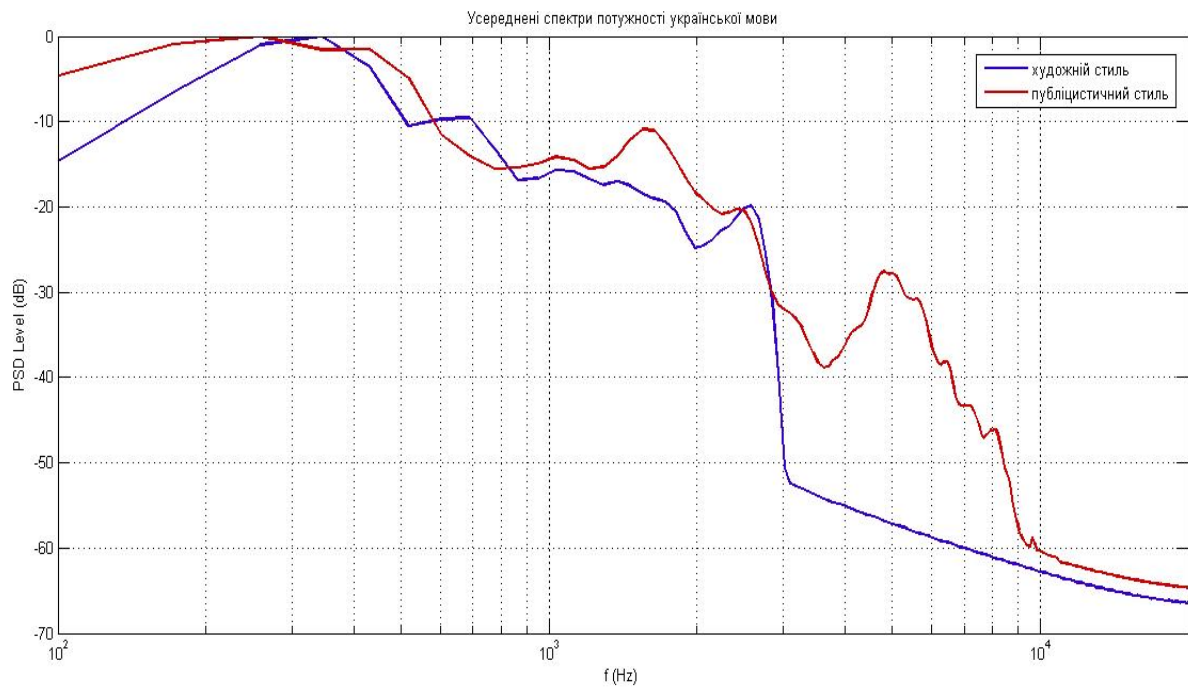


Рис.А.8. Усереднений спектр української мови

## Додаток Б. Порівняння усереднених спектрів

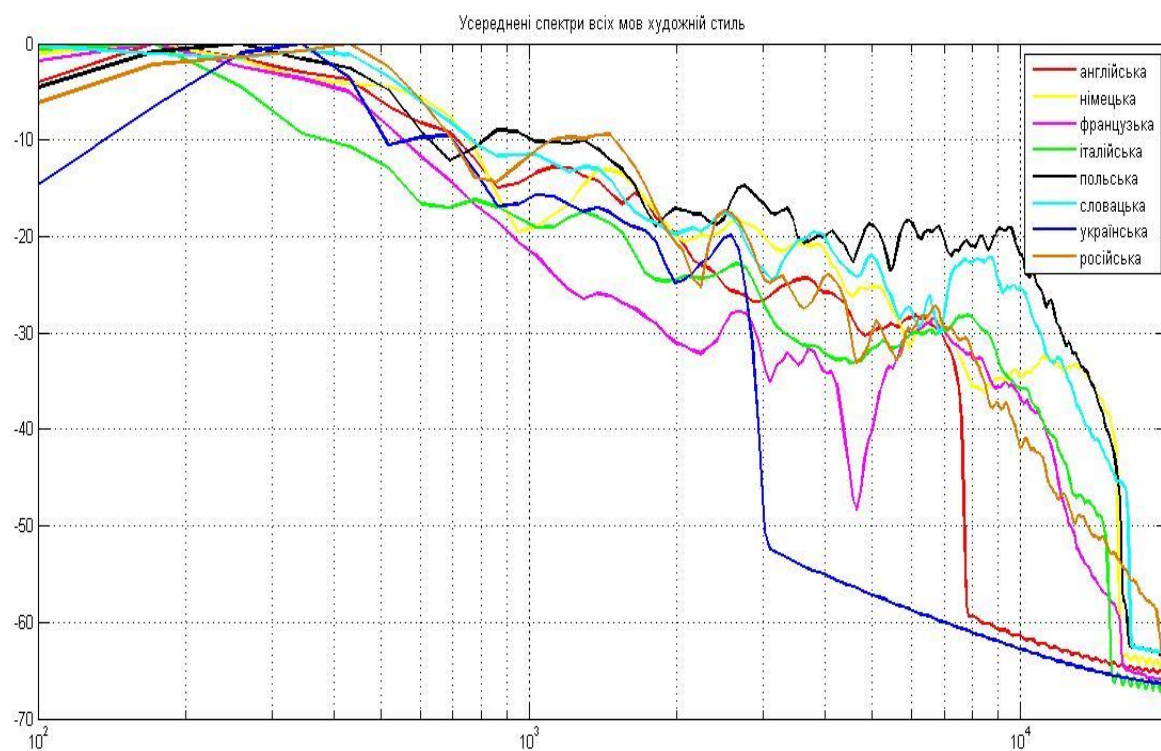


Рис.Б.1. Усереднені спектри всіх мов, художній стиль

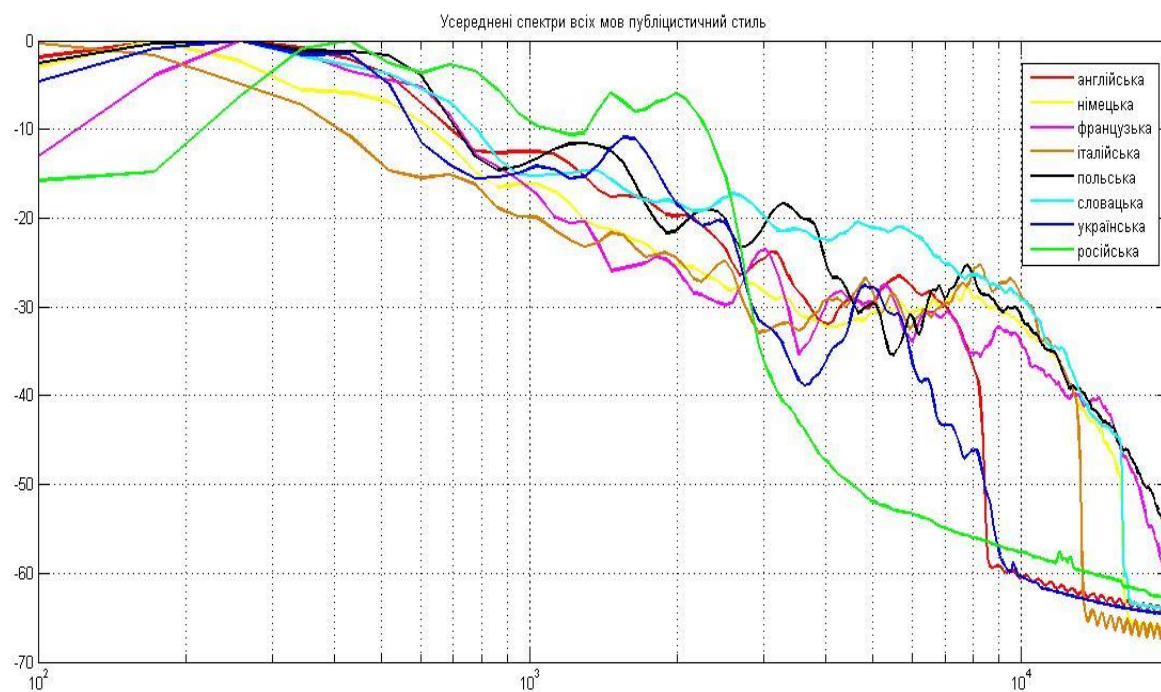




Рис.Б.2. Усереднені спектри всіх мов, публіцистичний стиль

### Додаток В. Код програми для побудови графіків усередненого спектру (основний код)

```

% англійська мова
[y1,Fs1]=audioread('01 - Chapter 1 (2).wav');
y1=y1(:,2);
[y2,Fs2]=audioread('01 - Chapter 1 (3).wav'); %диктор 1 художній
текст
y2=y2(:,2);
[y3,Fs3]=audioread('Truth Worth Telling-Part01 публицистический
(2).wav');
y3=y3(:,2);
[y4,Fs4]=audioread('Truth Worth Telling-Part01 публицистический
(3).wav'); %диктор 2 публіцистичний текст
y4=y4(:,2);
%стандартне відхилення сигналу
stdn1=std(y1);
stdn2=std(y2);
stdn3=std(y3);
stdn4=std(y4);
%нормування сигналу
yn1=y1./stdn1;
yn2=y2./stdn2;
yn3=y3./stdn3;
yn4=y4./stdn4;
% %спектр потужності
[ypn1,f1]=pwelch(yn1,hamming(441),[],[],Fs1);
[ypn2,f2]=pwelch(yn2,hamming(441),[],[],Fs2);
[ypn3,f3]=pwelch(yn3,hamming(441),[],[],Fs3);
[ypn4,f4]=pwelch(yn4,hamming(441),[],[],Fs4);
ypn1=ypn1';
ypn2=ypn2';
ypn3=ypn3';
ypn4=ypn4';
%усереднення оцінок спектрів потужності
y1=[ypn1;ypn2];%англійська художній
yuser1=mean(y1);
k1=max(abs(yuser1));
norm1=yuser1./k1;
norml1=10.*log10(norm1);
plot(f1,norml1)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності англійської мови художній
стиль');
grid on;

y2=[ypn3;ypn4];%англійська публіцистичний
yuser2=mean(y2);
k2=max(abs(yuser2));

```

```

norm2=yser2./k2;
norml2=10.*log10(norm2);
figure
plot(f3,norml2)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності англійської мови
публіцистичний стиль');
grid on;

y_1=[ypn1;ypn2;ypn3;ypn4];
yser_1=mean(y_1);
k_1=max(abs(yser_1));
norm_1=yser_1./k_1;
norml_1=10.*log10(norm_1);
figure
plot(f1,norml_1)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності англійської мови');
grid on;

%німецька
[y13,Fs13]=audioread('Schatzinsel_CD1-5 (2).wav');
y13=y13(:,2);
[y14,Fs14]=audioread('Schatzinsel_CD1-5 (3).wav');%диктор 7
художній текст
y14=y14(:,2);
[y15,Fs15]=audioread('Wissenschaftsmagazin_19-10-2019-
1240.1571477410059 (публіцистический) (2).wav');
y15=y15(:,2);
[y16,Fs16]=audioread('Wissenschaftsmagazin_19-10-2019-
1240.1571477410059 (публіцистический) (3).wav');%диктор 8
публіцистичний текст
y16=y16(:,2);
stdn13=std(y13);
stdn14=std(y14);
stdn15=std(y15);
stdn16=std(y16);
yn13=y13./stdn13;
yn14=y14./stdn14;
yn15=y15./stdn15;
yn16=y16./stdn16;
[ypn13,f13]=pwelch(yn13,hamming(441),[],[],Fs13);
[ypn14,f14]=pwelch(yn14,hamming(441),[],[],Fs14);
[ypn15,f15]=pwelch(yn15,hamming(441),[],[],Fs15);
[ypn16,f16]=pwelch(yn16,hamming(441),[],[],Fs16);
ypn13=ypn13';
ypn14=ypn14';
ypn15=ypn15';
ypn16=ypn16';
y7=[ypn13;ypn14];%німецька художній

```

```

yser7=mean(y7);
k7=max(abs(yser7));
norm7=yser7./k7;
norml7=10.*log10(norm7);
figure
plot(f13,norml7)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності німецької мови удожній
стиль');
grid on;

```

```

y8=[ypn15;ypn16];%німецька публіцистичний
yser8=mean(y8);
k8=max(abs(yser8));
norm8=yser8./k8;
norml8=10.*log10(norm8);
figure
plot(f15,norml8)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності німецької мови
публіцистичний стиль');
grid on;

```

```

y_4=[ypn13;ypn14;ypn15;ypn16];
yser_4=mean(y_4);
k_4=max(abs(yser_4));
norm_4=yser_4./k_4;
norml_4=10.*log10(norm_4);
figure
plot(f13,norml_4)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності німецької мови');
grid on;

```

```

%французька
[y25,Fs25]=audioread('Arthur_Conan_Doyle_-_Les_six_napoleons
(2).wav');
y25=y25(:,2);
[y26,Fs26]=audioread('Arthur_Conan_Doyle_-_Les_six_napoleons
(3).wav');%диктор 13 художній текст
y26=y26(:,2);
[y27,Fs27]=audioread('Charles_Darwin_-_
_L_Origine_des_especies_Chap01_P1 публицистический (2).wav');
y27=y27(:,2);
[y28,Fs28]=audioread('Charles_Darwin_-_
_L_Origine_des_especies_Chap01_P1 публицистический
(3).wav');%диктор 14 публіцистичний текст
y28=y28(:,2);
stdn25=std(y25);

```

```

stdn26=std(y26);
stdn27=std(y27);
stdn28=std(y28);
yn25=y25./stdn25;
yn26=y26./stdn26;
yn27=y27./stdn27;
yn28=y28./stdn28;
[ypn25,f25]=pwelch(yn25,hamming(441),[],[],Fs25);
[ypn26,f26]=pwelch(yn26,hamming(441),[],[],Fs26);
[ypn27,f27]=pwelch(yn27,hamming(441),[],[],Fs27);
[ypn28,f28]=pwelch(yn28,hamming(441),[],[],Fs28);
ypn25=ypn25';
ypn26=ypn26';
ypn27=ypn27';
ypn28=ypn28';
%усереднення оцінок спектрів потужності
y13=[ypn25;ypn26];% французький художній
yuser13=mean(y13);
k13=max(abs(yuser13));
norm13=yuser13./k13;
norml13=10.*log10(norm13);
figure
plot(f25,norml13)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності французької мови художній
стиль');
grid on;

y14=[ypn27;ypn28];%французький публіцистичний
yuser14=mean(y14);
k14=max(abs(yuser14));
norm14=yuser14./k14;
norml14=10.*log10(norm14);
figure
plot(f27,norml14)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності французької мови
публіцистичний стиль');
grid on;

y_3=[ypn25;ypn26;ypn27;ypn28];
yuser_3=mean(y_3);
k_3=max(abs(yuser_3));
norm_3=yuser_3./k_3;
norml_3=10.*log10(norm_3);
figure
plot(f25,norml_3)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності французької мови');
grid on;

```

```

%італійська
[y9,Fs9]=audioread('Verne-20 000 leghe sotto i mari (02)
(2).wav');
y9=y9(:,2);
[y10,Fs10]=audioread('Verne-20 000 leghe sotto i mari (02)
(3).wav');%диктор 5 художній текст
y10=y10(:,2);
[y11,Fs11]=audioread('Cambia la tua vita in 7 giorni
публицистический (2).wav');
y11=y11(:,2);
[y12,Fs12]=audioread('Cambia la tua vita in 7 giorni
публицистический (3).wav');%диктор 6 публіцистичний текст
y12=y12(:,2);
stdn9=std(y9);
stdn10=std(y10);
stdn11=std(y11);
stdn12=std(y12);
yn9=y9./stdn9;
yn10=y10./stdn10;
yn11=y11./stdn11;
yn12=y12./stdn12;
[ypn9,f9]=pwelch(yn9,hamming(441),[],[],Fs9);
[ypn10,f10]=pwelch(yn10,hamming(441),[],[],Fs10);
[ypn11,f11]=pwelch(yn11,hamming(441),[],[],Fs11);
[ypn12,f12]=pwelch(yn12,hamming(441),[],[],Fs12);

ypn9=ypn9';
ypn10=ypn10';
ypn11=ypn11';
ypn12=ypn12';

y5=[ypn9;ypn10];%італійська художній
yser5=mean(y5);
k5=max(abs(yser5));
norm5=yser5./k5;
norml5=10.*log10(norm5);
figure
plot(f9,norml5)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності італійської мови художній
стиль');
grid on;

y6=[ypn11;ypn12];%італійська публіцистичний
yser6=mean(y6);
k6=max(abs(yser6));
norm6=yser6./k6;
norml6=10.*log10(norm6);
figure
plot(f11,norml6)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');

```

```

title('Усереднений спектр потужності італійської мови
публіцистичний стиль');
grid on;

y_5=[ypn9;ypn10;ypn11;ypn12];
yser_5=mean(y_5);
k_5=max(abs(yser_5));
norm_5=yser_5./k_5;
norml_5=10.*log10(norm_5);
figure
plot(f9,norml_5)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності італійської мови');
grid on;
% словацька
[y21,Fs21]=audioread('04 Kapitola 1 (2).wav');
y21=y21(:,2);
[y22,Fs22]=audioread('04 Kapitola 1 (3).wav');%диктор 11
художній текст
y22=y22(:,2);
[y23,Fs23]=audioread('01 - Track 1 (2).wav');
y23=y23(:,2);
[y24,Fs24]=audioread('01 - Track 1 (3).wav');%диктор 12
публіцистичний текст
y24=y24(:,2);
stdn21=std(y21);
stdn22=std(y22);
stdn23=std(y23);
stdn24=std(y24);
yn21=y21./stdn21;
yn22=y22./stdn22;
yn23=y23./stdn23;
yn24=y24./stdn24;
[ypn21,f21]=pwelch(yn21,hamming(441),[],[],Fs21);
[ypn22,f22]=pwelch(yn22,hamming(441),[],[],Fs22);
[ypn23,f23]=pwelch(yn23,hamming(441),[],[],Fs23);
[ypn24,f24]=pwelch(yn24,hamming(441),[],[],Fs24);
ypn21=ypn21';
ypn22=ypn22';
ypn23=ypn23';
ypn24=ypn24';

y11=[ypn21;ypn22];% словацька художній
yser11=mean(y11);
k11=max(abs(yser11));
norm11=yser11./k11;
norml11=10.*log10(norm11);
figure
plot(f21,norml11)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');

```

```

title('Усереднений спектр потужності словацької мови художній
стиль');
grid on;

y12=[ypn23;ypn24];%словацька публіцистичний
yuser12=mean(y12);
k12=max(abs(yuser12));
norm12=yuser12./k12;
norml12=10.*log10(norm12);
figure
plot(f23,norml12)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності словацької мови
публіцистичний стиль');
grid on;

y_8=[ypn21;ypn22;;ypn23;ypn24];
yuser_8=mean(y_8);
k_8=max(abs(yuser_8));
norm_8=yuser_8./k_8;
norml_8=10.*log10(norm_8);
figure
plot(f21,norml_8)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності словацької мови');
grid on;
%польська
[y17,Fs17]=audioread('01-e-t-a-hoffmann-dziadek-do-orzechow
(2).wav');
y17=y17(:,2);
[y18,Fs18]=audioread('01-e-t-a-hoffmann-dziadek-do-orzechow
(3).wav');%диктор 9 художній текст
y18=y18(:,2);
[y19,Fs19]=audioread('002_POŁACTWO публицистический (2).wav');
y19=y19(:,2);
[y20,Fs20]=audioread('002_POŁACTWO публицистический
(3).wav');%диктор 10 публіцистичний текст
y20=y20(:,2);
stdn17=std(y17);
stdn18=std(y18);
stdn19=std(y19);
stdn20=std(y20);
yn17=y17./stdn17;
yn18=y18./stdn18;
yn19=y19./stdn19;
yn20=y20./stdn20;
[ypn17,f17]=pwelch(yn17,hamming(441),[],[],Fs17);
[ypn18,f18]=pwelch(yn18,hamming(441),[],[],Fs18);
[ypn19,f19]=pwelch(yn19,hamming(441),[],[],Fs19);
[ypn20,f20]=pwelch(yn20,hamming(441),[],[],Fs20);
ypn17=ypn17';

```

```

ypn18=ypn18';
ypn19=ypn19';
ypn20=ypn20';
y9=[ypn17;ypn18];% польська художній
yсер9=mean(y9);
k9=max(abs(yсер9));
norm9=yсер9./k9;
norml9=10.*log10(norm9);
figure
plot(f17,norml9)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності польської мови художній
стиль');
grid on;

y10=[ypn19;ypn20];%польська публіцистичний
yсер10=mean(y10);
k10=max(abs(yсер10));
norm10=yсер10./k10;
norml10=10.*log10(norm10);
figure
plot(f19,norml10)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності польської мови
публіцистичний стиль');
grid on;

y_6=[ypn17;ypn18;ypn19;ypn20];
yсер_6=mean(y_6);
k_6=max(abs(yсер_6));
norm_6=yсер_6./k_6;
norml_6=10.*log10(norm_6);
figure
plot(f17,norml_6)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності польської мови');
grid on;

%українська мова
[y29,Fs29]=audioread('knyaz ukr-1.wav');%художній
y29=y29(:,2);
[y30,Fs30]=audioread('knyaz ukr-2.wav');
y30=y30(:,2);
[y31,Fs31]=audioread('knyaz ukr-3.wav');
y31=y31(:,2);

[y32,Fs32]=audioread('gulag ukr-1.wav');%публіцистичний
y32=y32(:,2);
[y33,Fs33]=audioread('gulag ukr-2.wav');

```



```

y33=y33(:,2);
[y34,Fs34]=audioread('gulag ukr-3.wav');
y34=y34(:,2);

stdn29=std(y29);
stdn30=std(y30);
stdn31=std(y31);
stdn32=std(y32);
stdn33=std(y33);
stdn34=std(y34);

yn29=y29./stdn29;
yn30=y30./stdn30;
yn31=y31./stdn31;
yn32=y32./stdn32;
yn33=y33./stdn33;
yn34=y34./stdn34;
[ypn29,f29]=pwelch(yn29,hamming(441),[],[],Fs29);
[ypn30,f30]=pwelch(yn30,hamming(441),[],[],Fs30);
[ypn31,f31]=pwelch(yn31,hamming(441),[],[],Fs31);
[ypn32,f32]=pwelch(yn32,hamming(441),[],[],Fs32);
[ypn33,f33]=pwelch(yn33,hamming(441),[],[],Fs33);
[ypn34,f34]=pwelch(yn34,hamming(441),[],[],Fs34);

ypn29=ypn29';
ypn30=ypn30';
ypn31=ypn31';
ypn32=ypn32';
ypn33=ypn33';
ypn34=ypn34';

y15=[ypn29;ypn30;ypn31];%українська художній
y15=mean(y15);
k15=max(abs(y15));
norm15=y15./k15;
norml15=10.*log10(norm15);
figure
plot(f29,norml15)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності української мови художній
стиль');
grid on;

y16=[ypn32;ypn33;ypn34];%українська публіцистичний
y16=mean(y16);
k16=max(abs(y16));
norm16=y16./k16;
norml16=10.*log10(norm16);
figure
plot(f32,norml16)
xlabel('f (Hz)');

```

```

ylabel('PSD Level (dB)');
title('Усереднений спектр потужності української мови
публіцистичний стиль');
grid on;

y_2=[ypn29;ypn30;ypn31;ypn32;ypn33;ypn34];
yser_2=mean(y_2);
k_2=max(abs(yser_2));
norm_2=yser_2./k_2;
norml_2=10.*log10(norm_2);
figure
plot(f29,norml_2)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності української мови');
grid on;
%російська мова
[y35,Fs35]=audioread('viy-1.wav');%художній стиль
y35=y35(:,2);
[y36,Fs36]=audioread('viy-2.wav');
y36=y36(:,2);
[y37,Fs37]=audioread('viy-3.wav');
y37=y37(:,2);

[y38,Fs38]=audioread('nauki rus-1.wav');%публіцистичний
y38=y38(:,2);
[y39,Fs39]=audioread('nauki rus-2.wav');
y39=y39(:,2);
[y40,Fs40]=audioread('nauki rus-3.wav');
y40=y40(:,2);

stdn35=std(y35);
stdn36=std(y36);
stdn37=std(y37);
stdn38=std(y38);
stdn39=std(y39);
stdn40=std(y40);

yn35=y35./stdn35;
yn36=y36./stdn36;
yn37=y37./stdn37;
yn38=y38./stdn38;
yn39=y39./stdn39;
yn40=y40./stdn40;
[ypn35,f35]=pwelch(yn35,hamming(441),[],[],Fs35);
[ypn36,f36]=pwelch(yn36,hamming(441),[],[],Fs36);
[ypn37,f37]=pwelch(yn37,hamming(441),[],[],Fs37);
[ypn38,f38]=pwelch(yn38,hamming(441),[],[],Fs38);
[ypn39,f39]=pwelch(yn39,hamming(441),[],[],Fs39);
[ypn40,f40]=pwelch(yn40,hamming(441),[],[],Fs40);

ypn35=ypn35';
ypn36=ypn36';

```

```

ypn37=ypn37';
ypn38=ypn38';
ypn39=ypn39';
ypn40=ypn40';

y17=[ypn35;ypn36;ypn37];%російська художній
yser17=mean(y17);
k17=max(abs(yser17));
norm17=yser17./k17;
norml17=10.*log10(norm17);
figure
plot(f35,norml17)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності російської мови художній
стиль');
grid on;

y18=[ypn38;ypn39;ypn40];%російська публіцистичний
yser18=mean(y18);
k18=max(abs(yser18));
norm18=yser18./k18;
norml18=10.*log10(norm18);
figure
plot(f38,norml18)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності російської мови
публіцистичний стиль');
grid on;

y_7=[ypn35;ypn36;ypn37;ypn38;ypn39;ypn40];
yser_7=mean(y_7);
k_7=max(abs(yser_7));
norm_7=yser_7./k_7;
norml_7=10.*log10(norm_7);
figure
plot(f35,norml_7)
xlabel('f (Hz)');
ylabel('PSD Level (dB)');
title('Усереднений спектр потужності російської мови');
grid on;

```