

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»

Факультет електроніки

(повна назва інституту/факультету)

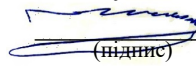
Кафедра акустичних та мультимедійних електронних систем

(повна назва кафедри)

«На правах рукопису»  
УДК 621.391.83

«До захисту допущено»

Завідувач кафедри

 С. А. Найда  
(підпис) (ініціали, прізвище)

“04” червня 2022 р.

## Магістерська дисертація

зі спеціальності (спеціалізації) 171 Електроніка

(код і назва спеціальності)

на тему: "Система комп'ютерної підтримки лікаря-фоніатра при розв'язанні  
діагностичних завдань"

Виконав (-ла): студент (-ка) 6 курсу, групи ДГ-01мн

(шифр групи)

Денисенко Олександр Ігорович

(прізвище, ім'я, по батькові)

  
(підпис)

Науковий керівник доктор технічних наук, професор кафедри  
акустичних та мультимедійних електронних систем КПІ ім.  
Сікорського Продеус А.М.

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

  
(підпис)

Консультант

(назва розділу)

(науковий ступінь, вчене звання, прізвище, ініціали)

(підпис)

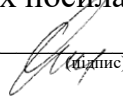
Рецензент канд. техн. наук, доц. каф. ел. інженерії Шуляк О.П.

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

  
(підпис)

Засвідчую, що у цій магістерській дисертації  
немає запозичень з праць інших авторів без  
відповідних посилань.

Студент

  
(підпис)

Київ – 2022 року

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

Інститут (факультет) електроніки \_\_\_\_\_  
(повна назва)

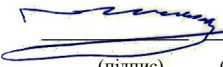
Кафедра акустичних та мультимедійних електронних систем \_\_\_\_\_  
(повна назва)

Рівень вищої освіти – другий (магістерський)

Спеціальність 171 Електроніка \_\_\_\_\_  
(код і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри

 С. А. Найда  
(підпис) (ініціали, прізвище)

«21» вересня 2020 р.

**ЗАВДАННЯ**  
**на магістерську дисертацію студенту**  
Денисенку Олександровичу  
(прізвище, ім'я, по батькові)

1. Тема дисертації: "Система комп'ютерної підтримки лікаря-фоніатра при розв'язанні діагностичних завдань"

науковий керівник дисертації Продеус Аркадій Миколайович, д.т.н., професор,  
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «14» квітня 2022 р. №НС/8/2022

2. Строк подання студентом дисертації 30 травня 2022 р. \_\_\_\_\_

3. Об'єкт дослідження: Процеси вимірювання діагностичних ознак при оцінюванні стану голосового апарату людини

4. Предмет дослідження: методи об'єктивного (інструментального) аналізу стану голосового тракту та інформативні, для задачі діагностики стану голосового апарату людини, параметри та алгоритми їх оцінювання

5. Перелік завдань, які потрібно розробити: 1) аналітичний огляд літературних джерел; 2) вибір інформативних, для задач фоніатрії, параметрів голосних звуків; 3) розробка та модернізація алгоритмів обчислення інформативних параметрів; 4) макетування системи оцінювання діагностичних ознак стану голосового апарату людини

6. Перелік графічного (ілюстративного) матеріалу: Презентація із 15 слайдів.

7. Орієнтовний перелік публікацій: Денисенко О. І. Комп'ютерне моделювання алгоритмів діагностики голосової системи людини / Олександр Ігорович Денисенко // Мікросистеми, Електроніка та Акустика. – 2022 р. (в публікації).

8. Консультанти розділів дисертації\*

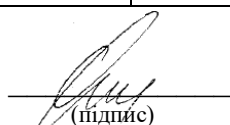
Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

9. Дата видачі завдання 03.09.2021

#### Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	Аналітичний огляд літературних джерел	4 листопада 2021 р.	Виконано
2	Вибір інформативних, для задач фоніатрії, параметрів голосних звуків	15 січня 2022 р.	Виконано
3	Опис алгоритмів інструментарію та його можливостей для створення програмного забезпечення	15 березня 2022 р.	Виконано
4	Макетування програмного продукту	10 травня 2022 р.	Виконано
5	Оформлення магістерської дисертації	30 травня 2022 р.	Виконано

Студент

  
(підпис)

О.І. Денисенко  
(ініціали, прізвище)

Науковий керівник дисертації



А.М. Продеус

\* Консультантом не може бути зазначено наукового керівника

## РЕФЕРАТ

Об'єктом роботи є процеси вимірювання діагностичних ознак при оцінюванні стану голосового апарату людини.

Предметом роботи є методи об'єктивного (інструментального) аналізу стану голосового тракту та інформативні, для задачі діагностики стану голосового апарату людини, параметри та алгоритми їх оцінювання.

Метою магістерської роботи є аналіз існуючих діагностичних ознак визначення стану голосу людини та розробка алгоритмів їх оцінювання на базі дослідженого та вивченого інструментарію розробки подібних систем.

У даній роботі проведено аналіз існуючих діагностичних ознак органів голосу людини та обрано методи та алгоритми, що найкращим чином підходять для їх розрахунку. За обраними методами та алгоритмами проведено модернізацію існуючого інтерфейсного рішення.

На базі проведеного аналізу виконано розробку власного програмного забезпечення, спрямованого на забезпечення вимог лікарів-фоніатрів. Воно має сучасний вигляд та є гнучким у сенсі можливості його майбутнього удосконалення.

Ключові слова: голосовий тракт, спектральний аналіз, мел-частотні кепстральні коефіцієнти, HNR, DSI, Shimmer, Jitter.

## **ABSTRACT**

The object of work is the process of measuring diagnostic signs in assessing the state of the human vocal apparatus.

The subject of the work is the methods of objective (instrumental) analysis of the state of the vocal tract and informative, for the task of diagnosing the state of the human vocal apparatus, parameters and algorithms for their evaluation.

The purpose of the master's work is to analyze the existing diagnostic features of determining the state of the human voice and the development of algorithms for their evaluation on the basis of researched and studied tools for the development of such systems.

This paper analyzes the existing diagnostic features of the human voice and selects methods and algorithms that are best suited for their calculation. According to the selected methods and algorithms the modernization of the existing interface solution was carried out.

On the basis of the conducted analysis the development of the own software directed on maintenance of requirements of phoniaticians is executed. It has a modern look and is flexible in terms of the possibility of its future improvement.

Key words: voice tract, spectral analysis, mel-frequency cepstral coefficients, HNR, DSI, Shimmer, Jitter.

# Зміст

<b>ВСТУП .....</b>	<b>7</b>
<b>РОЗДІЛ 1. ГОЛОСОВИЙ АПАРАТ ТА КОМП'ЮТЕРНИЙ АНАЛІЗ ЙОГО СТАНУ .....</b>	<b>10</b>
1.1    Анатомія та фізіологія голосового апарату, а також його фізичні та комп'ютерні моделі .....	10
1.2.    Фізико-математичні моделі голосових зв'язок .....	15
1.3.    Акустичні дослідження голосу з використанням комп'ютерних технологій .....	25
1.4.    Характеристика голосової функції за патології голосового апарату за даними комп'ютерного спектрального аналізу голосу.....	30
1.5    Характеристичні ознаки голосу в діагностичній системі LingWaves.....	42
<b>ВИСНОВКИ ДО 1 РОЗДІЛУ .....</b>	<b>53</b>
<b>РОЗДІЛ 2. АЛГОРИТМИ ТА МАКЕТУВАННЯ ВІДПОВІДНИХ ПРОГРАМНИХ ФУНКЦІЙ РОЗРАХУНКУ ДІАГНОСТИЧНИХ ОЗНАК .....</b>	<b>54</b>
2.1    Трасування частоти основного тону .....	54
2.2.    Розрахунок Shimmer та Jitter.....	57
2.3    Розрахунок відношення гармонік до шуму (HNR) .....	59
2.4.    Час максимальної фонації.....	59
2.5    Індекс тяжкості дисфонії .....	60
2.6    Мел-частотні кепстральні коефіцієнти .....	60
<b>ВИСНОВКИ ДО 2 РОЗДІЛУ .....</b>	<b>65</b>
<b>РОЗДІЛ 3. РОЗРОБКА ІНТЕРФЕЙСУ ПРОГРАМИ КОМП'ЮТЕРНОЇ ПІДТРИМКИ ФОНІАТРА .....</b>	<b>67</b>
3.1    Демонстрація попередньої розробки.....	67
3.2.    Впровадження нових розроблених функцій в існуючий інтерфейс .....	68
<b>ВИСНОВКИ ДО 3 РОЗДІЛУ .....</b>	<b>71</b>
<b>ВИСНОВКИ .....</b>	<b>72</b>
<b>Літературні джерела .....</b>	<b>73</b>
<b>ДОДАТОК А. Скрипти програмних модулів .....</b>	<b>76</b>

## ВСТУП

Загальні екологічні та економічні тенденції розвитку суспільства приводять до збільшення випадків захворювань в усіх системах організму людини. Дуже важливим залишається питання стану голосового апарату. Серйозною проблемою досі залишається вчасне діагностування патологій голосового апарату, особливо поміж працівників мовних професій та осіб похилого віку.

Демографічне старіння населення за останні десятиліття стало глобальним явищем, характерним не тільки для розвинутих країн, але практично для всього світу. Україна є однією з «найстаріших» країн світу. Згідно з національним демографічним прогнозом, до 2025 року частка осіб віком понад 60 років становитиме 25,0 % від загальної кількості населення, а людей віком 65 років і старше – 18,4 % [1].

У цій ситуації збереження високої соціальної активності осіб похилого віку стає одним з важливих завдань сучасної медицини. Це актуально як для представників професій з активним використанням голосу (вокалісти, актори, педагоги, лікарі та ін.), які нерідко продовжують працювати до дуже похилого віку, так і для непрофесіоналів голосу, яким чистий, звучний, стійкий голос необхідний для продовження соціального життя.

У фоніатрії застосовується ціла низка діагностичних методик - загальноклінічні дослідження, оториноларингологічний огляд, збір анамнезу та скарг хворих, спеціалізований фоніатричний огляд з використанням функціональних навантажень, ларингостробоскопія, мікроларингоскопія, мікроларингостробоскопія, фіброназоларингоскопія, сонографія, рентгенографія гортані, комп'ютерна та магнітно-резонансна томографія гортані, рентгенокінематографія гортані, електроміографія, хронаксиметрія, глотографія, електрокімографія. Серед алгоритмів дослідження голосової функції та визначення стану голосового апарату найбільш важливими є спеціалізований фоніатричний огляд, перцептивна оцінка голосу, вимірювання часу максимальної фонації, спектральний аналіз голосу, оцінка

голосового поля, дослідження оцінок стану голосу з використанням цифрової обробки звуку [2].

З використанням обробки акустичних параметрів успішно визначаються наступні захворювання [2]:

- Гіпотонусна дисфонія
- Гіпертонусна дисфонія
- Гіпогіпертонусна дисфонія
- Фонастенія
- Хронічний катаральний ларингіт
- Крововиливи в голосову складку
- Вазомоторний монохордит
- Вузлики голосових складок
- Хронічний набряково-поліпозний ларингіт

Методи дослідження традиційних акустичних параметрів голосової функції, що застосовуються в широкій практиці, незважаючи на їх діагностичну значущість, не завжди дозволяють повною мірою виявити його розлади в **ранній** стадії захворювання і провести детальну оцінку голосової функції.

Для підвищення ефективності лікування, ранньої діагностики та профілактики захворювань голосового апарату доцільне впровадження у практику лікаря-фоніатра сучасних об'єктивних методів оцінки стану голосової функції гортані, які дозволяють здійснити ранню діагностику порушень голосової функції. Особливого значення набуває об'єктивна оцінка голосу в осіб мовних професій (педагогів, перекладачів, дикторів тощо), оскільки до його якості, як правило, висувуються високі вимоги. Голос має бути виразним, приємним для сприйняття слухачами. Крім того голосовий тракт наведених професій зазнає значних навантажень та потребує особливого догляду.





# РОЗДІЛ 1. ГОЛОСОВИЙ АПАРАТ ТА КОМП'ЮТЕРНИЙ АНАЛІЗ ЙОГО СТАНУ

## 1.1 Анатомія та фізіологія голосового апарату, а також його фізичні та комп'ютерні моделі

### *Анатомія артикуляційного апарату*

Людський голос складається зі звуків, які видає людина за допомогою голосового тракту, включаючи розмову, спів, сміх, плач, крик. Мовний апарат складається з двох частин: центрального (або регулюючого) мовного апарату та периферичного (або виконавчого). Центральний мовний апарат складається з кори головного мозку (переважно лівої півкулі), підкіркових вузлів, що проводять шляхів, ядер стовбура (насамперед довгастого мозку) та нервів, що йдуть до дихальних, голосових та артикуляторних м'язів. Мовленнєва артикуляція здійснюється в результаті діяльності складної сенсо-моторної системи, кінцевим підсумком якої є приведення в готовність оральної, фонаторної та дихальної мускулатури до виконання артикуляторної задачі відповідно до інтегрованої в мозку програми.

Артикуляційний апарат – це система органів, які задіяні у процесі мовоутворення. До них відносяться: ротова порожнина (мова, м'яке та тверде піднебіння, зуби верхньої та нижньої щелепи, губи, щоки), глотка та гортань (рис.1.1.1).

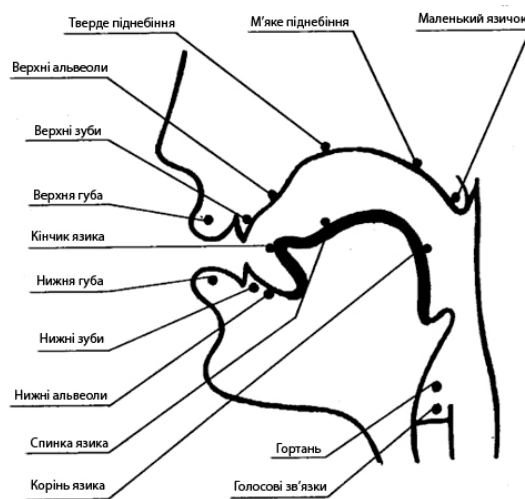


Рис.1.1.1 Будова артикуляційного апарату людини [3]

Наведені органи поділяються на рухливі та нерухомі, активні та пасивні. До рухливих відносяться язик, губи, м'яке небо з маленьким язичком, нижня щелепа та голосові зв'язки. Нерухомі: верхня щелепа, зуби, альвеолярна дуга та тверде небо.

У дорослого чоловіка загальна довжина голосового тракту становить приблизно 17 см. Площа поперечного перерізу голосового тракту, яка визначається положенням язика, губ, щелеп та піднебінної фіранки, може змінюватися від нуля (тракт повністю перекритий) до приблизно 20 см<sup>2</sup>. Носова порожнина починається у піднебінній фіранки і закінчується ніздрями. При опущеній піднебінній завісці носова порожнина акустично з'єднана з голосовим трактом і бере участь в утворенні носових звуків мови.

### *Цифрова модель голосового апарату*

На рис. 1.1.2 зображена типова схема цифрового мовоутворення.

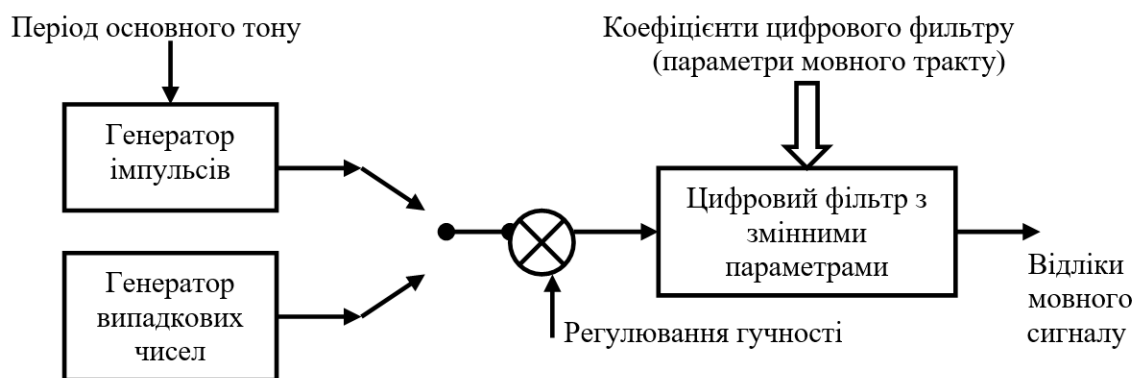


Рис.1.1.2 Цифрова модель утворення мови за Шафнером [4]

Фільтр тут застосовується як система, що моделює резонансні властивості порожнини рота і носа людини. Алгоритм цифрової фільтрації можна реалізувати по-різному. Наприклад, можна для кожного відрізка часу заново розраховувати коефіцієнти рекурсивного фільтра, кількість і розташування полюсів якого визначається кількістю і розташування формант в синтезованому звуці.

Інший підхід - використовувати гребінку вузькосмугових фільтрів з ЧХ, що перекриваються таким чином, що АЧХ результуючої системи практично рівномірна

в смузі частот мовного сигналу. Очевидно, варіюючи коефіцієнтом ослаблення сигналу в кожному з окремих фільтрів гребінки, можна синтезувати сумарну ЧХ практично будь-якої форми. Незавжди бачити, що такий підхід повинен забезпечувати, при правильній організації обчислень, результати синтезу, практично ідентичні для першого підходу. Тим часом, другий підхід значно ефективніше в плані обчислювальних витрат. Гребінку фільтрів при цьому можна організувати або за допомогою класичних методів розрахунку цифрових фільтрів, або за допомогою алгоритмів ДПФ і ШПФ.

Мовний сигнал можна представити у якості послідовності стаціонарних сегментів, довжина яких наближена до довжини відповідних мовних звуків. Відповідно кожному окремому сегменту ставиться у відповідність свій цифровий фільтр. Отримані сегменти відповідно «склеюються». Протяжність сегментів для синтезу мови обираються співрозмірно з найкоротшими звуками мови, що відповідає 10-20 мс.

### *Основні характеристики голосу*

Основними фізичними характеристиками голосу є сила (інтенсивність) і висота (частота). Сила голосу – це суб'єктивне сприйняття інтенсивності чи рівня звукового тиску голоса. Вона знаходиться у прямій залежності від величини повітряного тиску в підскладковому просторі, ступеня напруги голосових складок та визначається амплітудою їх коливань. Сила голосу вимірюється в децибелах (дБ). Висота голосу – суб'єктивне слухове сприйняття частоти ущільнення та розрідження повітря, що утворюються змиканням та розмиканням голосових складок. Висота голосу визначається герцах (Гц).

Голос людини, як і всі інші звуки, що нас оточують, має складну будову і складається з комбінації простих звуків, серед яких виділяють основний тон та часткові тони, іншими словами обертони.

Також у голосі крім висоти, сили звуку необхідно оцінювати тембр голосу і діапазон голосу. Тембр голосу – сума та поєднання обертонів, зумовлених

резонансними властивостями замкнутих об'ємів окремих ділянок голосового тракту, наявність яких обумовлює індивідуальне «забарвлення» голосу, надає промові людини різні емоційні відтінки, дозволяє впізнавати людей за голосом. Ці властивості голосу спеціалісти описують як голос дзвінкий чи глухий, такий, що ллється чи жорсткий, металевий чи матовий тощо. Все це нескінченне розмаїття з погляду акустики є наслідком зміни у часі лише трьох найважливіших акустичних параметрів звуку: частоти коливань, їх амплітуди та звукового спектра. Тембр голосу залежить не лише від спектрального складу звуку, а й вібрато, що надає голосу виразності.

Вібрато – складне акустичне явище, що залежить від періодичної зміни всіх характеристик звуку: висоти та сили. При частоті коливань 5-7 разів на секунду голос красивий, що ллється. При більш рідкісних коливаннях голос сприймається як такий, що хитається, при більш частих - як тремолуючий.

Діапазон голосу – це кількість тонів, в межах мінімальної та максимальної частоти, яка може бути відтворена голосом здорової людини без особливих зусиль. На рис. 1.1.3 голосовий тракт та носова порожнина показані у вигляді труб зі змінними площинами поперечного перерізу.

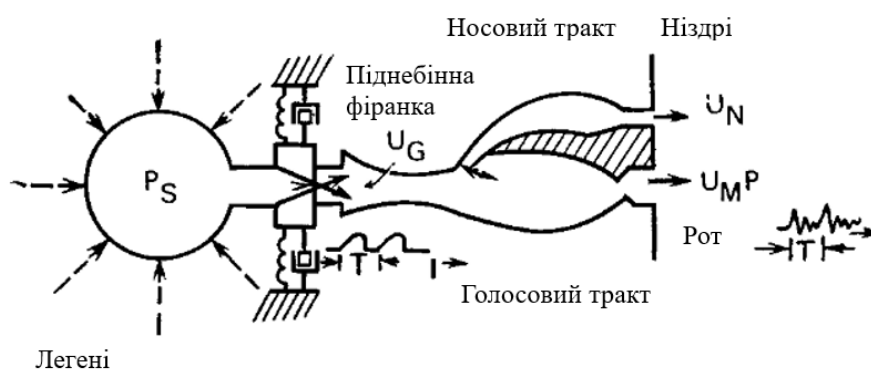


Рис.1.1.3 Схема мовоутворюючих органів [4]

При проходженні звукових хвиль через ці труби їх частотний спектр змінюється відповідно до резонансних властивостей цих труб (рис. 1.1.4). При опису мовоутворення резонансні частоти труби мовного тракту називають формантними частотами. Форманти залежать від конфігурації та розмірів порожнин мовного

тракту і визначаються площею поперечного перерізу. Деякий зсув резонансних частот виникає за рахунок втрат. В таблиці 1.1.1 наведено формантні області деяких голосних фонем:

Таблиця 1.1.1

Символ по системі IPA	Типове слово	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>
i	(beet)	270	2290	3010
ɪ	(bit)	390	1990	2550
ε	(bet)	530	1840	2480
æ	(bat)	660	1720	2410
ʌ	(but)	520	1190	2390
a	(hot)	730	1090	2440
ɔ	(bought)	570	840	2410
u	(food)	440	1020	2240
μ	(boot)	300	870	2240
ɝ	(bird)	490	1350	1690

У більшості спектр голосних фонем характеризується наявністю двох основних формантних областей: в області низьких частот (300-600 Гц) та області високих частот (2500-3000 Гц). Окремі дослідження подають 3 формантні області. Перша форманта, яка надає тембру голосу м'якість, округлість та масивність, була названа низькою формантою (НФ). Друга, що зумовлює дзвінкість, польотність, яскравість голосу, отримала назву верхньої чи високої форманти (ВФ).

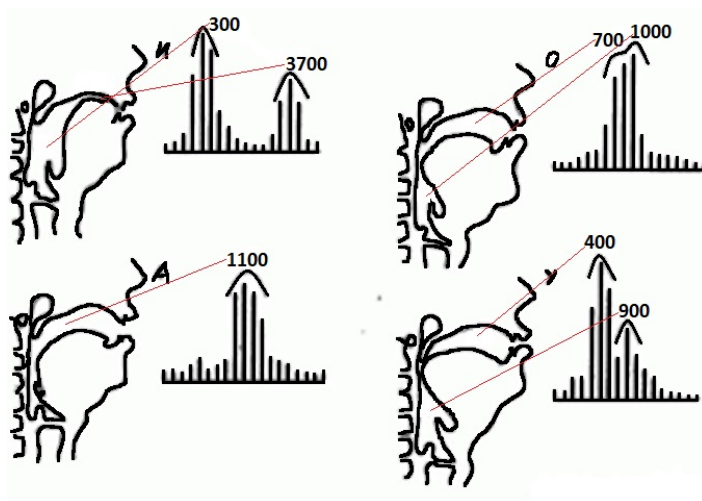


Рис.1.1.4 Положення артикуляторних органів при вимові голосних і, а, о, у за результатами рентгенограм і спектри цих голосних [5]

Дослідження спектрального складу голосу показали важливу роль як високих, так і низьких спектральних його складових. Для низьких голосів (бас, баритон) висока форманта становить 2500-2800 Гц, для вищих (тенора, сопрано) вона сягає 3200 Гц. У професіоналів голосу з великим стажем роботи в області ВФ зосереджено до 30-35% всієї звукової енергії голосу. У початківців вона сягає лише 5-7%.

Встановлено, що вирішальне значення в утворенні основного тону, і особливість основного звучання всього голосу в цілому має характер змикання голосових складок.

Основний тон видозмінюється за рахунок резонансу у верхніх (гортань, ротоглотка, придаткові пазухи носа) та нижніх (трахея, грудна клітка) резонаторах. Формування формантів залежить від клініко-функціонального стану гортані, а остаточне їх звучання від стану резонаторних порожнин. Висока та низька форманти постійно присутні у голосі кваліфікованого працівника голосових професій.

## **1.2. Фізико-математичні моделі голосових зв'язок**

*Базові положення, що використовуються під час моделювання голосових зв'язок[6]*

Терміни «частота основного тону» і «висота тону» часто використовують у однаковому сенсі, проте слід враховувати відмінність цих термінів. Висота тону - це міра суб'єктивна, що оцінюється слуховою системою людини. Частота основного тону, за визначенням, є об'єктивна міра, що оцінюється за допомогою спектроаналізатора. Таким чином, частота основного тону (ЧОТ)  $F_0$  – це точніший науковий термін.

Частотою  $F_0$  можна керувати за допомогою:

- Зміни довжини голосових зв'язок;
- Зміни натягу тканин голосових зв'язок.

*Будова та характеристики тканин*

Для глибокого розуміння механізму управління висотою тону корисно спочатку розглянути будову голосових зв'язок і дати спеціальні назви різних видів тканин, що мають особливі біомеханічні властивості (рис. 1.2.1). Епітелій, що тонким шаром покриває поверхню зв'язок, є пухким і нездатним забезпечити напружений стан зв'язок. Проте глибокі тканини голосових зв'язок – колаген, еластин і м'язи – здатні чинити певний опір натягу і, отже, здатні забезпечити жорсткість голосових зв'язок.

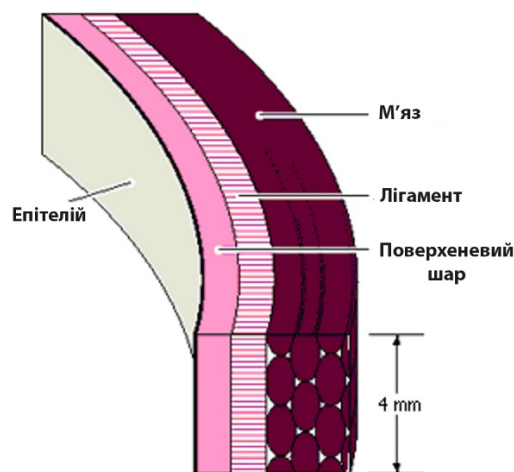


Рис. 1.2.1 Будова тканин голосових зв'язок [6]

При тихій розмові або співі фальцетом вібрає лише еластична поверхня голосових зв'язок - епітелій і жирова тканина, що знаходиться під ним. При розгляді властивостей цього поверхневого шару керування  $F_0$  можна описати за допомогою порівняно простої моделі: поверхневої моделі керування  $F_0$ . Однак при розмові з нормальною гучністю, а також при звичайному співі в процес вібрації залучені глибші шари голосових зв'язок. Через залучення безлічі шарів використовують складнішу модель голосових зв'язок: глибинно-поверхневу модель управління (Body-Cover Model of  $F_0$  Control).

#### *Роль крикотиреоїдної (СТ) та тиреоаритеноїдної (ТА) м'язів*

Первинним параметром, що визначає  $F_0$ , є натяг голосових зв'язок. Аналогічно тому, як гітарна струна або гумова стрічка можуть бути натягнуті для отримання більш високого тону при щипці, збільшення напруги голосових тканин



зв'язок призводить до збільшення  $F_0$ . Звертаючись до основ анатомії голосового апарату, можна сказати, що це забезпечується скороченням або подовженням голосових зв'язок за допомогою крикотиреоїдного та тиреоаритеноїдного м'язів.

Слід пам'ятати:

- ТА м'язи скорочують голосові зв'язки;
- СТ м'язи подовжують голосові зв'язки виштовхуванням щитовидного хряща.

Також важливо розуміти, що ці м'язи можуть працювати незалежно один від одного при керуванні  $F_0$ . В результаті багатьох досліджень активації м'язів вдалося з'ясувати, що диктори та співаки – особливо треновані співаки – намагаються збалансувати активність СТ та ТА м'язів, що дозволяє максимально використати їх потенціал. За допомогою тренування можна допомогти вокалістам краще використовувати можливості голосових м'язів.

Дослідники також виявили, що підвищення  $F_0$  зазвичай має місце зі збільшенням активності ТА у той час, коли активність СТ не є максимальною. На рис. 1.2.2 показано, як м'язи СТ та ТА рухаються в гортані.

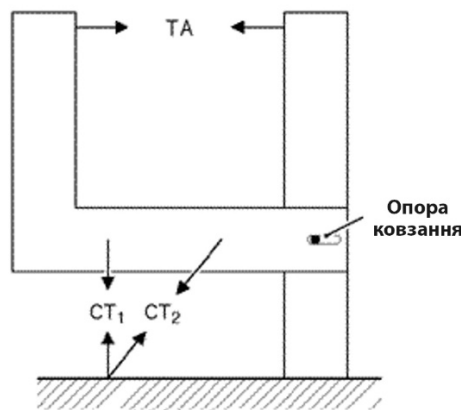


Рис. 1.2.2 Демонстрація руху СТ та ТА м'язів в гортані [6]

Зв'язок між довжиною голосової зв'язки та  $F_0$  нелінійний, що легко перевірити на простій моделі у вигляді гумового джгута. Досвід показує, що з малого збільшенні довжини джгута частота  $F_0$  не змінюється. І лише при значному

збільшенні довжини джгута відбувається зростання  $F_0$ . Для голосової зв'язки ця залежність описується співвідношенням:

$$F_0 = \frac{1}{2L_m} \sqrt{\frac{\sigma_c}{\rho}}$$

де  $L_m$  - довжина голосової зв'язки;  $\sigma_c$  - натяг голосової зв'язки;  $\rho$  - щільність тканини голосової зв'язки.

Дане співвідношення отримано з використанням найпростішої моделі коливальної системи, що містить лише одну масу  $m$  (рис. 1.2.3):

$$F_0 = \frac{1}{2\pi} \sqrt{\frac{k}{m}}$$

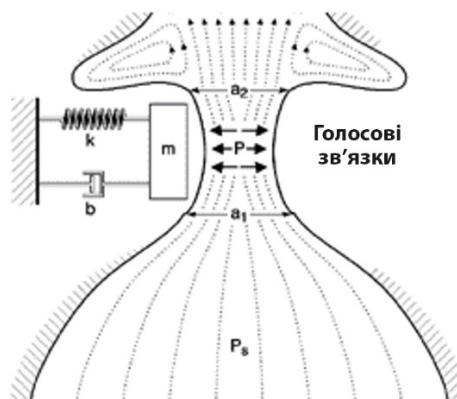


Рис. 1.2.3 Найпростіша модель з використанням однієї маси  $m$  [11]

Вводячи поняття довжини  $L$  зв'язки та її напруженості  $\sigma$ , отримуємо співвідношення, аналогічне відомому у фізиці співвідношенню для частоти коливань натягнутої струни:

$$F_0 = \frac{1}{2L} \sqrt{\frac{\sigma}{\rho}}$$

Зв'язок між жорсткістю  $k$ , напруженістю  $\sigma$ , масою  $m$ , щільністю  $\rho$  та довжиною зв'язки  $L$  визначається співвідношеннями:

$$k = \frac{\pi^2 \sigma m}{L^2 \rho}; \quad m = \rho L T D;$$

де  $T$  та  $D$  - товщина і глибина голосових зв'язок, відповідно.

### *Поверхнева модель управління $F_0$*

Головною концепцією поверхневої моделі і те, що з тихому розмові чи співі фальцетом, коли амплітуда вібрацій мала, вібрує лише поверхню голосових зв'язок. Напруга голосових зв'язок при цьому визначається лише їх довжиною. Дослідники припустили, що голосовий лігамент (область між епітелієм та ТА м'язом) може відігравати головну роль у послабленні напруги.

Кількісно це означає, що подовження м'яза ТА призводить до зниження  $F_0$ . Отже, збільшення активності СТ призведе до зростання  $F_0$ . По суті, дія ТА і СТ позначається зміні довжини м'язи і, як наслідок, її напруженості:

$$\varepsilon = \Delta L_1 / L_0$$

де  $\Delta L_1$  - подовження, спричинене обертанням;  $L_0$  - залишкова довжина;  $\varepsilon$  - напруга голосових зв'язок.

### *Глибинно-поверхнева модель управління $F_0$*

Дана модель складніша за попередню, де дія СТ є вирішальною. Тепер в управлінні грають роль такі параметри: 1) кількість глибинної тканини, що вібрує, і 2) активація м'яза ТА. Зі зростанням розмаху вібрацій, активація ТА значно сильніше позначається на  $F_0$  і, звичайно, якщо ТА не активна, глибинна модель нічим не відрізняється від поверхневої моделі. Це означає, що управління описується співвідношенням:

$$F_0 = \frac{1}{2L_m} \sqrt{\frac{\sigma_p}{\rho}} \left( 1 + \frac{d_a}{d} \cdot \frac{\sigma_{am}}{\sigma_p} a_{TA} \right)^{1/2}$$

де  $d_a/d$  - відношення глибини коливання м'яза ТА до загальної глибини коливання;  $\sigma_{am}$  - максимум активної напруги;  $\sigma_p$  - пасивна напруга.

### *Вплив тиску повітря легень на $F_0$*

Коли повітря виштовхується з легень,  $F_0$  зазвичай підвищується. Цей ефект більш помітний на низьких частотах, коли голосові зв'язки коротші, ніж на високих частотах.

Таким чином, людина керує  $F_0$  за допомогою наступних «важелів»:

- зміна активності м'яза ТА;
- зміна активності м'яза СТ;
- збільшення кількості повітря, що видихається з легенів.

Співаки та диктори зацікавлені в оптимальному співвідношенні цих «важелів» керування  $F_0$ . На рис. 1.2.4 проведено порівняння активності м'язів ТА та СТ звичайної людини та тренованого співака:

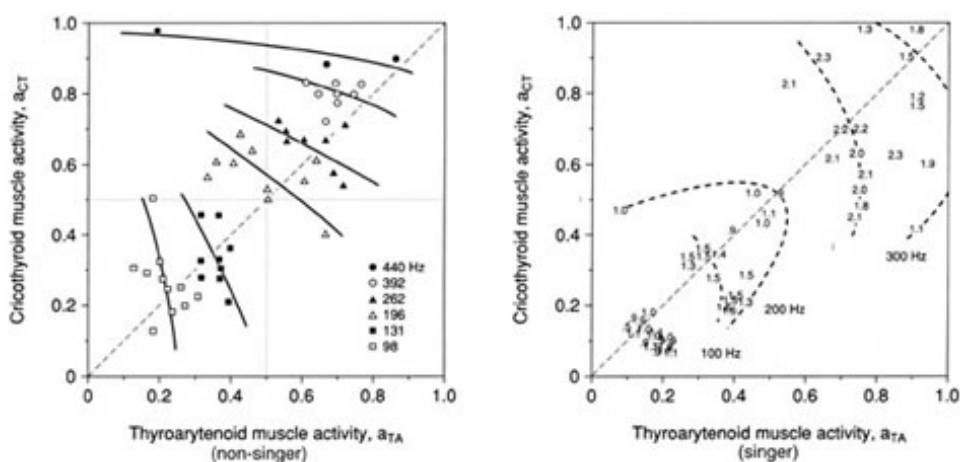


Рис. 1.2.4 Порівняння активності м'язів ТА та СТ звичайної людини та тренованого співака [6]

Цікаво, що графіки дуже схожі. В обох випадках використання ТА та СТ м'язів практично однаково, що видно по діагональній пунктирній лінії, нахиленій під кутом 45 градусів.

### *Міоеластична аеродинамічна теорія фонації*

Міоеластична аеродинамічна теорія фонації (міо-м'язова, еластична – враховується еластичність зв'язок, аеродинамічна – враховується рух повітря з легенів), що виникла у 50-60-ті роки, пояснювала коливання голосових зв'язок наступним чином. Зімкнені голосові зв'язки з підвищенням повітряного тиску

розмикаються, і відбувається одиничний «поштовх» повітря зв'язками. Далі за рахунок еластичності зв'язок вони повертаються на своє місце і стуляються, далі процес циклічно повторюється. Повітряні коливання, що створюються, відбуваються багато разів на секунду, в результаті генерується звукове коливання, яке далі піддається фільтрації голосовим трактом.

Ця модель коливань пояснюється рис. 1.2.3, у лівій частині якого в зоні розташування голосових зв'язок вміщена схема механічної моделі голосових зв'язок у вигляді маси  $m$ , жорсткості  $k$  та активного опору  $b$ .

Пізніші дослідження показали, що дана модель надто спрощена, і коливання, що збуджуються, через активний опір невдовзі мали б припинитися.

На рис. 1.2.5 показана ускладнена модель де додано повітряний стовп голосового тракту. Справа в тому, що для змикання зв'язок тиск повітря має бути зниженим, тим часом легені створюють підвищений тиск. Це зниження тиску і забезпечується існуванням повітряного стовпа з боку ротової порожнини.

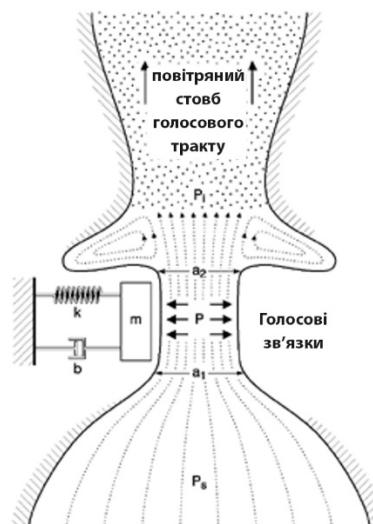


Рис. 1.2.5 Ускладнена модель де додано повітряний стовп голосового тракту [11]

Існують і складніші моделі, одна з яких показана на рис. 1.2.6.

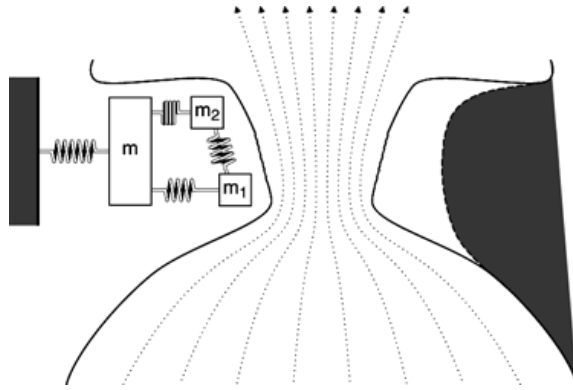


Рис. 1.2.6 Приклад складнішої системи [11]

Це модель із трьома масами. Найбільша маса  $m$  становить ТА м'яз. Дві менші маси  $m_1$  і  $m_2$  становлять поверхневі тканини. Всі маси рухаються незалежно, проте пов'язані з гнучкістю та активними опорами.

Зауважимо, що на даний час досліджуються (з допомогою комп'ютерних технологій) ще складніші моделі, наприклад, з 16 елементарних мас.

#### *Огляд моделей керування висотою тону [8]*

Перша найпростіша фізико-математична модель голосових зв'язок запропонована в роботі Фланагана 1968 [9], який використовував електроакустичні аналогії. Згодом ця модель використовувалася багатьма авторами. Це модель осцилятора з однією масою та втратами, що описується співвідношенням

$$m\ddot{x} + c\dot{x} + kx = \frac{1}{2}ld\left(P_s - 0,935\frac{\rho v^2}{L^2}\right)$$

де  $\dot{x} = \frac{dx}{dt}$ ,  $\ddot{x} = \frac{d^2x}{dt^2}$ , - перша та друга похідні за часом величини, відповідно.

Автоколивання в даній моделі підтримуються завдяки впливу повітряного потоку змінного тиску (рис. 1.2.7), що рухається зі швидкістю  $v$ . Це модель, зручна пояснення основних властивостей голосових зв'язок. Її недолік - неможливість пояснити механізм виникнення коливань.

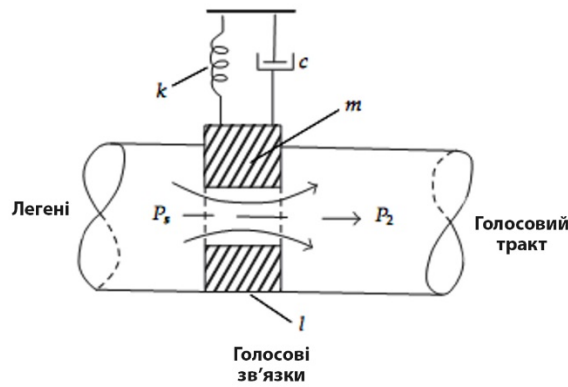


Рис.1.2.7 Повітряний потік змінного тиску [9]

Модель для однієї маси з двома ступенями свободи (маса рухається у двох перпендикулярних напрямках), запропонована у роботі Adachi та Yu [10] (рис. 1.2.8).

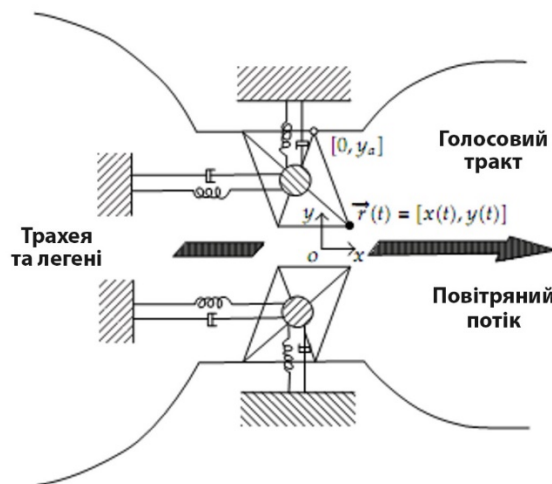


Рис.1.2.8 Модель для однієї маси з двома ступенями свободи [10]

Математична модель такої системи:

$$\frac{1}{2}m\ddot{z} + \frac{1}{2}r\dot{z} = f_{Bz}(t) + f_{\Delta pz}(t) + f_{Cz}(t) + f_{Rz}(t)$$

де  $r = 2\sqrt{mk}/q$  - втрати у системі, що має добротністю  $q$ , а правої частини рівняння зазначені сили, що діють голосові зв'язки.

У роботі [11] Фланагана розглянута модель у вигляді двох мас, що коливаються (рис.1.2.9). Рівняння коливань такої системи має вигляд:

$$m_1\ddot{x}_1 + r_1\dot{x}_1 + k_1x_1 + k_c(x_1 - x_2) = f_1,$$

$$m_2 \ddot{x}_2 + r_2 \dot{x}_2 + k_2 x_2 + k_c (x_2 - x_1) = f_2.$$

У правій частині цієї системи рівнянь вказані сили, що діють голосові зв'язки.

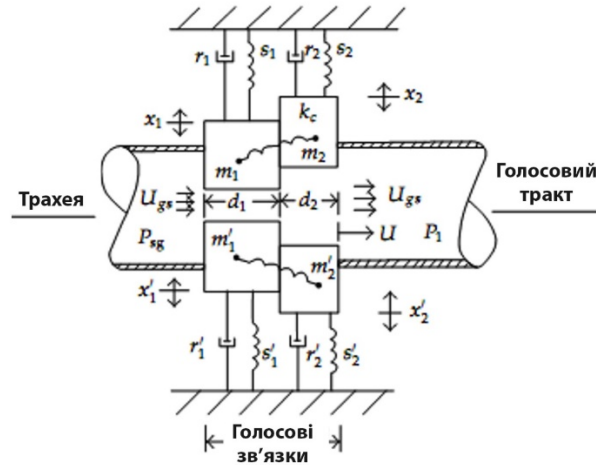


Рис.1.2.9 Модель у вигляді двох мас, що коливаються [11]

Звук, що генерується голосовими зв'язками, є полігармонічним (тобто спектр звуку складається з безлічі дискретних складових), і запропонована відповідна математична модель:

$$m_1 \ddot{x}_1 + r_1 \dot{x}_1 + k_1 x_1 + k_c (x_1 - x_2) + \Theta(-a_1) c_1 \frac{a_1}{2L} = f_1,$$

$$m_2 \ddot{x}_2 + r_2 \dot{x}_2 + k_2 x_2 + k_c (x_2 - x_1) + \Theta(-a_2) c_2 \frac{a_2}{2L} = f_2.$$

Подальші ускладнення моделей полягали не тільки у збільшенні кількості мас, що коливаються, але і в обліку наступних факторів:

- нелінійний характер системи, що генерує коливання;
- асиметричність форми голосових зв'язок;
- нестационарний характер системи, що генерує коливання.

Незважаючи на те, що ускладнення моделей дозволяє все точніше описувати поведінку голосових зв'язок, проте точність такого опису все ще залишається недостатньо високою. У зв'язку з цим у [8] перераховані такі напрями перспективних досліджень у галузі фізико-математичного моделювання голосових зв'язок:



- у нелінійних моделях диференціальні рівняння повинні містити члени як цілого, так і дробового ступенів, а для розв'язання таких рівнянь мають залучатися не лише чисельні методи, а й апроксимативно-аналітичні методи;
- для всіх мас у моделях голосової складки передбачається рух у трьох напрямках: бічне, поздовжнє та вертикальне, що не цілком відповідає дійсному стану речей;
- у моделях голосових зв'язок негативне згасання вважається лінійним; для покращення моделі слід припустити нелінійність загасання цілого або дробового порядку, як того вимагають емпіричні клінічні виміри;
- нестационарний характер коливань голосових зв'язок типовий для багатьох голосових розладів, тому включення змінних параметрів та реактивної сили, що діє за рахунок зміни маси в часі, дозволить врахувати додатковий внесок від впливу нестационарності на коливання зв'язок;
- особливу увагу слід приділити моделі нерегулярності руху голосових зв'язок (обумовленої аномалією або захворюванням), якісному та кількісному аналізу диференціальних рівнянь, що описують цю нерегулярність.

### 1.3. Акустичні дослідження голосу з використанням комп'ютерних технологій

Сучасні акустичні дослідження голосу, засновані на комп'ютерних технологіях, дозволяють швидко та з високим ступенем точності визначити діапазон голосу, силу голосу, частоту його основного тону, кількість обертонів, формантний склад звукового сигналу та характеристики вібрато.

Слід наголосити, що перш ніж проводити комп'ютерний аналіз голосової функції пацієнта необхідно зібрати анамнез захворювання, провести загальноклінічне обстеження, ларингоскопію – метод візуального обстеження гортані, мікроларингоскопію.

Основні можливі помилки експерименту поділяються на помилки запису голосу, помилки вибору апаратури звукозапису та помилки оцінювання спектру сигналу. Відповідно перелічимо за наведеними видами:

1. До проблем з записом відноситься підготовка приміщення та налаштування апаратури. Перше це боротьба з завищеним часом реверберацією в приміщенні. Відповідно кабінет лікаря має бути невеликим та має бути мебльованим. Стіни мають бути виконані з гіпсокартонових панелей, стеля з підвісних акустичних панелей, на підлозі має бути кавролін. Відносно зашумленості бажано ізолювати кабінет від шуму з коридору та встановити табличку «ведеться запис», яка може загорятися тоді, коли лікар починає запис контрольного сигналу.

Відносно налаштування апаратури, необхідно встановити мікрофон на відстані близько 30 см від обличчя пацієнта. Висота мікрофонної стійки має бути на рівні рота. Для запобігання потрапляння потоків повітря на мембрану мікрофону, між пацієнтом та мікрофоном, враховуючи можливі ситуації, коли лікарю треба записати не лише голосні звуки, а й приголосні, або навіть цілі фрази, встановлюється поп-фільтр. Підсилення мікрофону встановлюється таким чином, щоб запобігти нелінійним спотворенням сигналу в тракті підсилювача та тракті аналого-цифрового перетворювача (АЦП). Відповідно підсилення має бути максимальним при якому не відбувається ефект кліпування.

2. Для звукозапису необхідно обрати мікрофон з можливістю максимально передавати реальний сигнал до підсилювача без змін. До таких відноситься клас вимірювальних мікрофонів. Підсилювач та АЦП не має вносити жодних спотворень на сигнал, що записується. Відповідно до теореми Котельникова, АЦП в складі аудіоінтерфейсу має працювати з частотою дискретизації мінімум вдвічі більшою за верхню граничну частоту частотної характеристики сигналу, що записується. На ринку поширені аудіоінтерфейси з частотою дискретизації від 44100 Гц. Позаду мікрофону бажано встановити акустичний екран для запобігання потрапляння відбиттів на тильний бік мікрофону.

Як було наведено вище, мікрофон має бути з класу вимірювальних мікрофонів. Мікрофонний підсилювач не має вносити частотних чи нелінійних спотворень до записаного сигналу. АЦП має бути обрано відповідно з частотою

дискретизації 44.1 кГц. З огляду на все наведене вище, структурна схема експерименту має наступний вигляд:

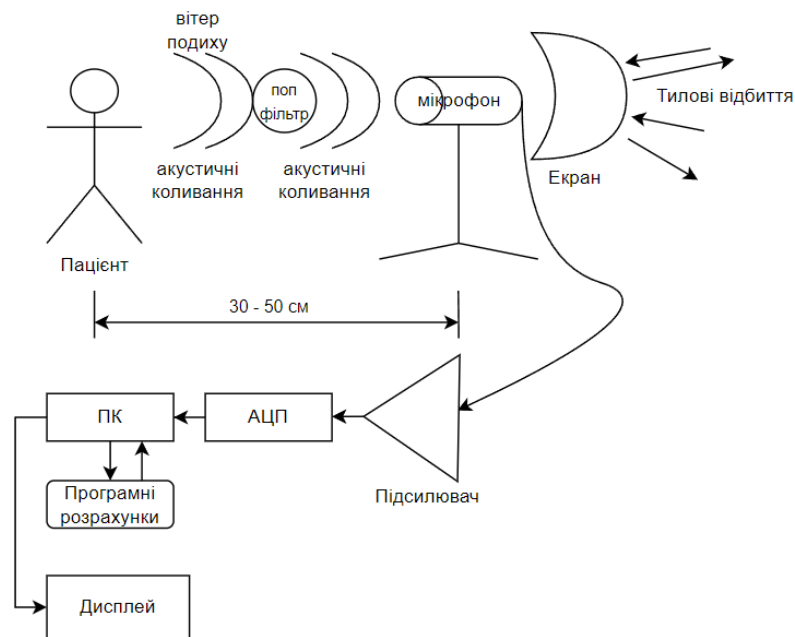


Рис.1.3.1 Структурна схема експерименту

Відповідно до наведених вище невизначеностей та структурної схеми експерименту, маємо наступний план проведення експерименту:

1. Лікар робить первинний огляд пацієнта та робить висновок про потребу проведення тестового запису.
2. Або пацієнт підходить до готового стенду з мікрофоном, або лікар, чи помічник лікаря встановлює стенд з мікрофоном, акустичним екраном, та поп-фільтром.
3. Необхідно провести налаштування висоти мікрофонної стійки та повідомити пацієнту на якій відстані він має стати від поп-фільтру.
4. Проконсультувати пацієнта як потрібно провести запис (максимально довго проспівати набір голосних фонем).
5. Виконати тестовий запис, під час якого налаштувати чутливість підсилювача (максимальна чутливість, при якій не виникає ефекту кліпування).
6. Виконати запис.
7. Подати запис на обробку та отримати результат на екрані.

Пацієнт, після 5-хвилинного голосового навантаження (читання), необхідного для «розігріву» голосу, інтонує всі ноти свого передбачуваного робочого діапазону від нижніх до верхніх на “forte” і “piano” з інтервалом у терцію. Кожна нота визначається на синтезаторі.

Результати аналізу можуть бути представлені у кількох видах.

2.1. У вигляді графіка, що відображає в режимі реального часу весь записаний діапазон голосу пацієнта, за яким можна буде судити про рівномірність наростання та загасання звукового сигналу, оцінити вид «атаки звуку», оцінити стійкість голосних звуків. По осі абсцис фіксується реальний час фонації за секунди, по осі ординат — сила перетвореного сигналу в мілівольтах.

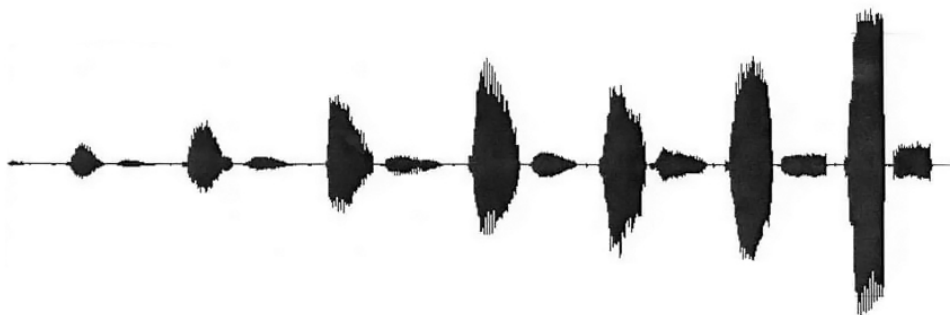


Рис.1.3.2 Загальний вигляд графічно записаного голосу вокалісту в нормі [12]

2.2. У вигляді спектральної діаграми (спектрограми), що змінюється в реальному часі, де по осі абсцис представляється частота звуку (у герцах), а по осі ординат - відповідна їй інтенсивність звуку (у децибелах).

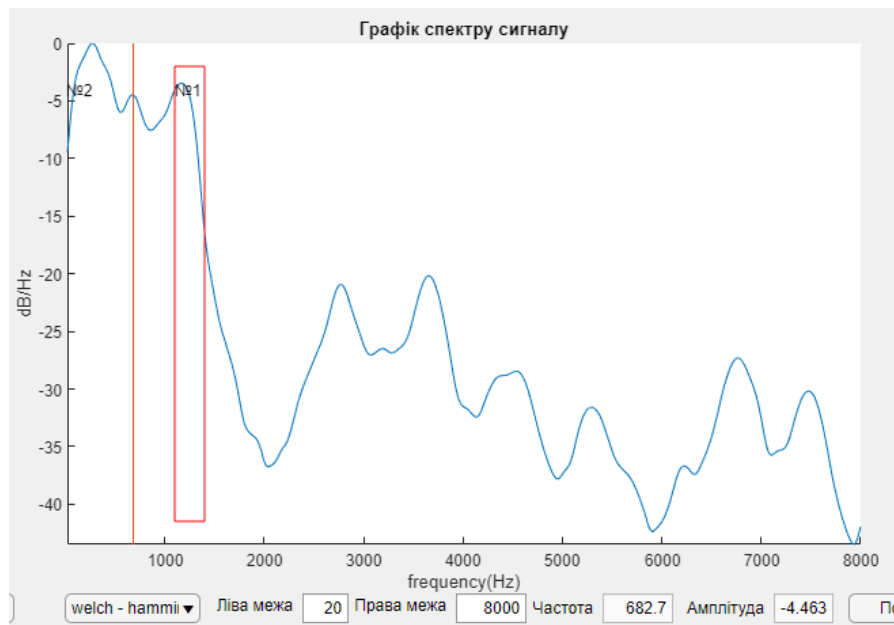


Рис.1.3.3 Спектральна діаграма голосу

Велику діагностичну цінність спектральний комп'ютерний аналіз має у визначенні типу голосу у пацієнтів, які висувають до голосу дуже високі вимоги, наприклад, у вокалістів.

Встановлено, що нерідко захворювання голосового апарату у співаків викликається не лише використанням неправильної вокальної техніки, а й виконанням партій, які не відповідають його типу голосу.

Тип голосу (сопрано, мецо-сопрано, баритон тощо) визначається його діапазоном, тобто. можливістю вокаліста співати у певній ділянці нотного стану (так званої тессітури). Також, для кожного різновиду вокального голосу існують певні перехідні ноти (ноти на яких відбувається перебудова механізму звукознавства з одного регістру на інший, наприклад з головного на грудній) і наявністю в його голосі певних найбільш зручних для виконання тонів, які називаються марними.

Найбільш відповідні типу голосу співака спектрограми характеризуються чітко вираженим сигналом основного тону, регулярністю піків обертонів, однаковими часовими інтервалами між ними та чітко вираженими областями резонансного посилення. У тессітурі, що менш відповідає типу голосу вокаліста,

фіксується зміщення частоти основного тону у бік його підвищення або зниження, нерегулярність піків обертонів і «розмитість» областей резонансного посилення.

#### 1.4. Характеристика голосової функції за патології голосового апарату за даними комп'ютерного спектрального аналізу голосу.

Вікові зміни, що відбуваються в гортані та голосі, пов'язані з морфологічними, ендокринними, біохімічними факторами, змінами діяльності вищої нервової системи та нервово-м'язового апарату. Захворювання гортані ускладнює мовленнєве спілкування, знижують працездатність та створюють загрозу професійній діяльності.

Отже, справедливо говорити, що зменшення звучності голосу чи його повну відсутність прийнято розглядати як ознака захворювання гортані. Але при проведенні найретельнішого обстеження в осіб із зміненим голосом не виявляється органічних змін у гортані. Тоді варто вважати, що захворювання голосового апарату має функціональний характер.

Функціональними дисфоніями вважають такі стани, при яких під час огляду гортані голосові складки та інші її елементи мають нормальну конфігурацію, слизова оболонка гортані звичайного кольору, відсутні порушення артикуляції, дихання[13].

Вирізняють такі основні причини розвитку функціональних дисфоній: патологічні зміни у м'язах гортані; патологічні зміни в органах та системах, що беруть участь у голосоутворенні; патологічна зміна в органах та системах, що не входять до голосового апарату.

Інакше кажучи, дисфонії функціонального характеру спостерігаються у осіб, які страждають на різні соматичні захворювання і користуються неправильною технікою голосознавства.

Також нерідко розлад голосу виникає після стресу (переляк, страх, різні невротичні стани та ін.).

У пацієнтів старше 60 років виникають дегенеративні зміни у грудній клітці та хребті, що обмежують їх рухливість, знижується тонус міжреберної мускулатури. Це призводить до зменшення ємності легень, укорочення фонаційного видиху. Зниження еластичності тканин, зниження загального м'язового тону (зокрема зниження тону лицьової та жувальної мускулатури, м'язів глотки) зумовлюють гіпотонус голосових складок, порушення резонаторної функції глотки, ускладнення артикуляції, що викликає додаткове навантаження на голосовий апарат та сприяє розвитку дисфоній. У літньому віці за рахунок кальцифікації та осифікації хрящів відбуваються обмеження рухливості суглобів (у тому числі і в гортані), що також погіршує якість голосу[13].

Після 50 років відбувається зміна частоти основного тону (підвищення на 30-35 Гц у чоловіків та зниження на 10-15 Гц у жінок), виникають труднощі при підтримці тонової висоти голосу, знижується його сила, звужується динамічний діапазон, нижні та верхні межі якого зміщуються вниз. Змінюються також якісні характеристики голосу. З'являються додаткові призвуки (захриплість, диплофонія), голос стає нестабільним (тремоляція, хитання) [13].

В даний час прийнято ділити функціональні дисфонії на гіпокінетичні (гіпотонусні), гіперкінетичні (гіпертонусні), гіпогіпертонусні, спастичні. Виділяють також функціональну афонію та фонастенію.

Найбільш поширеними порушеннями голосової функції у пацієнтів похилого віку є гіпотонусна дисфонія, гіпотонус голосових складок у поєднанні з вестибулярно-складковою фонацією (гіпо-гіпертонус гортані), рідше зустрічається гіпертонусна дисфонія. З органічних захворювань, що впливають на якість голосу у даної категорії пацієнтів, найчастіше діагностуються хронічний катаральний ларингіт та хронічний набряково-поліпозний ларингіт[13].

Спектральний аналіз голосу використовувався при діагностиці різних захворювань голосового апарату у поєднанні з традиційними методами обстеження,

що дозволяло уточнити ступінь тяжкості порушення голосової функції за різних видів патології та призначити адекватне лікування.

Також, при тривалому, рецидивному перебігу функціональних дисфоній, дослідження спектрального складу голосу, зокрема аналіз стійкості основного тону і тривалості фонації різних голосних звуків, допомагає дати прогноз перебігу захворювання, оцінити ступінь ймовірності переходу функціонального захворювання в органічне.

Найбільш інформативними для оцінки голосової функції є спектрограми, на яких можна в режимі реального часу оцінити частоту основного тону, кількість обертонів та їх регулярність.

### *Гіпотонусна дисфонія*

Велику групу функціональних розладів голосового апарату становлять гіпотонусні порушення. Гіпотонусна дисфонія відноситься до функціональних розладів голосу, що найбільш часто зустрічаються. Питома вага становить приблизно 70-80% від загальної кількості дисфоній. В основному гіпотонусною дисфонією страждають особи жіночої статі, це можна пов'язати з меншою витривалістю у них голосового апарату, лабільністю нервової системи, схильністю до вегетосудинної дистонії та особливостями ендокринної системи, які створюють сприятливе тло для виникнення порушень голосу. Також можна відзначити, що більшою мірою до цього захворювання схильні професіонали: артисти, вокалісти, викладачі, а також представники інших професій, де мається на увазі щоденне голосове навантаження. У таких хворих відзначається різка стомлюваність голосу після голосового навантаження, голос стає хрипким або сиплим, з'являється напруга та біль у м'язах шії, потилиця. Функціональні зміни при гіпотонусній дисфонії найчастіше мають оборотний характер[13].

Дефекти голосу можуть бути виражені від легкої осиплості до різкої захриплості. Гіпотонусна дисфонія проявляється швидкою стомлюваністю голосу, захриплістю різного ступеня після голосового навантаження. Знижується сила



голосу, зникає його польотність, виникають труднощі під час розмови, у пацієнтів може бути відчуття скупчення слизу, «кома у горлі», іноді виникає кашель. При розмові чи співі виникає відчуття «відпливу» повітря, з'являються додаткові призвуки.

При проведенні акустичного аналізу голосу відзначається [12]:

- зниження сили голосу на 10-15 дБ,
- скорочення часу максимальної фонації (ВМФ) до 8-12 с,
- зниження частоти основного тону (ЧОТ) у жінок та підвищення його у чоловіків на 8-10 Гц,
- фіксується нестабільність ЧОТ за амплітудою у короткостроковому та довгостроковому періодах (Jitter  $1,53 \pm 0,32\%$ , Shimmer  $9,2 \pm 1,1\%$ ),
- збільшуються показники рівня шумових компонентів у спектрі (NHR  $0,16 \pm 0,53\%$ , VTI  $> 0,06$ ).

Гіпотонусна дисфонія є порушенням голосу, викликане зниженням тону голосових складок. Причинами, що ведуть до виникнення гіпотонусної дисфонії, є постійне перенапруження голосового апарату, неправильне користування ним, робота на тлі гострих запальних захворювань верхніх дихальних шляхів (ларингіти, трахеїти). Ця патологія займає 17-18% у структурі захворювань голосового апарату.

На спектрограмах, за рахунок слабкості голосових м'язів, відзначається усунення частоти основного тону, як правило, у бік його зниження, нерегулярність піків обертонів і зменшення їх кількості, нестійкість голосних звуків, зміна частоти вібрато (як у бік його збільшення - тремоляція, так і в сторону зменшення - "хитання" голосу). Має місце падіння інтенсивності голосу області окремих тонів перехідного регістру, звуження динамічного діапазону [12].

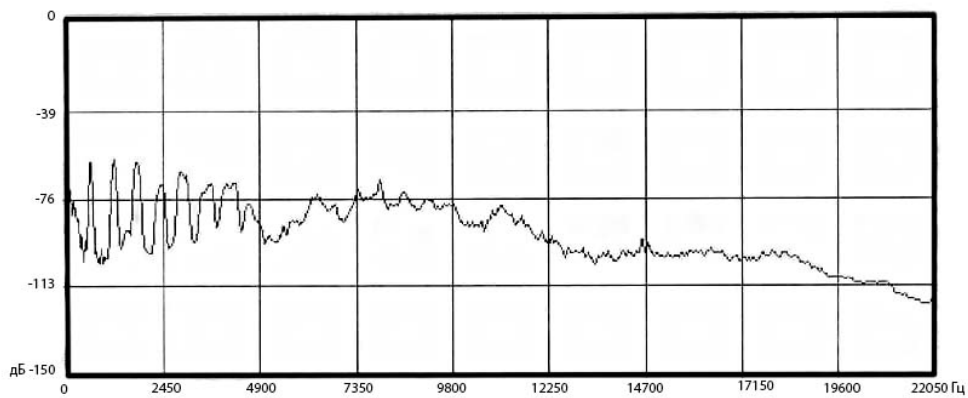


Рис.1.4.1 Типовий вигляд спектру голосної фонемі хворого з гіпотонусною дисфонією [12]

### *Гіпертонусна дисфонія*

Наступною формою функціональних порушень голосу є гіпертонусна. Дана форма дисфонії є розладом голосу, викликане підвищеним м'язовим тонусом внутрішніх м'язів гортані. Причинами, що ведуть до виникнення гіперкінетичної дисфонії, є постійна гучна розмова, особливо за наявності виробничого шуму; використання твердої атаки голосу при зниженні слуху. Подібні порушення голосу найчастіше спостерігаються у чоловіків при напруженій, силовій манері розмови або співу. Пацієнти скаржаться на захриплість, неприємні відчуття в ділянці ший; частий кашель із бажанням відкашляти мокротиння; швидку голосову стомлюваність. Голос звучить напружено, пронизливо, різко.

Причинами, що ведуть до виникнення гіпертонусної дисфонії, часто є постійне перенапруження голосу на роботі (особливо за наявності виробничого шуму) або в побуті (зазвичай за наявності в сім'ї туговухих або багатодітних сім'ях). Хронічному перенапруженню м'язів гортані сприяють зловживання жорсткою атакою та розмова на підвищеній гучності. Іноді гіпертонусну дисфонію називають «хворобою командирів», оскільки досить часто пацієнти з цією патологією у минулому обіймали командні посади у армії чи були керівниками великих колективів. Розвитку гіпертонусного розладу може сприяти наявність хронічного запального процесу в гортані, оскільки досягнення чистого звучного голосу пацієнти змушені постійно перенапружувати голосовий апарат. Найчастіше (80% випадків) гіпертонусна дисфонія діагностується у чоловіків. Голос у цих хворих нерідко з

металевим тембром, пронизливий, різкий, іноді хрипкий та грубий, відзначається жорстка атака звуків.

При проведенні акустичного аналізу голосу при гіпертонусній дисфонії характерними ознаками є[12]:

- підвищення індексу м'якої фонації (SPI) у 1,5-2 рази – до  $27,6 \pm 2,7\%$ ,
- збільшення показників Jitter -  $1,56 \pm 0,23\%$ , Shimmer  $7,95 \pm 2,31\%$ ,
- також  $F_0 - 1.35 \pm 0.34\%$  та  $VTI > 0,07$  що свідчить про значне перенапруження м'язів гортані.

При проведенні спектрального дослідження голосу відзначалося різке усунення частоти основного тону (убік її підвищення), його нестійкість. Сила голосу у пацієнтів не перевищувала 70-75 дБ. Також фіксується вкорочення часових інтервалів між піками обертонів та зменшення числа обертонів, мінімальна кількість областей резонансного посилення голосу та аритмічне, нестабільне вібрато, з тенденцією до збільшення його частоти (тремоляція). Відзначається зниження динамічного діапазону голосу.

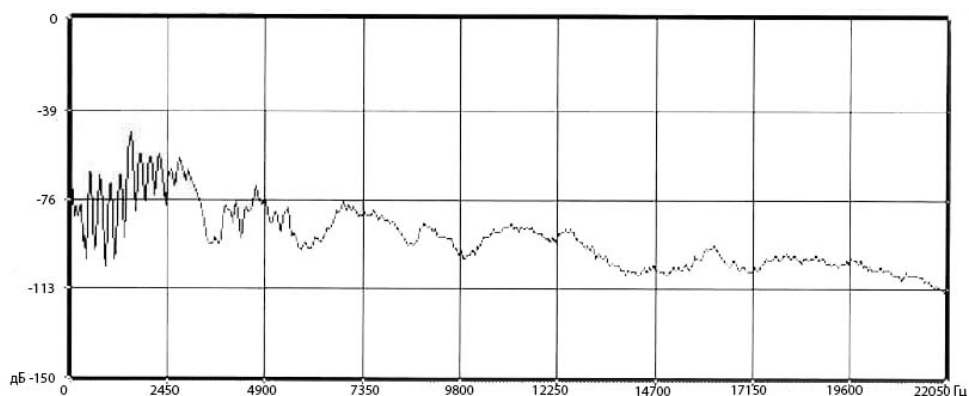


Рис.1.4.2. Типовий вигляд спектру голосної фонемі хворого з гіпертонусною дисфонією [12]

#### *Гіпотонусна дисфонія з хибноскладковим голосом (Гіпо-гіпертонус гортані)*

Гіпотонусна дисфонія з хибноскладковим голосом (гіпо-гіпертонус гортані) є порушенням голосу, що характеризується зниженням тону м'язів голосових складок, і підвищенням тону вестибулярних. Дана патологія, як правило, зустрічається у пацієнтів, які на тлі гіпотонусу голосових складок (гострий ларингіт,

перевантаження голосового апарату), коли для досягнення звучного голосу вони змушені включати у процес голосоутворення вестибулярні складки. Ця патологія зустрічається однаково часто як у чоловіків, так і у жінок. Голос звучить глухо, має затиснутий, здавлений характер [12].

У пацієнтів похилого віку дана патологія розвивається, як правило, як наслідок вже наявного гіпотонусного розладу, коли з метою домогтися звучного голосу пацієнт починає активно використовувати вестибулярні складки. В даному випадку вестибулярно-складкова фонація сприймається як компенсаторний механізм зниженого тону голосових складок. Частіше ця форма дисфонії зустрічається у хворих, які страждають на хронічний ларингіт.

Проведення ендовідеоларингостробоскопії у хворих з гіпо-гіпертонусною дисфонією представляє певні труднощі, які зумовлені з одного боку значним недозмиканням голосових складок і швидким витоком повітря, внаслідок чого голос цих пацієнтів слабкий і швидко виснажується, що є недостатнім для утворення відповідних синхронних світлових спалахів. З іншого боку, не завжди доступні для огляду самі голосові складки, частково або повністю прикриті вестибулярними, що унеможливорює оцінку коливального циклу [13].

Показники акустичного голосу при гіпо-гіпертонусній дисфонії близькі до таких при гіпертонусній дисфонії – індекс м'якої фонації (SPI) значно підвищений, показники нестабільності голосу у короткостроковому та довгостроковому періодах, а також показники співвідношення шуму та гармонік значно підвищені.

При спектральному аналізі голосової функції відзначається звуження діапазону голосу пацієнта, не вираженість основного тону, переважання сигналу обертонів над основним тоном, усунення та зменшення кількості областей резонансного посилення сигналу.

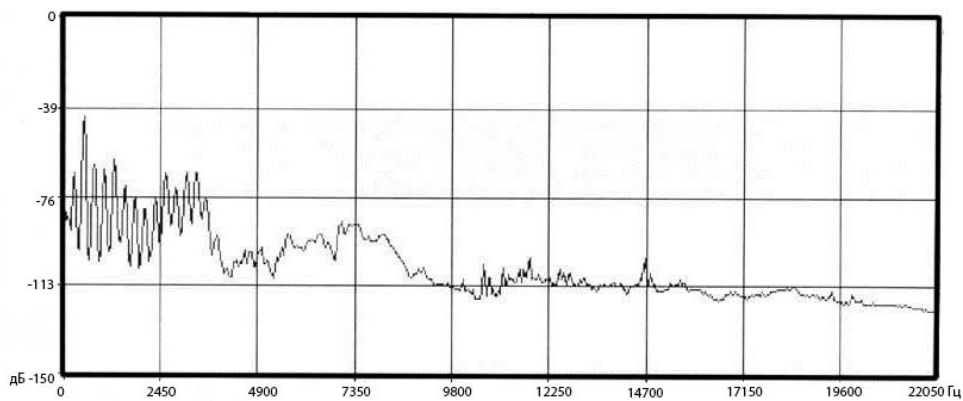


Рис.1.4.3. Типовий вигляд спектру голосної фонемі хворого з гіпо-гіпертонусною дисфонією [12]

### *Фонастенія*

Фонастенія — є функціональним розладом голосу, що характеризується порушенням координації між диханням, фонацією, артикуляцією та функцією резонаторів на тлі неврастенічної схильності. Найчастіше захворювання зустрічається у жінок. Це пояснюється особливостями їхньої нервової та ендокринної системи, що створюють фон для розвитку фонастенії, а також меншою витривалістю, порівняно з чоловіками, голосового апарату [12].

При спектральному аналізі голосу фіксується відсутність сигналу основного тону, стертість піків обертонів та відсутність областей резонансного посилення сигналу. При цьому ларингоскопічна та стробоскопічна картина можуть бути в межах норми.

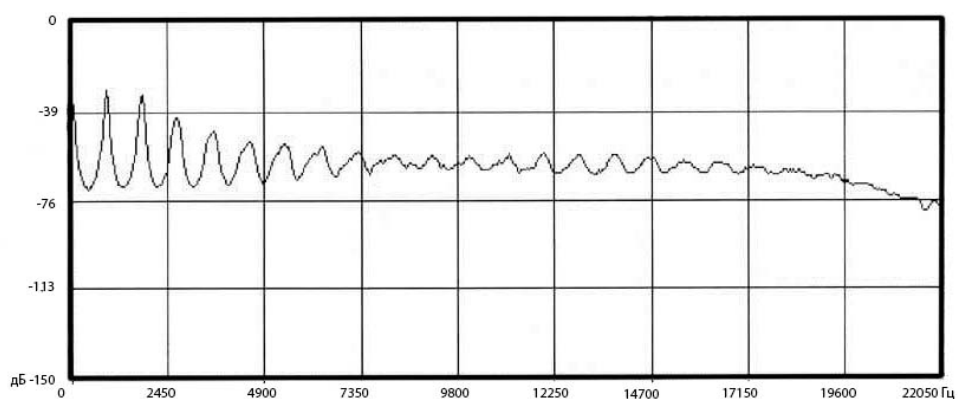


Рис.1.4.4. Типовий вигляд спектру голосної фонемі хворого з фонастенією [12]

### *Хронічний катаральний ларингіт*

Хронічний катаральний ларингіт характеризується наявністю стійкої гіперемії, потовщенням слизової оболонки, пастозністю слизової оболонки голосових складок, вираженою ін'єкцією судин. При фонації нерідко відзначається скупчення в'язкого слизу у вузликів зоні[12].

Хронічний катаральний ларингіт є досить поширеним захворюванням у пацієнтів похилого віку і зустрічається більш ніж у 75% пацієнтів старше 60 років.

Скаргами пацієнтів є, як правило, відчуття сухості та дискомфорту в горлі, незначна захриплість, що посилюється при голосовому навантаженні. З метою домогтися звучного голосу пацієнти перенапружують горло, що провокує розвиток функціональних розладів голосу.

При мікроларингоскопії визначається блідість, субатрофія, витончення та підвищена сухість слизової оболонки гортані та голосових складок. Судинний малюнок часто посилено, проте зберігає лінійну форму. Змінюються еластичні властивості слизової оболонки голосових складок, при фонації відсутня їх щільний контакт, при змиканні визначається лінійна щілина, часто визначається скупчення слизу у вузликів зоні. Ендовідеоларингостробоскопія, як правило, не виявляє значних відхилень від норми.

Акустичні показники голосу неспецифічні та їхня виразність залежить від наявності функціонального порушення голосу. При проведенні акустичного аналізу голосу відзначаються[13]:

- зниження сили голосу на 5-10 дБ,
- час максимальної фонації (ВМФ), як правило, відповідає віковій нормі - 12-14 сек,
- частота основного тону також зазвичай відповідає віковій нормі ( $132 \pm 5,5$  Гц у чоловіків та  $197 \pm 4,5$  Гц у жінок),
- показники  $vF0$ , Jitter, Shimmer, NHR, VTI - зазвичай немає діагностично значущих відхилень.

При спектральному аналізі голосу відзначається невиразність сигналу основного тону, стертість обертонів та відсутність областей резонансного посилення, що свідчить про порушення генераторної функції гортані внаслідок хронічного запального процесу. Відзначалося зниження сили глосу в міру підвищення частоти основного тону області перехідного регістру і крайніх тонів діапазону. Сила голосу не перевищувала 80 дБ. Вібрато характеризувалося нестійкістю, з тенденцією у тремоляції.

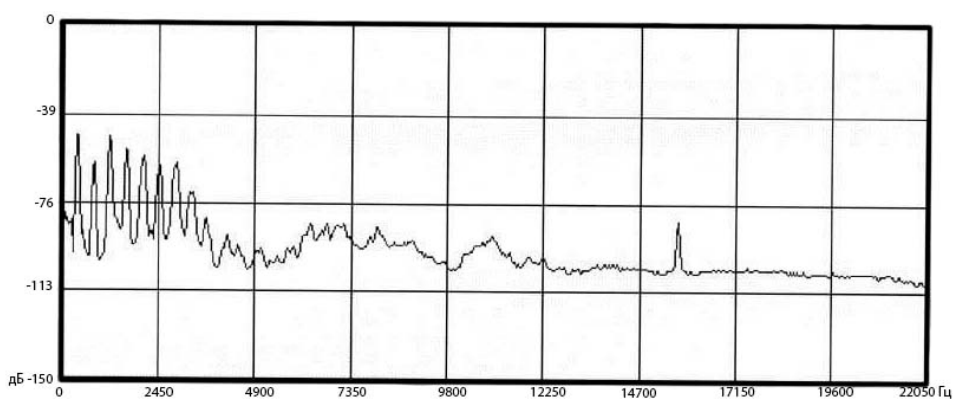


Рис.1.4.5. Типовий вигляд спектру голосної фонемі хворого з хронічним катаральним ларингітом [12]

#### *Крововиливи в голосову складку*

Крововилив у голосову складку виникає, як правило, при форсованому голосовому навантаженні на тлі запальних захворювань верхніх дихальних шляхів (ГРВІ, гострі ларингіти) або під час місячних у жінок на тлі підвищеного артеріального тиску. Слизова оболонка у своїй виглядає набряковою, гіперимированою. Нерідко спостерігаються варикозно змінені судини[12].

На спектрограмі відзначається звуження діапазону голосу, неточне відтворення основного тону, нерівномірні часові інтервали між піками обертонів, звуження областей резонансного посилення голосу. Відзначалася тенденція до зменшення кількості формант, значне зниження сили голосу на всьому діапазоні, що реєструється. Сила голосу не перевищувала 65-70 дБ. Частота вібрата зменшувалася до 3-4 сек.

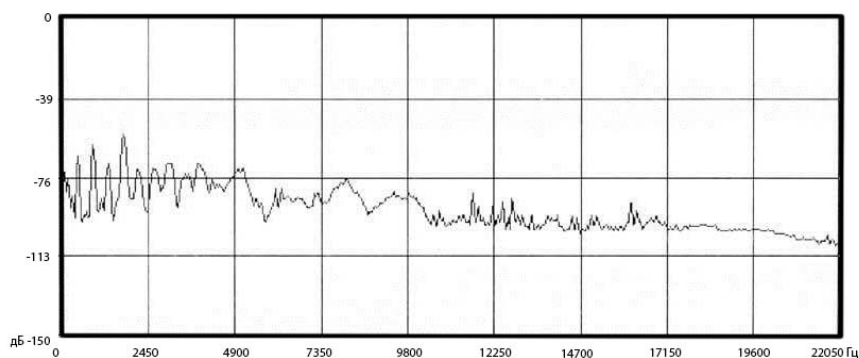


Рис.1.4.6. Типовий вигляд спектру голосної фонемі хворого з кривавиливами в голосову складку [12]

### *Вазомоторний монохордит*

Вазомоторний монохордит характеризується наявністю стійкої запальної реакції на одній із голосових складок, набряком вільного краю слизової оболонки голосової складки, ін'єкцією судин. При фонації часто фіксується лінійна щілина протягом усієї голосової щілини[12].

При спектральному дослідженні голосової функції встановлено звуження голосового поля з допомогою випадання тонів верхнього регістру діапазону. На спектрограмах відзначається нестабільність утримання частоти основного тону, розширення площі піків обертонів, поява додаткових частотних піків, нехарактерних для цього типу голосу. Зменшується сила звуку до 75-80 дБ. Вібрато нестійке та має тенденцію до збільшення частоти (тремоляція).

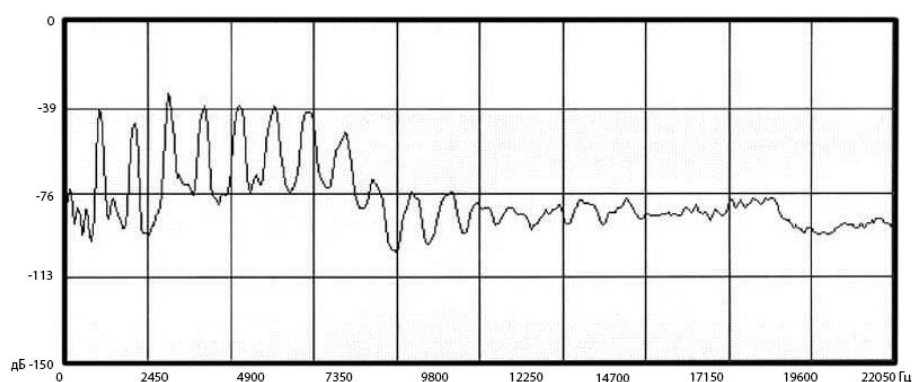


Рис.1.4.7. Типовий вигляд спектру голосної фонемі хворого з вазомоторним монохордитом [12]

### *Вузлики голосових складок*

При спектральному дослідженні голосу спостерігалось зміна його діапазону рахунок зниження рівня високої форманти на “ріано”. Наголошували також на



нестійкості основного тону при фонації на “*riano*”, переважання сигналу обертонів над основним тоном, поява додаткових призвуків, зменшення кількості обертонів та їх регулярності. Кількість співочих формантів не змінювалася[12].

Залежно від ступеня вираженості змін можна визначити, наскільки ефективним буде консервативне лікування, а також коригувати тактику лікування у хворих після видалення вузликів.

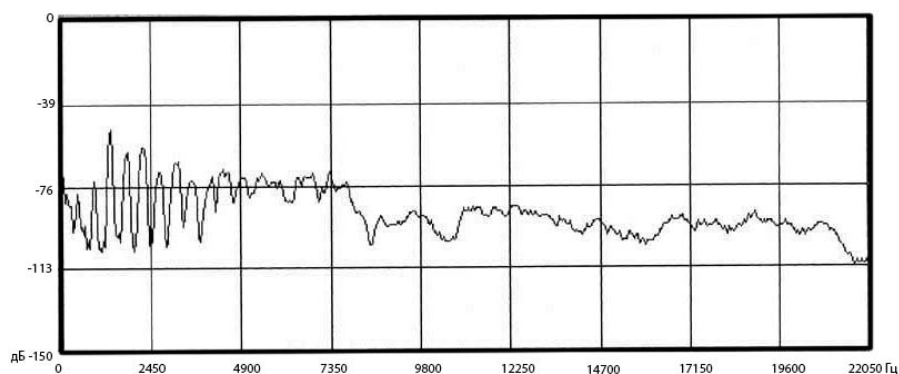


Рис.1.4.8. Типовий вигляд спектру голосної фонемі хворого з вузликами голосових складок [12]

#### *Хронічний набряково-поліпозний ларингіт*[13]

Хронічний набряково-поліпозний ларингіт (ХНПЛ) є поліповидною гіперплазією голосових складок.

Патогенез захворювання пов'язаний із формуванням поліповидного набряку простору Рейнеке, розташованого між епітелієм та еластичною мембраною голосової складки від комісури гортані до головки голосового відростка черпаловидного хряща, тобто займає всю мембранозну частину голосової складки.

Етіологічними факторами захворювання є куріння, голосові навантаження, робота у жарких приміщеннях, професійний контакт із фарбами, розчинниками, деякими газами. Основною скаргою пацієнтів з ХНПЛ є захриплість різного ступеня виразності, огрублення голосу. Деякі пацієнти відзначають появу задишки під час фізичних навантажень.

Розлад голосової функції при цьому захворюванні має як органічну, і функціональну природу. Поліповидний набряк простору Рейнеке, змінюючи структуру голосової складки, порушує її вібраторну функцію. Змикання при фонації

здійснюється поліпозною тканиною, при тривалому перебігу захворювання сама голосова складка стає тоншою, розвивається атрофія її м'яза.

Для підвищення гучності мови ці пацієнти використовують форсовану фонацію, яка супроводжується перенапругою внутрішніх та зовнішніх м'язів гортані. У частини пацієнтів із хронічним набряково-поліпозним ларингітом формується патологічний тип фонації – частково вестибулярно-складковий голос. Виражена вестибулярно-складкова фонація найбільш характерна для пацієнтів з великою давністю захворювання та досить великим обсягом поліпозної тканини.

Пацієнти похилого віку з ХНПЛ найчастіше звертаються до лікаря на пізніх стадіях захворювання, коли зміни голосу стають настільки вираженими, що створюють складності у звичайному соціальному спілкуванні або виникає порушення дихання.

При мікроларингоскопії визначається характерна картина – як правило, двостороннє поліповидне потовщення голосових складок, рідше процес носить односторонній характер. Проведення ендоларингостробоскопії у пацієнтів з ХНПЛ необхідне проведення диференціального діагнозу зі злоякісними новоутвореннями.

Показники акустичного аналізу голосу характеризуються:

- зниженням ЧОТ у чоловіків та жінок на 10-20 Гц,
- збільшенням показників нестабільності ЧОТ у довгостроковому та короткостроковому періоді,
- зміною NTR у бік збільшення шумової компоненти.

Дані показники не є специфічними для ХНПЛ. Застосування акустичного аналізу у літніх пацієнтів є доцільним для оцінки ступеня вираженості функціонального компонента та визначення тактики лікування.

### 1.5 Характеристичні ознаки голосу в діагностичній системі LingWaves

Всі особи з порушеннями голосу не залишаються байдужими до свого стану, помітніше це проявляється при функціональних розладах голосоутворення. У деяких переживання має характер ситуаційно-психологічних реакцій, вираженість яких залежить від ролі голосу в трудовій діяльності, особливостей особистості,

діяльності голосового розладу, багаторазовості малоуспішного лікування в минулому. В роботі [13] були проаналізовані дані пацієнтів з функціональними порушеннями голосу, які звертаються за консультативною допомогою у фоніатричний центр. У дослідженні брали участь 20 осіб із різними функціональними порушеннями голосу, з них 16 жінок (80%) та 4 чоловіків (20%). Здебільшого це особи працездатного віку від 25 до 49 років. Професія кожного пацієнта була пов'язана із підвищеним голосовим навантаженням.

До завдань констатуючого етапу дослідження входило: вивчення порушення голосу учасників дослідження; проведення клінічного логопедичного обстеження з метою виявлення у них ступеня порушення голосу; оцінка стану голосу за допомогою акустичного аналізу голосу; узагальнення одержаних результатів. Під час вивчення медичної документації фіксувалися такі показники: клінічний діагноз; скарги, які пред'являються пацієнтом; обстеження лікарями-фоніатром, неврологом, ендокринологом та ін; оцінка результатів анкетування. Під час дослідження голосової функції використовувалися непряма ларингоскопія, ендовідеоларингостробоскопічний метод обстеження голосового апарату, програма LingWaves для аналізу акустичних параметрів голосу.

LingWAVES стала однією з найбільш використовуваних систем для професійного аналізу голосу та мовлення, біологічного зворотного зв'язку та документування за останні 20 років у всьому світі. Поєднання стандартних і нових технологій аналізу та обробки разом із простим використанням є ключовими характеристиками цієї унікальної системи [14].

Система складається з різних модулів, керованих інтерфейсом користувача на основі lingWAVES. Менеджер клієнтів дозволяє аналізувати та документувати на основі пацієнта/клієнта з перевагою порівняння та відстеження результатів у часі.

Dysphonia Severity Index (DSI) розроблений як об'єктивний і кількісний відповідник якості голосу. Він заснований на багатофакторному аналізі розширеної бази даних з більш ніж 1000 нормальних і патологічних голосів для

багатоцентрового дослідження, ініційованого Бельгійською дослідницькою групою з розладів голосу. На одного пацієнта було зібрано більше 45 голосових характеристик та вимірювань.[15]

Крім того, якість голосу, сприйняту клініцистом і виражену за допомогою ступеня осиплості, була отримана для всіх суб'єктів. Індекс тяжкості дисфонії потім був розроблений за допомогою багатофакторного статистичного підходу, який розраховував комбінацію вагових змінних таким чином, щоб різні типи голосів відповідно до сприйнятої якості голосу по шкалі GRBAS були оптимально визначені.

Шкала GRBAS є слухово-перцептивним методом оцінки захриплості Японського товариства логопедії та фоніатрії, яка дає оцінки 0, 1, 2 або 3 для ступеня захриплості, шорсткості, дихання, астенії та напруженості, де 0 нормальний, 1 – легкий ступінь, 2 – середній ступінь, 3 – високий ступінь. [16]

В результаті отримано формулу, з якою за допомогою лише чотирьох характеристик голосу та вагових коефіцієнтів можна розрахувати індекс DSI. Ці характеристики голосу: максимальний час фонації (MPT, сек.), найвища частота (F0-high, Гц), найнижча інтенсивність (I-low, dB(A)) і jitter (%). Розраховані ваги наведені в такому виразі:

$$DSI = 0.13 \cdot MPT + 0.0053 \cdot F0High + 0.26 \cdot ILow - 1.18 \cdot jitter + 12.4$$

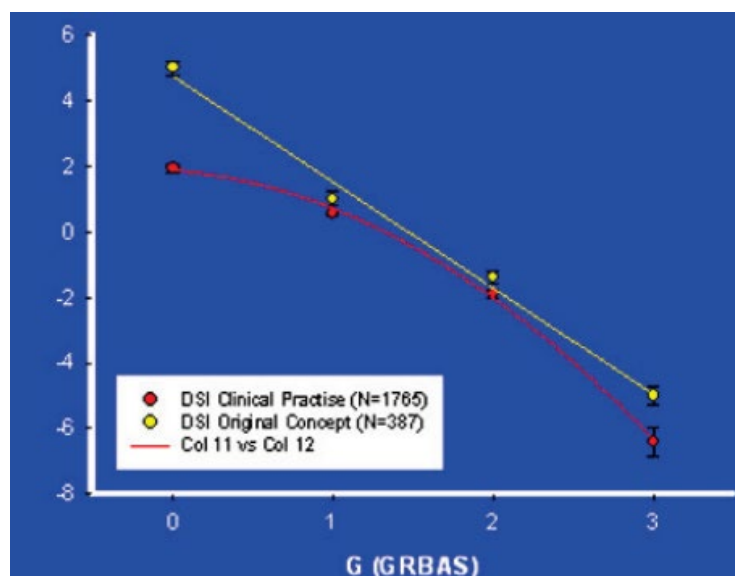


Рис.1.5.1. Демонстрація зв'язку між DSI і параметром "G", як описано в шкалі GRBAS [15]

Жовта лінія показує лінійне відношення, яке було розраховане за отриманою формулою. Червона лінія є результатом аналізу великої клінічної бази даних і показує подібну тенденцію.

На основі аналізу бази даних Бельгійської дослідницької групи з розладів голосу, яка містить понад 1000 випадків було розроблено таблицю відповідності рівня порушення голосу за Therapy Outcome Measures (Таблиця 1.4.1).

Таблиця 1.4.1:

ТОМ		DSI
0	Важка персистуюча афонія: у хворого немає голосу, не може фонувати	< - 4.3
1	Постійна дисфонія: можуть бути спорадичні періоди фонації, афонічні періоди	- 4.2 до -2.3
2	Помірна дисфонія: пацієнт може видавати голос, але є часті періоди дисфонії	-2.2 до -0.4
3	Легка або помірна дисфонія: менш часті періоди приглушеної фонації або незначна персистуюча дисфонія	- 0.3 до 0.7
4	Незначна дисфонія: спорадичні моменти дисфонії протягом коротких періодів	0.8 до 1.7
5	Немає дисфонії	> 1.8

МРТ, максимальний час фонації - це максимальний час (у секундах), протягом якого людина може витримати голосний звук при одному глибокому вдиху з відносно комфортною висотою і гучністю. Оскільки він неінвазивний і не вимагає спеціального обладнання, крім секундоміра, його можна легко використовувати в різних умовах. Існує давня традиція дослідників, які використовують цю міру, і вони використовували різноманітні голосні звуки: найпоширенішими є «ах», «е» та «оо». Як правило, найкраща з трьох спроб підтримати голосну використовується як МРТ пацієнта [17].

З цього ж запису, відповідно до протоколу роботи з пацієнтом з програмою LingWaves, можна отримати такі параметри як Jitter, Shimmer, Irregularity, Noise, Overall severity. Параметр Jitter використовується для отримання індексу вираженості дисфонії [18].

Jitter та shimmer є акустичними характеристиками голосових сигналів, які викликані нерегулярною вібрацією голосових складок. Вони сприймаються як грубість, дихість або хрипість голосу мовця. Уся природна мова містить певний рівень тремтіння та мерехтіння, але їх вимірювання є поширеним способом виявлення голосових патологій. Особисті звички, такі як куріння або вживання алкоголю, можуть підвищити рівень тремтіння і мерехтіння в голосі. Однак багато інших факторів також можуть мати вплив, наприклад, гучність голосу, мова або стать. Оскільки тремтіння і мерехтіння представляють індивідуальні характеристики голосу, які люди можуть використовувати для розпізнавання знайомих голосів, ці заходи можуть бути навіть корисними для систем розпізнавання мовців.

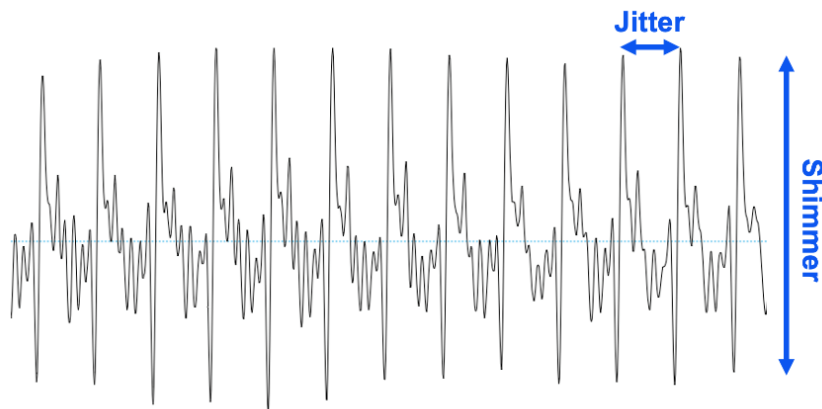


Рис.1.5.2. Демонстрація параметрів Shimmer та Jitter [19]

Щоб виконати вимірювання Jitter на відрізок звукового сигналу потрібно[20]:

- обчислити періоди сигналу;
- обчислити різниці між періодами по модулю;
- обчислити середній період сигналу;
- обчислити середню різницю між періодами.

Jitter - представляє середню абсолютну різницю між двома послідовними періодами:

$$Jitter = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i-1}|$$

Для отримання параметру Jitter у відносному відсотковому вигляді потрібно:

$$Jitter(\%) = \frac{Jitter}{\frac{1}{N} \sum_{i=1}^N T_i} \cdot 100$$

Відповідно, для вимірювання Shimmer на відрізку звукового сигналу необхідно:

- обчислити середню амплітуду сигналу;
- обчислити різниці між амплітудами по модулю;
- обчислити середню різницю між амплітудами.

Shimmer дорівнює середній різниці між амплітудами, поділеній на значення середньої амплітуди:

$$Shimmer(\%) = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - A_{i+1}|}{\frac{1}{N} \sum_{i=1}^N A_i} \cdot 100$$

Наступним тестовим записом пацієнт має зробити наступне: проспівати, піднімаючись від найнижчої висоти до найвищої, яку пацієнт може досягти, використовуючи голосний звук а, потім спускаючись від найвищої до найнижчої висоти, якої пацієнт може досягти, використовуючи звук /а/. Відповідно з запису зі спадним тоном треба взяти максимальну частоту основного тону F0-high, Гц.

Для отримання індексу DSI залишається визначити найнижчу інтенсивність (I-low, dB(A)). Для цього робиться вимірювання шумоміром, коли пацієнт має почати спів фонему /а/ своїм звичайним голосом, а потім перейти до найтихішого голосу, не переходячи на шепіт.

HNR - це оцінка співвідношення між періодичними та неперіодичними компонентами голосового сигналу[20]. Періодичні компоненти виникають

внаслідок вібрації голосових зв'язок, а неперіодичні – це шум, що утворюється в голосових зв'язках. Оцінка між цими двома компонентами відображає ефективність мови, тобто, чим більший потік повітря, що витісняється з легенів, перетворюється в енергію вібрації голосових зв'язок, тим більшим буде HNR. Таким чином, високий показник HNR асоціюють з дзвінким і гармонійним голосом. Визначення співвідношення гармонік до шуму ґрунтується на математичних основах. Щоб визначити HNR використовують автокореляційну функцію голосового сигналу, приклад якої продемонстровано на рис. 1.5.3.

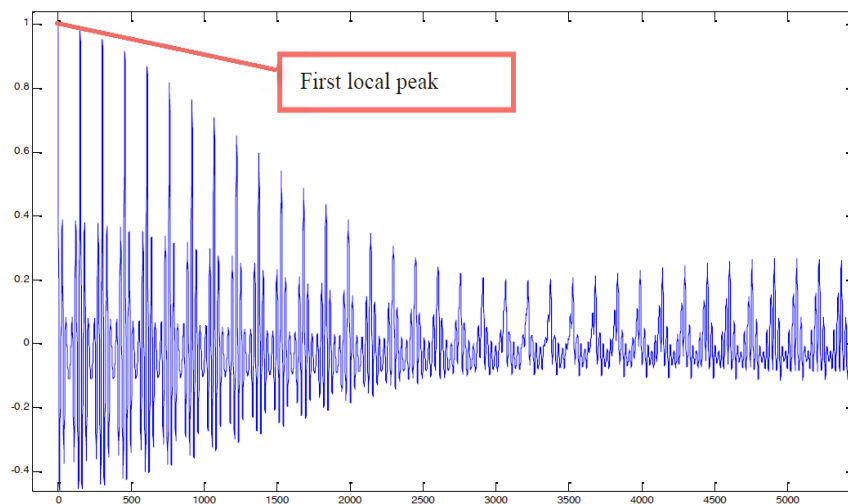


Рис. 1.5.3. Автокореляційна функція.

Значення HNR отримується за наступною формулою:

$$HNR = 10 \lg \frac{AC_v(T)}{AC_v(0) - AC_v(T)}$$

$AC_v(T)$  - значення автокореляційної функції в позиції індексу, що відповідає періоду сигналу. Тому очікувані значення для частоти основного тону визначають положення цього піку між двома індексами. Після визначення індексів локальний максимум виявляється в межах першого та другого індексу, знаходячи їх відповідну амплітуду.

Таким чином, при обстеженні пацієнтів із функціональними порушеннями голосу були отримані наступні результати.



Після проведення ларингостробоскопічного дослідження можна назвати різні форми функціональних порушень гортані:

- При гіпотонусній дисфонії спостерігаємо, що слизова оболонка горла має рожевий колір; обидві половини рівномірно рухливі; вестибулярні складки не змінені; справжні голосові складки каламутно-сірі, поверхня гладка, краї рівні; при фонації недозмикаються, щілина лінійна по всій довжині. Голос осиплий.

- Гіпертонусна дисфонія характеризується тим, що слизова оболонка горла рожевого кольору; обидві половини рівномірно рухливі; вестибулярні складки напружені, прикривають собою голосові складки; справжні голосові складки каламутно-сірі, поверхня гладка, краї рівні; при фонації стуляються по всій довжині. Голос грубий.

- При обстеженні пацієнтів із гіпогіпертонусною дисфонією відзначається: слизова оболонка гортані рожевого кольору; обидві половини рівномірно рухливі; вестибулярні складки напружені; справжні голосові складки каламутно-сірі, поверхня гладка, краї рівні; при фонації недозмикаються, прикриті вестибулярними складками. Голос охриплий, затиснутий.

- У пацієнтів із функціональною афонією було виявлено таку картину: слизова оболонка гортані рожевого кольору; обидві половини рівномірно рухливі; вестибулярні складки не змінені; справжні голосові складки каламутно-сірі, поверхня гладка, краї рівні; при фонації змикання по всій довжині. Звучного голосу немає, каже пошепки.

Після проведення ларингостробоскопії у 20 пацієнтів були виявлені такі форми дисфонії: у 12 пацієнтів (60%) діагностована гіпотонусна дисфонія; у 4 пацієнтів (20%) – гіпертонусна дисфонія; у 3 (15%) – гіпергіпотонусна дисфонія, у одного пацієнта (5%) – функціональна афонія.

Під час проведення акустичного аналізу голосу можна назвати такі результати: при всіх видах функціональних дисфоній відзначається різке скорочення часу максимальної фонації (MPT); негативний DSI; виражена пертурбація частоти

основного тону; значне звуження частотного діапазону та деяке звуження динамічного діапазону.

Таблиця 1.4.2

Група пацієнтів (n=20)							
Нозологія	Irregularity	Noise	Overall Severity	MPT, с	ЧОТ, Гц	Висота, Гц	Сила, Дб
Гіпотонусна	0,86	0,74	0,67	12,4	202,8	334,9	70
Гіпертонусна	1,26	0,98	1,07	10,6	185,6	217,4	65
Гіпергіпотонусна	1,06	1,04	1,07	14,2	223,7	315,7	63

З даних табл. 1.4.2 видно, що у всіх пацієнтів з патологією голосового апарату різко змінені такі акустичні параметри, як Irregularity (відчутна нерівномірність джерела голосу), Noise, Overall Severity (глобальне, інтегроване враження про відхилення голосу), у бік збільшення показників від нормативних. Що ж до індексу вираженості дисфонії (DSI), то для кожної нозологічної групи він наочно представлений у табл. 1.4.3.

Таблиця 1.4.3

Група пацієнтів N=20	
Нозологія	Індекс DSI
Гіпотонусна дисфонія	I степені
Гіпертонусна дисфонія	I-II степені
Гіпогіпертонусна дисфонія	II-III степені

При проведенні анкети-опитувальника у пацієнтів основними питаннями були: стомлюваність голосу після підвищеного мовного та голосового навантаження, зміна тембру голосу до та після стресових ситуацій, захворювання верхніх дихальних шляхів. У процесі анкетування респонденти відзначали зміни голосу. Дані критерії (втома голосу) відзначали 12 обстежуваних (60%). Однією з ознак дисфонії було «першіння» в гортані, бажання відкашлятися. Аналіз анкети-опитувача показав, що ці симптоми були у 15 пацієнтів (75%). У 10 респондентів

(50%) відзначалася сухість у горлі. Ці відчуття посилювалися при інтенсивному мовному та голосовому навантаженні, а також в умовах стресу.

Таким чином, ґрунтуючись на отриманих результатах експериментального дослідження, можна дійти невтішного висновку, що основний контингент пацієнтів становили люди працездатного віку професій з активним використанням голосу; аналіз медичної документації та анкетування показав, що всі пацієнти, які взяли участь у дослідженні, скаржилися на великі голосові навантаження, психоемоційні стресові фактори; клінічне обстеження голосового апарату дало нам розуміння структури дефекту, допомогло виявити не тільки анатомічні зміни в гортані, а й порушення акустичних параметрів голосу, які у всіх пацієнтів значно відрізнялися від нормативних показників унаслідок неправильного голосознавства та незнання основних профілактичних заходів щодо гігієни голосу.

Комплексне обстеження пацієнтів із функціональними порушеннями голосу має велике значення при розробці та виборі тактики методик та реабілітаційних заходів, спрямованих на відновлення голосової функції. Це дуже важливо у повсякденній роботі спеціаліста логопеда-фонопеда, так як одним з найбільш часто використовуваних методів при функціональних дисфонії є фонопедія. Фонопедичні вправи сприяють відновленню звучності голосу та покращенню акустичних параметрів голосу.

Крім людей похилого віку та професіоналів голосу важливо працювати також над питанням голосу дітей. Як і всі інші хвороби, патології голосу потребують завчасного виявлення. Крім того важливим питанням є збереження голосу дітей вокалістів. Для цього в роботі [21] було проаналізовано дані акустичного скрінингу 27 дітей, що займаються співом. З них 15 дітей віком 7-12 років (до мутації), 12 дітей віком 13-15 років (з мутацією).

В групі дітей молодшого віку без мутацій відсутні відмінності між голосами дітей чоловічої та жіночої статі та всіма акустичними параметрами. (Таблиця 1.4.4). У дітей в періоді мутації виражаються статеві відмінності та більшістю акустичних

параметрів голосу. У молодших дітей голос характеризувався невисоким Jitter (нижче 1%), достатньою силою (89 - 97 дБ), широким діапазоном (Fdelta - 352,1 Гц) та позитивним значенням DSI, що не перевищує 3,2. Частота основного тону в межах від 175,1 до 296,2 Гц, середнє значення - 243,8 Гц; MPT -10,4 с.

У періоді мутації у хлопчиків і дівчаток виявлено подовження MPT, в середньому до 12,9 с, але статистичних відмінностей в порівнянні з дітьми 7-12 років не встановлено.

У хлопчиків відзначалося звуження частотного діапазону голосу і зміщення його в бік низьких частот, що проявилось у статистично достовірній зміні параметрів FO, Fmax, Fmin і Fdelta. Тембр голосу відрізнявся нестійким характером і вираженим тремтінням. Показник Jitter часом перевищував 5% (в середньому 1,6%). В цілому, голос у хлопчиків 13-15 років характеризувався низьким, негативним значенням DSI: від -6,3 до -1,7. Середній показник склав -1,6, що, безумовно, свідчить про невисоку якість вокального голосу. У дівчаток в періоді мутації, в порівнянні з дітьми 7-12 років, виявлено розширення частотного діапазону голосу у бік високих і в бік низьких частот. Показник Fmax склав 740,2 Гц, Fmin -138,1 Гц, Fdelta - 520,1 Гц. У дівчат 13-16 років діапазон виявився в 2 рази ширше, ніж у хлопчиків того ж віку. Голос дівчаток відрізнявся незначним тремтінням (Jitter - 0,8%) і кращою, ніж у хлопчиків і у дітей 7-12 років, динамікою гучного співу. Сила голосу при співі на forte склала 92,9 дБ, при співі на piano - 56,9 дБ. Отримані результати свідчать про високу здатність дівчат з мутацією користуватися як гучним, так і тихим голосом, що підтверджувалося параметрами розмовного профілю. У хлопчиків з мутацією і у дітей до мутації ці параметри істотних відмінностей не мали. Максимальна сила голосу при співі в середньому не перевищувала 88,3 дБ, мінімальна - 60,2 дБ. Отже, голос дівчат з мутацією, в порівнянні з хлопчиками і дітьми домутаційного віку, відрізнявся гарною якістю, на що вказували досить високі значення DSI: від 0,2 до 0,8, середні значення 2,3. Невід'ємною перевагою акустичного аналізу є можливість порівняння даних в процесі спостереження за дитиною і становлення вокальних навичок.

Таблиця 1.4.4

Групи	Акустичні характеристики									
	F0 Гц	MPT с	Jitter %	DSI	Fmax Гц	Fmin Гц	Fdelta Гц	SPLmax дБ	SPLmin дБ	SPLdelta дБ
Діти 7-12 років	243	10.4	1.2	-0.4	548	212	352	86.7	61.3	28.5
Мутація										
Хлопчаки 13-15 років	206	12.9	1.6	-1.6	351	138	214	89.4	61	27.9
Дівчата 13-15 років	243	12.9	0.8	2.3	740	221	520	92	56.9	35.9

## ВИСНОВКИ ДО 1 РОЗДІЛУ

Спектральний комп'ютерний аналіз голосової функції є високоінформативним методом дослідження, що дозволяє оцінити такі характеристики голосу, як частота основного тону, кількість областей резонансного посилення звуку (формант), число та регулярність обертонів, частота вібрато.

При обстеженні пацієнтів із застосуванням вище описаної методики можливе виявлення, що не діагностується традиційними методами раніше, патології: фонастенії; порушення реєстрів голосознавства, що провокує розвиток як функціональних, так і органічних дисфоній. Також, спектральний аналіз голосу є визначальним методом у вирішенні спірних питань визначення типу голосу в осіб професій з активним використанням голосу, що пред'являють високі вимоги до якості голосу (педагоги, перекладачі, диктори, вокалісти).

Отже, комп'ютерний аналіз голосу є об'єктивним, інформативним методом оцінки голосової функції. Метод забезпечує ранню та диференціальну діагностику голосових порушень, а також може бути використаний при визначенні тактики лікування у пацієнтів як з функціональними, так і з органічними захворюваннями голосового апарату. Особливо ефективний цей метод в осіб, які є представниками професій з використанням голосу.

## РОЗДІЛ 2. АЛГОРИТМИ ТА МАКЕТУВАННЯ ВІДПОВІДНИХ ПРОГРАМНИХ ФУНКЦІЙ РОЗРАХУНКУ ДІАГНОСТИЧНИХ ОЗНАК

### 2.1 Трасування частоти основного тону

Серед існуючих методів розрахунку частоти основного тону варто навести наступні [22]:

- Кепстральний метод;
- Метод з визначенням першого спектрального піку на спектрі сигналу;
- Автокореляційний метод.

Кепстр – це результат зворотного перетворення Фур’є від логарифму спектру потужності сигналу. Відповідно область визначення на графіку кепстру є часовою областю, яку називають сachtотою. Типовий графік кепстру зображено на рис. 2.1.1:

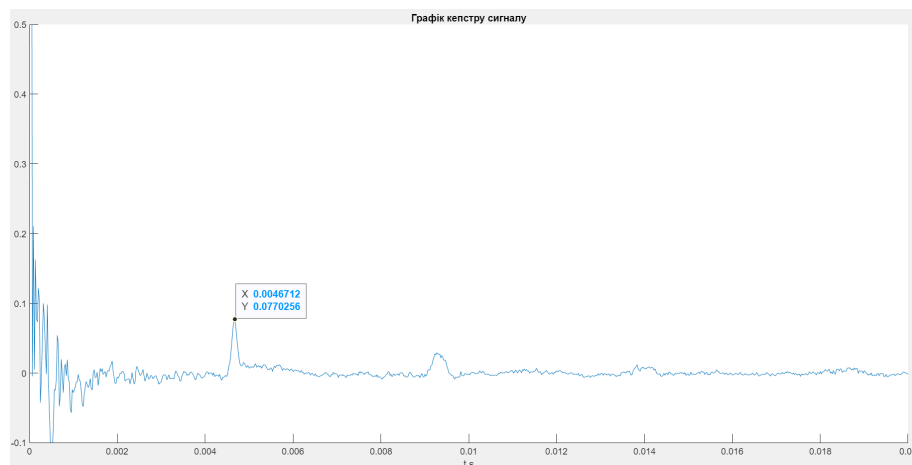


Рис. 2.1.1 Графік кепстру сигналу

Маркером позначено низькочастотний пік, що відповідає частоті:

$$F0 = \frac{1}{0.0046} = 217 \text{ Гц}$$

Відповідно, враховуючи те, що частота основного тону голосу людини знаходиться в певних обмежених рамках, можна розробити алгоритм визначення сachtоти піку в рамках сachtот, що відповідають обраним частотним рамкам.

З іншого боку частоту основного тону можна оцінити по графіку спектру сигналу. Для цього необхідно визначити частоту першого частотного піку (рис. 2.1.2).

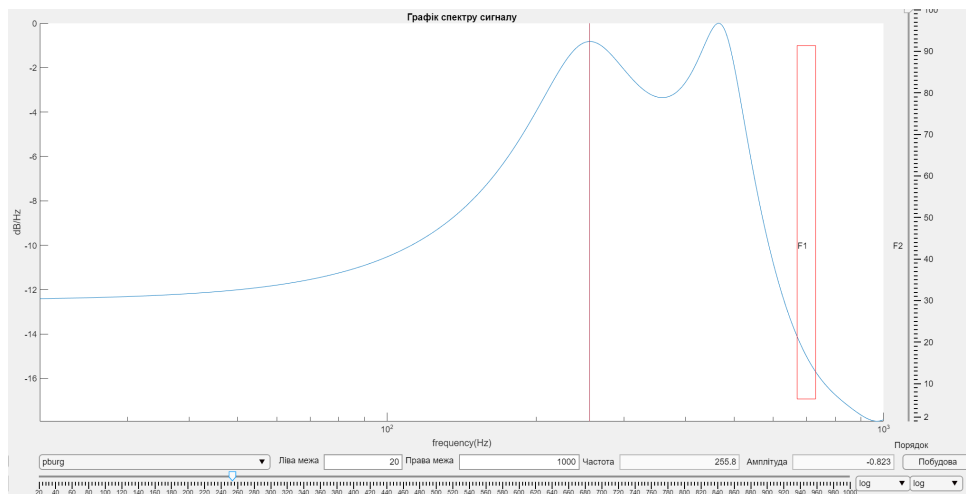


Рис. 2.1.2 Графік спектру сигналу

Звідси видно, що частота основного тону сигналу становить 255 Гц, що відрізняється від встановленого значення по кепстру цього ж сигналу. Це пов'язано з основним недоліком цього методу – неточністю, що пов'язана з вибором точності оцінки спектру (порядком, для випадку авторегресійної оцінки спектру, та розміру сегменту, для випадку непараметричного методу модифікованої періодограми Уелча).

Автокореляційний метод будується на тому, що результатом автокореляції гармонічного сигналу буде гармонічний сигнал з тією ж частотою. Для випадку сигналу, що має велику кількість обертонів, як у голосу, необхідно враховувати наявність додаткових гармонік на графіку автокореляції. Графік автокореляції запису голосу людини наведено на рис. 2.1.3.

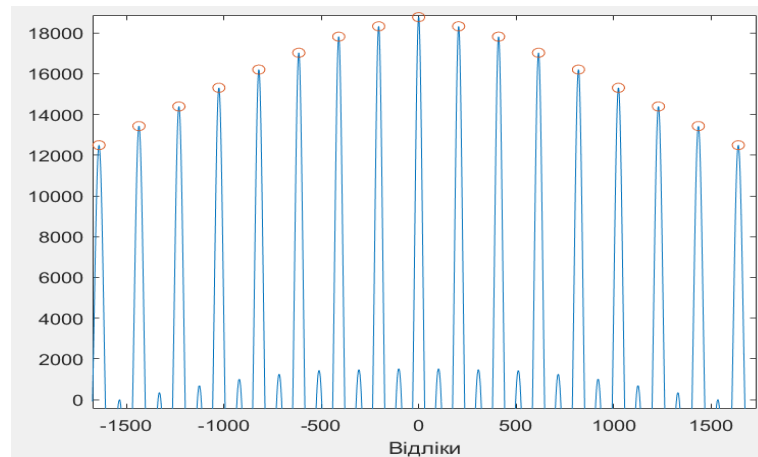


Рис.2.1.3 Автокореляція запису голосу людини

Кружечками виділено піки автокореляційної функції, що відповідають частоті основного тону, знизу піки, що відповідають обертонам. Враховуючи те, що у голосі людини в більшості випадків частота основного тону потужніша за обертони в декілька разів, можна обмежитися аналізом піків, що знаходяться в межах 50-100% від максимуму автокореляційної функції. У випадку наявності у записі потужної шумової складової, необхідно видалити потужний нульовий пік.

Відповідно лістинг розробленої М-функції має наступний вигляд:

```
function F0 = FundamentalFreqAcorr(vec,fs)
[corr,~]=xcorr(vec,vec); %автокореляційна функція
[~,lc]=max(corr); corr(lc)=0;
[maxc,~]=max(corr); corr(lc)=maxc;
[~,lcm] = findpeaks(corr,'MinPeakHeight',0.5*max(corr));
lcmd=zeros(1,length(lcm)-1);
for i=2:length(lcm)
    lcmd(i-1)=lcm(i)-lcm(i-1);
end
F0=fs/mean(lcmd);%розрахунок частоти основного тону як обернене до
% середнього значення відстані між максимумами кореляційної функції
end
```

Спершу отримується автокореляційна функція, з якої другим кроком вирізається нульовий максимум. Третім кроком знаходиться наступний після нульового максимум, значення якого присвоюється нульовому.

Функція `findpeaks` дозволяє отримати позиції максимумів автокореляційної функції. Значення атрибуту `'MinPeakHeight'` обирається як 50 відсотків від нового максимуму, завдяки чому можна відсіяти гармоніки обертонів.



Циклом перебираються позиції максимумів та розраховуються відстані між сусідніми, після чого частота основного тону відповідає частоті дискретизації помноженій на середню частоту появ максимумів.

Розрахована частота основного тону цим методом становить 215 Гц.

Для того, щоб отримати графік частоти основного тону, необхідно розбити запис на сегменти. Довжина сегментів впливає на точність та плавність графіку, тому в фінальному інтерфейсі необхідно передбачити вибір тривалості сегменту розбиття. Експериментально визначено, що задовільними є тривалості сегменту в межах 30-150 мс.

Для згладжування графіку ЧОТ використовується функція `envelope` з аргументом `'Peak'`. Результатом є верхня та нижня огинаючі, що визначаються за допомогою сплайн-інтерполяції за локальними максимумами, розділеними обраною кількістю вибірок (обрано 5). Поверх графіку ЧОТ будується середнє арифметичне між верхньою та нижньою огинаючою.

Іноді при малій тривалості сегменту трапляються випадки неможливості визначення ЧОТ. В таких випадках функція визначення частоти основного тону видає значення NaN. Щоб позбутися цього використовується `isnan`, що видає одиницю у випадку, коли її аргументом є значення NaN. Тоді відібраний умовним оператором локальний ЧОТ замінюється попереднім.

В результаті маємо наступну М-функцію:

```
function [tvec,vecF0,f0vecEnv] = FundamentalFreqAcorrVec(vec,fs,tseg)
N=floor(tseg*fs); %довжина сегменту
Nseg=floor(length(vec)/N); %кількість сегментів
vecF0=zeros(1,Nseg);
for t=1:Nseg
    seg=vec(((t-1)*N+1):(t*N));%розбиття на сегменти
    vecF0(t)=FundamentalFreqAcorr(seg,fs); %розрахунок частоти основного тону
    if isnan(vecF0(t)) == 1
        vecF0(t)=vecF0(t-1);
    end
end
[up,lo]=envelope(vecF0,5,'Peak');
f0vecEnv=(up+lo)/2;
tvec=tseg*(0:Nseg-1);
end
```

## 2.2. Розрахунок Shimmer та Jitter

Як було вказано вище параметр Jitter визначається як частотно-фазова модуляція (власне тремтіння) сигналу. Jitter розраховується як середнє значення різниць періодів сигналу. Для отримання відносного значення jitter необхідно віднести то середньої тривалості періоду сигналу.

Тривалості періодів та їх різниці визначаються за нижніми та верхніми піками графіку сигналу. Для цього використовується функція findpeaks. Її аргументом є 'MinPeakHeight' та 'MinPeakDistance'. Перший аргумент вказує на мінімальну висоту піків. Її необхідно визначати для кожного запису персонально. Для цього в інтерфейсі необхідно передбачити слайдер вибору верхньої та нижньої межі. На рис.2.2.1 продемонстровано вірний вибір верхньої межі, та невірний вибір нижньої межі:

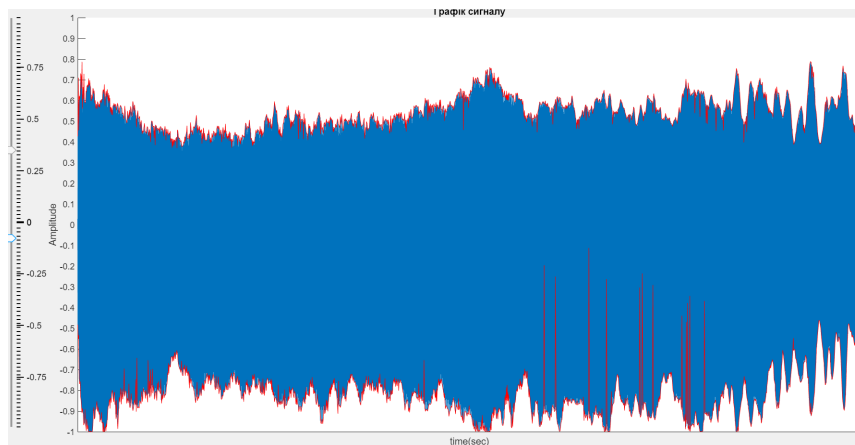


Рис.2.2.1 Графік побудови верхніх та нижніх піків для визначення Shimmer та Jitter

Параметр 'MinPeakDistance' визначається як відсоток від періоду основного тону. Вибір цього відсотку має бути реалізовано у фінальному інтерфейсі.

Параметр Shimmer визначається як відношення середнього значення різниць амплітуд до середнього значення амплітуд. В результаті маємо наступну М-функцію:

```
function [Jitter,JitterPercent,ShimmerPercent,peaksUp,nUp,peaksDown,nDown] =
ShimmerJitter(vec,fs,MPHD,MPHU,perc)
vec=vec-mean(vec); vec=vec/max(abs(vec));
F0 = FundamentalFreqAcorr(vec,fs);
F0=floor(F0);

[peaksUp,nUp] = findpeaks(vec,'MinPeakHeight',MPHU,'MinPeakDistance',perc*(fs/F0));
[peaksDown,nDown] = findpeaks(-vec,'MinPeakHeight',MPHD,'MinPeakDistance',perc*(fs/F0));

N=min(length(nUp),length(nDown));
Tup(1) = nUp(2)-nUp(1);
Tdown(1) = nDown(2)-nDown(1);
Aup(1) = peaksUp(2)-peaksUp(1);
```

```

Adown(1) = peaksDown(2)-peaksDown(1);
for i=2:N-1
    Tup(i) = nUp(i+1)-nUp(i);
    Tdown(i) = nDown(i+1)-nDown(i);
    Tdifdown(i-1) = abs(Tdown(i)-Tdown(i-1));
    Tdifup(i-1) = abs(Tup(i)-Tup(i-1));
    Adifup(i) = abs(peaksUp(i)-peaksUp(i-1));
    Adifdown(i) = abs(peaksDown(i)-peaksDown(i-1));
end

Tmean = mean([Tup Tdown]);
Tdiffmean = mean([Tdifdown Tdifup]);
Amean = mean([peaksUp' peaksDown']);
Adiffmean = mean([Adifdown Adifup]);
Jitter=Tdiffmean/fs;
JitterPercent = 100*Tdiffmean/Tmean;
ShimmerPercent = 100*Adiffmean/Amean;
end

```

## 2.3 Розрахунок відношення гармонік до шуму (HNR)

Як вказувалося раніше, HNR - це оцінка співвідношення між періодичними та неперіодичними компонентами голосового сигналу. Значення HNR оцінюється з автокореляційної функції цього сигналу. Відомо, що нульовий максимум автокореляційної функції є сумою автокореляції гармонічної та шумової складової, тоді як наступний за ним максимум, в позиції першого періоду коливальності основного тону, містить вже інформацію про потужність лише гармонічної складової. Відповідно  $AC_v(T)$  - це потужність гармонічної складової, тоді як  $AC_v(0) - AC_v(T)$  потужність шумової складової. Результатом їх відношення є значення HNR.

Для реалізації цього знову використовується функція `findpeaks`, параметр `'MinPeakDistance'` якої відповідає частоті основного тону. Після чого знаходиться позиція та значення нульового та наступного за ним максимуму. Результируюча M-функція має наступний вигляд:

```

function [HNRdB, Autocor, Peak, lc, N] = HNR(vec, fs)
vec=vec-mean(vec); vec=vec/max(abs(vec));
F0 = FundamentalFreqAcorr(vec, fs);
F0=floor(F0);

Autocor=xcorr(vec, vec);
[Peak, lc] = findpeaks (Autocor, 'MinPeakDistance', 0.5*(fs/F0));

[MaxPeak, N] = max(Peak);
HNRdB = 10*log10(Peak(N+1)/(MaxPeak-Peak(N+1)));
end

```

## 2.4. Час максимальної фонації

Крім найпростішого наочного методу визначення часу максимальної фонації, шляхом встановлення часових слайдерів на графіку сигналу в положення початку та кінця запису, відокремлюючи ділянки тиші, можна реалізувати алгоритм підрахунку вибірок, огинаюча нормованого до одиниці запису яких більша за обраний поріг. В даному випадку обрано поріг в -6 дБ. Результуюча М-функція має наступний вигляд:

```
function [MPT] = MPTTime(vec,fs)
vec=vec-mean(vec); vec=vec/abs(max(vec));
env=envelope(vec,150,'rms');
envdB=10*log10(env/max(env));
mpn=0;
for i=1:length(envdB)
    if envdB(i) > -6
        mpn=mpn+1;
    end
end
MPT=mpn/fs;
end
```

## 2.5 Індекс тяжкості дисфонії

Індекс тяжкості дисфонії DSI є показником, що розраховується з чотирьох важелів з їхніми коефіцієнтами. Для цього використовуються вже визначені показники Jitter та час максимальної фонації. Крім того необхідно визначити мінімальну ЧОТ з запису зі спадною ЧОТ, а також мінімальну інтенсивність з вправи, де пацієнт проспівує /a/, плавно переходячи з нормальної мови до найтихішої, не переходячи на шепіт. Мінімальну інтенсивність необхідно оцінювати або каліброваною системою запису, або шумоміром. Для розрахунку DSI необхідно передбачити віконце для вписування мінімальної інтенсивності.

М-функція визначення індексу тяжкості дисфонії має наступний вигляд:

```
function DSI = DSI(MPT,F0High,ILow,JitterPercent)
DSI=0.13*MPT+0.0053*F0High+0.26*ILow-1.18*JitterPercent+12.4;
end
```

## 2.6 Мел-частотні кепстральні коефіцієнти

Як альтернативний спосіб опису голосового сигналу можна використати його мел-частотне кепстральне уявлення. Розрахунок мел-кепстральних коефіцієнтів (МЧКК) відноситься до методів короточасного аналізу голосового сигналу, які припускають розбиття сигналу на інтервали аналізу. Як правило, в інтервалі від 10

до 30 мс голосовий сигнал вважається стаціонарним. В даному випадку аналіз виконувався на інтервалах тривалістю 20 мс, які мали половинне перекриття в 10 мс. Обчислення МЧКК виконується у частотній шкалі мелів, яка враховує специфіку сприйняття висоти звуку людським вухом.

Внаслідок короткочасного аналізу формується великий набір МЧКК, які описують локальну структуру сигналу. Для переходу до більш глобального подання та зменшення обсягу даних МЧКК, що виходять, усереднюються для формування надсегментного вектора ознак.

Маємо наступний алгоритм мел-частотного кепстрального аналізу:

Крок 1. Корекція спектра сигналу, що полягає у вирівнюванні енергій високочастотної та низькочастотної складових за допомогою фільтра, який має підйом амплітудно-частотної характеристики приблизно 6 дБ на октаву:

$$s(n) = x(n) - 0,82 \cdot x(n-1)$$

Крок 2. Розбиття сигналу на кадри тривалістю N відліків (і з перекриттям N/2 відліків) і зважування їх за допомогою вікна Хеммінга:

$$w(n) = 0,54 + 0,46 \cdot \cos(2\pi n / N)$$

Далі для всіх кадрів  $s_t(n)$ , де t – номер кадру, виконуються кроки 3–6.

Крок 3. Розрахунок короткочасного спектру за допомогою швидкого перетворення Фур'є для кожного аналізованого кадру:

$$S_t(k) = \sum_{n=0}^{N-1} w(n) \cdot s_t(n) \cdot e^{\frac{-j2\pi kn}{N}}, \quad k = 0, 1, \dots, N-1.$$

Крок 4. Розрахунок набору з M фільтрів (m=1, 2, ..., M) трикутної форми:

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \frac{k - f(m-1)}{f(m) - f(m-1)}, & f(m-1) \leq k \leq f(m); \\ \frac{f(m+1) - k}{f(m+1) - f(m)}, & f(m) \leq k \leq f(m+1); \\ 0, & k > f(m+1), \end{cases}$$

які використовуються для усереднення спектру поблизу центральних частот  $f(m)$ , розташованих рівномірно у шкалі мелів (рис. 2.6.1). У даному розрахунку параметр  $M = 19$  позначає кількість критичних смуг у аналізованому частотному діапазоні.

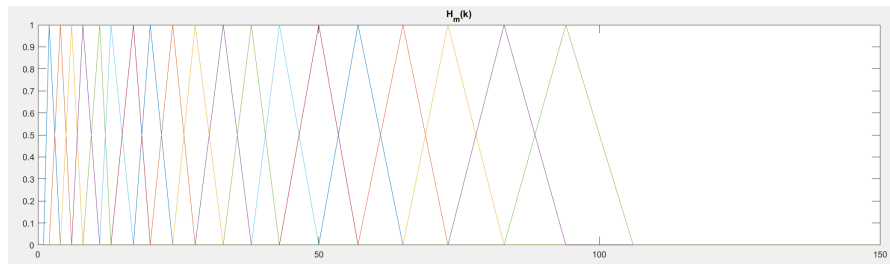


Рис.2.6.1 Трикутні фільтри для розрахунку мел-спектру

Якщо  $f_l$  і  $f_h$  – нижня та верхня межі частотного діапазону (Гц), що покриває набір фільтрів, то граничні частоти  $f(m)$ , рівномірно розташовані в шкалі мелів, розраховуються за формулою

$$f(m) = \frac{N}{F_s} B^{-1} \left( B(f_l) + m \frac{B(f_h) - B(f_l)}{M+1} \right),$$

Де функція  $B$  являє собою перехід від шкали герців до шкали мелів:

$$B(f) = 1127 \cdot \ln(1 + f / 700).$$

Значення  $f_l$  і  $f_h$  вибиралися рівними 50 та 5300 Гц відповідно. Зазначений діапазон частот відповідає першим 19 критичним смугами слуху. З іншого боку, у нього потрапляє основна енергія аналізованого у роботі голосного звуку /a/.

Крок 5. Застосування набору фільтрів, отриманого на кроці 4 для розрахунку логарифму від енергії на виході кожного фільтра:

$$Y_t(m) = \ln \left( \sum_{k=0}^{N-1} |S_t(k)|^2 \cdot H_m(k) \right), \quad 0 \leq m \leq M.$$

Крок 6. Обчислення МЧКК за допомогою дискретного косинусного перетворення:

$$c_t(k) = \sum_{m=0}^{M-1} Y_t(m) \cdot \cos(\pi k(m+1/2)/M)$$

Де  $k=0, 1, \dots, M-1$ .

Крок 7. Обчислення кінцевих різниць МЧКК, які застосовуються поряд із МЧКК як інформаційні ознаки:

$$\Delta c_t(k) = c_t(k) - c_{t-1}(k)$$

де  $c_t(k)$  – це  $k$ -й коефіцієнт, обчислений для кадру з номером  $t$ .

Крок 8. Розрахунок надсегментних ознак на основі МЧКК для отримання характеристичного вектора. Для цього послідовності  $c_t(k)$  та  $\Delta c_t(k)$  розбиваються на сегменти тривалістю 0,8 с, за якими обчислюються їх середні значення  $c^\mu(k)$ ,  $\Delta c^\mu(k)$  та середньоквадратичні відхилення (СКВ)  $c^\sigma(k)$  та  $\Delta c^\sigma(k)$ .

В результаті маємо наступний скрипт розрахунку МЧКК:

```
[x,fs]=audioread('Шлях до файлу');
x=x/max(abs(x)); x=x-mean(x);
tseg=20e-3;
tnseg=800e-3;
N=floor(tseg*fs); N=N-mod(N,2);
M=19;
fl=50;
fh=5300;

%Крок перший: корекція сигналу з вирівнюванням енергій ВЧ та НЧ складових
s=zeros(1,length(x)-1);
for i=2:(length(x))
    s(i)=x(i)-0.82*x(i-1);
end

%Крок другий: розбиття сигналу на сегменти довжиною N з перекриттям N/2 та
%взважування їх вікном Хеммінга

w=hamming(N)'; w=1-w+min(w);

Nsize=length(s)-mod(length(s),N); %Округлення досліджуваного відрізка до кратності N
Nseg=2*Nsize/N; %кількість сегментів
```

```

sseg=zeros(Nseg-1,N);
for t=1:Nseg-1
    sseg(t,:)=s(((t-1)*N/2+1):((t-1)*N/2+N));
end

%Крок третій: розрахунок короткочасного спектру за допомогою ШПФ для
%кожного сегменту, що аналізується

Sseg=zeros(Nseg-1,N);
for t=1:Nseg-1
    Sseg(t,:)=fft(sseg(t,:).*w);
end
figure,plot(Sseg(:,1));
%Крок четвертий: розрахунок набору з фільтрів трикутної форми

bfl=1127*log(1+f1/700); %менша частота досліджуваного діапазону в мелах
bfh=1127*log(1+fh/700); %більша частота досліджуваного діапазону в мелах
f=zeros(1,M);
for m=1:M
    bf=bfl+m*(bfh-bfl)/(M+1); %центральна частота кожного фільтра в мелах
    f(m)=floor((N/fs)*(exp(bf/1127)-1)*700); %перерахунок в Гц
end
fffh=floor((N/fs)*(exp(bfh/1127)-1)*700); %перерахунок верхньої частоти в
H=zeros(M,N);
for k=1:N %формування першого фільтра
    if (k>=0)&&(k<=f(1))
        H(1,k)=(k-1)/(f(1)-1);
    elseif (k>=f(1))&&(k<=f(2))
        H(1,k)=(f(2)-k)/(f(2)-f(1));
    end
end
figure; plot(H(1,:)), hold on;
for m=2:M %формування фільтрів, що лишилися
    for k=1:N
        if (k>=f(m-1))&&(k<=f(m))
            H(m,k)=(k-f(m-1))/(f(m)-f(m-1));
        elseif (k>=f(m))&&(m<M)&&(k<=f(m+1))
            H(m,k)=(f(m+1)-k)/(f(m+1)-f(m));
        elseif (k>=f(m))&&(m==M)&&(k<=fffh)
            H(m,k)=(fffh-k)/(fffh-f(m));
        end
    end
    plot(H(m,:))
end
title('H_m(k)'), xlim([0 150])
Y=zeros(Nseg-1,M);

%Крок п'ятий: застосування набору фільтрів та розрахунок логарифму енергій
%на виході кожного фільтра

for m=1:M
    for t=1:Nseg-1
        Y(t,m)=log(sum(abs(Sseg(t,:)).^2.*H(m,:)));
    end
end

%Крок шостий: розрахунок мел-частотних кедральних коефіцієнтів (МЧКК) за
%допомогою дискретно-конусного перетворення

for t=1:Nseg-1

```



```

for k=1:M
    c(t,k)=sum(Y(t,:)*cos(pi*(k+1)*(m+1.5)/(M+1)));
end
end
%Крок сьомий: розрахунок кінцевих різниць МЧКК - додаткової інформаційної
%ознаки

for t=2:Nseg-1
    dc(t-1,:)=c(t,:)-c(t-1,:);
end
%Крок восьмий: відбір сорока сегментів з отриманих, що відповідають
%0.02*40=0.8 с тривалість сегменту та розрахунок середніх значень та їх
%середньоквадратичних відхилень
figure;
cmu=mean(c(40:80,:)); subplot(221), plot(cmu), title('c^\mu')
dcmu=mean(dc(40:80,:)); subplot(222), plot(dcmu), title('\delta c^\mu')
cs=std(c(40:80,:)); subplot(223), plot(cs), title('c^\sigma')
dcs=std(dc(40:80,:)); subplot(224), plot(dcs), title('\delta c^\sigma')

```

Результатом виконання скрипту є наступні графіки:

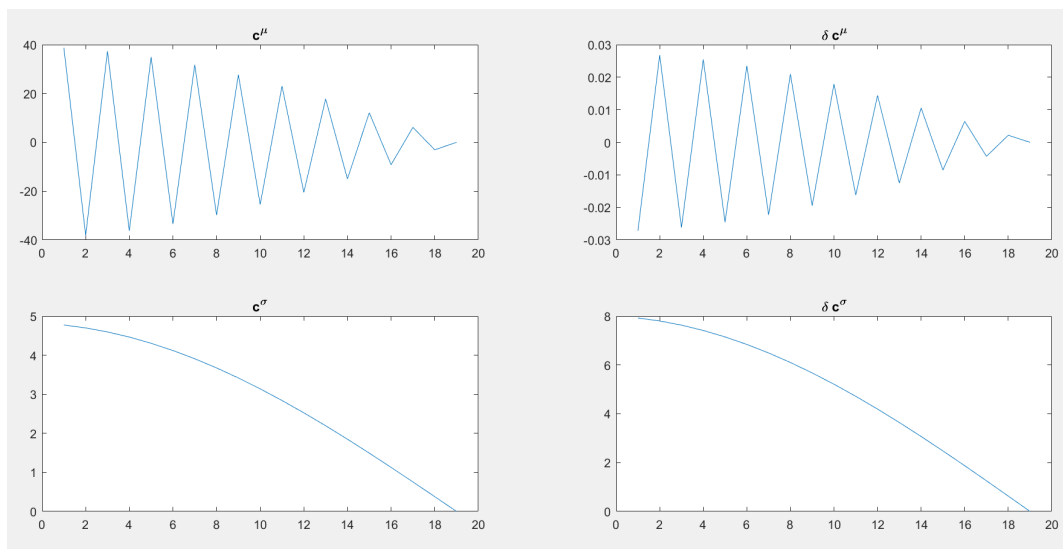


Рис.2.6.2 Результат розрахунку МЧКК

В роботі [22] продемонстровано зв'язок між діагностичними ознаками програмного середовища LingWaves та мел-частотними коефіцієнтами. Для цього проводиться відбір ознак методом LASSO, заснований на вирішування задачі лінійної регресії зі штрафною функцією, базуючись на базі 60 пацієнтів.

## ВИСНОВКИ ДО 2 РОЗДІЛУ

В даному розділі продемонстровано алгоритми та результати макетування відповідних програмних функцій розрахунку діагностичних ознак: частота основного тону (в тому числі мінімальна, максимальна та різниця між ними),

Shimmer, Jitter, відношення гармонік до шуму (HNR), час максимальної фонації та індекс тяжкості дисфонії (DSI).

Більшість цих параметрів потребує налаштування під час роботи з ними. Це означає, що в фінальному інтерфейсному рішенні необхідно передбачити тривалість сегментів розбиття для трасування ЧОТ, відсоток від ЧОТ при визначенні min peak distance, слайдери вибору min peak height верхньої та нижньої. Також, враховуючи неможливість оцінки мінімальної інтенсивності за допомогою мікрофону з звуковою картою, необхідно залишити віконце для запису результату оцінки за допомогою шумоміру.

Крім того було отримано алгоритми та виконано макетування обчислення мел-частотних кепстральних коефіцієнтів, які мають великі перспективи у подальшій розробці автоматизованої системи діагностування патологій. При тому, що МЧКК володіють значною інформативністю, вони програють іншим в своєму суттєвому недоліку – відсутність зрозумілої інтерпретації для лікаря-фоніатра.

Усі наведені діагностичні ознаки візуалізуються для суб'єктивної діагностики людиною-фоніатром. Розробка автоматизованої об'єктивної системи потребує складної співпраці з лікарями та їх пацієнтами, збору бази записів їх голосів та відповідності обчислених параметрів їх захворюванням.

## РОЗДІЛ 3. РОЗРОБКА ІНТЕРФЕЙСУ ПРОГРАМИ КОМП'ЮТЕРНОЇ ПІДТРИМКИ ФОНІАТРА

### 3.1 Демонстрація попередньої розробки

Попередньо [23] було розроблено інтерфейс програмного рішення для допомоги лікарю фоніатру. На рис.3.1.1 зображено зовнішній вигляд цієї програми.

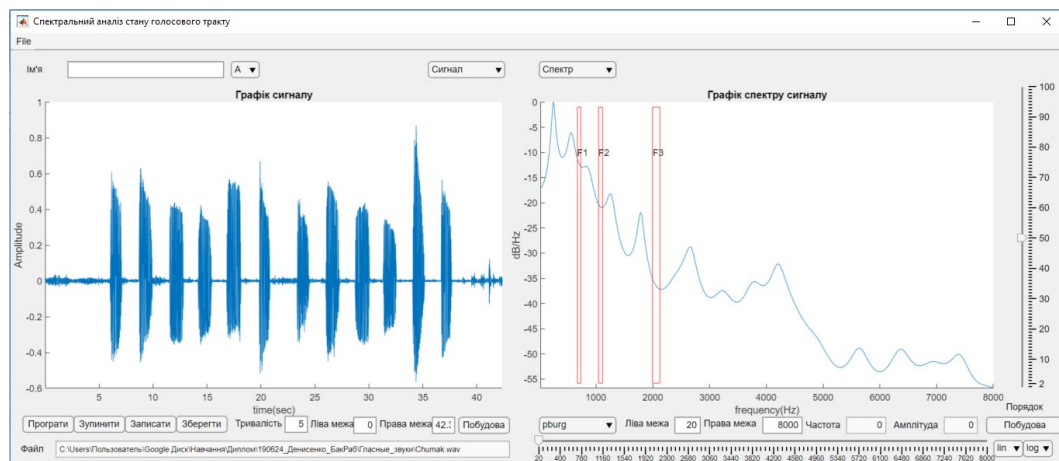


Рис.3.1.1 Існуюче програмне рішення [23]

Як було продемонстровано у першому розділі, одним з методів діагностики є суб'єктивний аналіз лікарем графіку спектру сигналу. Спираючись на лікарський досвід, фоніатр візуально визначає формантний розподіл потужності оберτονів на графіку спектру. В першому розділі наведено межі формантних областей голосних фонем, вони ж будуються поверх графіку спектру для допомоги лікарю. Випадаючий список вгорі відповідає за те, яку саме фонему аналізується.

Для початку роботи є два варіанти: обрати існуючий запис, обравши «Open» з меню «File», або записати новий, натиснувши «Записати». Також, для випадку використання файлу, на якому є щось крім необхідної фонемі полями «Ліва межа» та «Права межа» необхідно обрати межі початку й кінця сегменту, який треба аналізувати.

Текстове поле «Тривалість» вказує на те, скільки буде тривати запис. Натиснувши «Програти» можна прослухати записаний або завантажений файл. За допомогою «Зберегти» можна зберегти записаний, або завантажений та відредагований файл на диск.

Для оцінки спектру обрано два методи: параметричний авторегресійний метод Берга та непараметричний метод модифікованої періодограми Уелча. Є можливість перемкнутися між ними, обрати порядок моделі в першому випадку та розмір сегменту розбиття в другому випадку. Окремі перемикачі відповідають за вибір лінійної чи логарифмічної шкали абсцис та ординат[24].

Як було також продемонстровано у другому розділі, лікар може опиратися при оцінці частоти основного тону на графік кепстру сигналу. Для цього вгорі спектральної частини програми є перемикач. Графік кепстру було продемонстровано у другому розділі.

Окремим корисним інструментом є спектрограма. Її перемикач знаходиться над графіком сигналу. На рис.3.1.2 продемонстровано вигляд спектрограми сигналу:

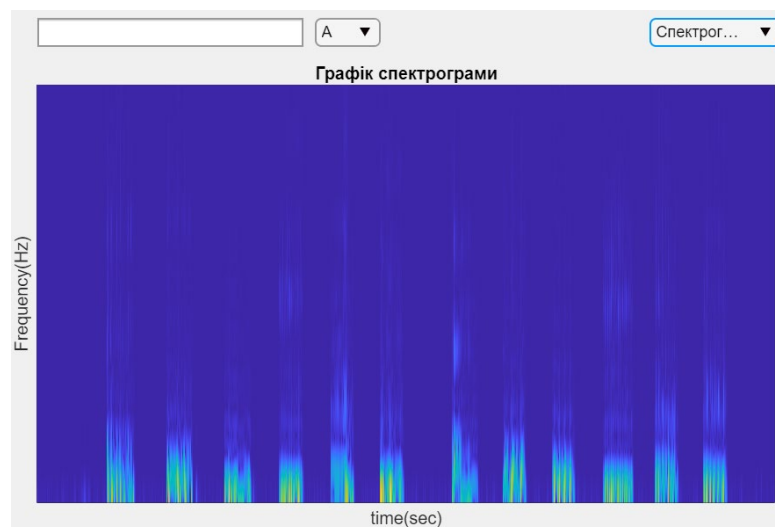


Рис.3.1.2 Графік спектрограми запису голосу [23]

### 3.2. Впровадження нових розроблених функцій в існуючий інтерфейс

В висновку до попереднього розділу описано які інтерфейсні рішення мають бути передбачені у новому програмному продукті. Це низка слайдерів для вибору мінімальної висоти та ширини визначення піків функцією `findpeaks`, та слайдер вибору ширини сегментного розбиття для трасування частоти основного тону.

Крім того необхідно передбачити поле для введення мінімальної інтенсивності голосу при співі затихаючої голосної фонеми не переходячи на шепіт. Лікар фоніатр

має виміряти цю інтенсивність за допомогою шумоміру. Для візуального аналізу будуються графіки спадного тону, та того, що зростає. Також для візуального контролю вірності визначення відношення гармонік до шуму будується графік автокореляції запису, масштабований до перших двох піків.

На рис.3.2.1 зображено зовнішній вигляд програми після виконання всіх етапів розрахунку:

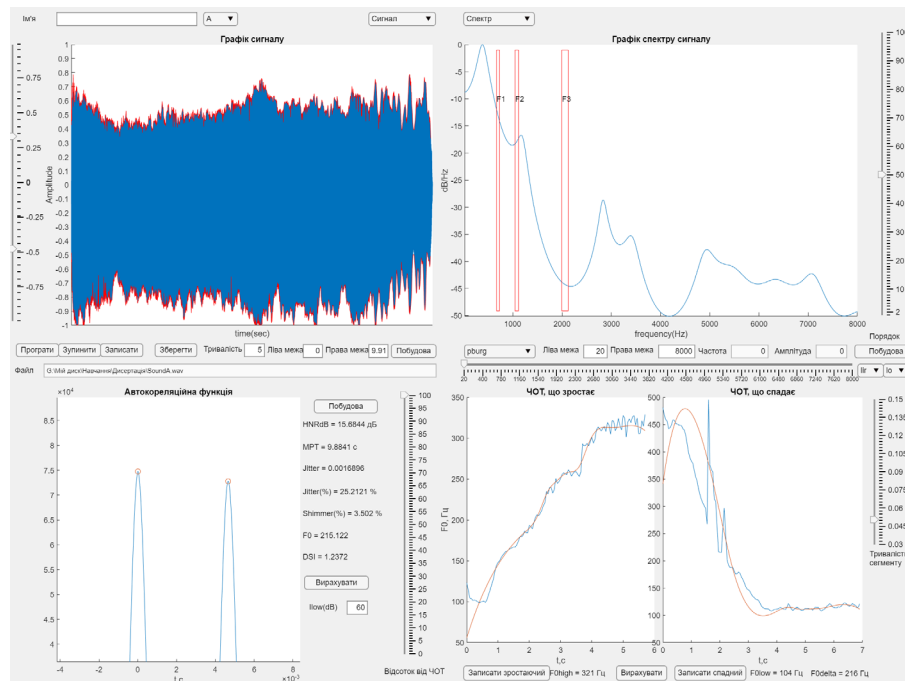


Рис.3.2.1 Зовнішній вигляд фінального програмного рішення

В першому розділі наведено приклади того, як лікарі працюють з діагностичними ознаками, в другому розділі того, як ці діагностичні ознаки обчислювати. В цьому ж необхідно продемонструвати алгоритм роботи з отриманим програмним рішенням.

Першим кроком необхідно обрати з диску або зробити новий запис співу голосної фонемі максимальної протяжності такої гучності та висоти, що є типовою для пацієнта. Після цього необхідно обрати ліву та праву межу сегменту для того, щоб обрізати тишу на початку та кінці запису.

Після вибору відсотку від ЧОТ на відповідному слайдері, та верхньої та нижньої межі визначення мінімальної висоти піків на графіку сигналу червоним визначиться контур огинаючої, що складається з цих піків. У відповідних текстових полях з'являться розраховані значення Jitter, Jitter(%) та Shimmer, а також MPT – час максимальної фонації.

Далі необхідно зробити два нових записи: фонему зі зростаючим тоном, та фонему зі спадним тоном. Для цього в коді функції кнопки передбачено запис тривалістю 10 секунд. Важливо наголосити на репетиції цієї вокальної вправи та виконанні декількох записів. Слайдером вибору тривалості сегменту розбиття можна модифікувати графіки. Кнопка «Вирахувати» будує графіки заново та обчислює мінімальну та максимальну ЧОТ, демонструє їх положення на графіках та в окремому текстовому полі вписується різниця між максимальним та мінімальним ЧОТ.

Якщо натиснути кнопку «Побудова», на лівому нижньому графіку побудується графік автокореляції, а поруч у текстовому полі розрахується відношення гармонік до шуму. Слайдером «Відсоток від ЧОТ» можна відредагувати атрибут мінімального проміжку між піками для визначення HNR.

Вписавши значення  $I_{min}$  та натиснувши «Вирахувати» у відповідному полі відобразиться значення ЧОТ та індексу тяжкості дисфонії.

Для отримання мел-частотних кепстральних коефіцієнтів у верхньому полі меню знаходиться кнопка «МЧКК». Натиснувши її в окремому вікні будуються наступні графіки:

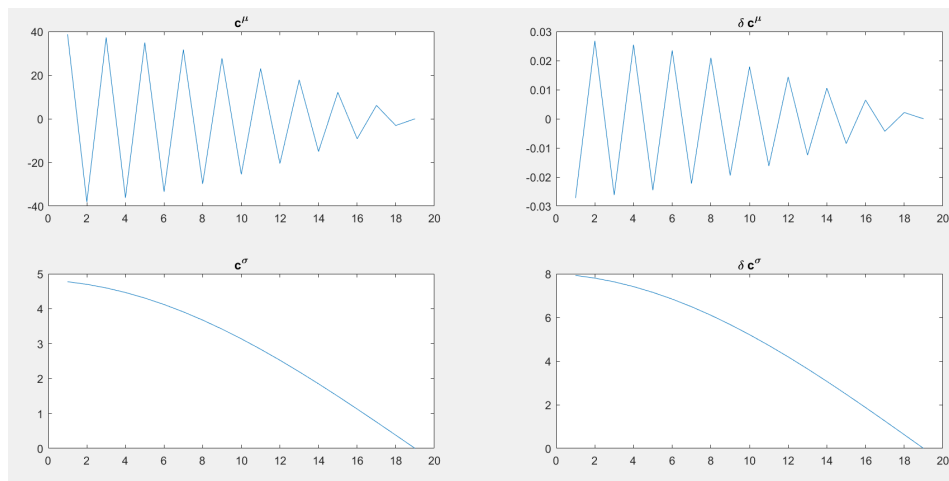


Рис.3.2.2 МЧКК запису голосу

## ВИСНОВКИ ДО 3 РОЗДІЛУ

Макетування модернізованої системи інструментального оцінювання параметрів голосового тракту дозволяє дійти висновку, що отриманий програмний продукт є зручним та інтуїтивним для використання лікарем фоніатром. Вивчивши порядок роботи з ним можна швидко отримати всі діагностичні ознаки, необхідні для роботи з ним.

Дане алгоритмічне та програмне рішення потребує подальшої роботи з впровадження у реальну роботу з лікарем та апробацію на можливі помилки. Ці помилки можуть бути наприклад тоді, коли якісь розрахунки проводяться тоді, коли в програму не завантажено аудіофайл. Крім того всі використані функції потребують спрощення та оптимізації розрахунку.

## ВИСНОВКИ

Хвороби голосу залишається серйозною проблемою в світі, що прогресує. Поруч з цим розвивається напрямок фоніатрії по завчасній діагностиці патологій органів голосу. При виникненні перших вказівок на вади голосу, пацієнт звертається до лікаря, котрий суб'єктивно оцінює об'єктивні діагностичні ознаки.

В роботі продемонстровано ряд діагностичних ознак, що часто використовуються фоніатрами. До таких відносяться: Shimmer, Jitter, HNR, частота основного тону (в тому числі мінімальна та максимальна, а також їх діапазон), мінімальна та максимальна інтенсивність, час максимальної фонації. По цим показникам розраховуються такі індекси захворювань, як DSI, Noise, Irregularity, Overall severity of dystonia. Розробники розкривають зміст індексу DSI, як похідне від наведених діагностичних ознак.

Проте досі поширеним залишається лише демонстрація цих ознак, по яких лікар має сам прийняти рішення. Повноцінна об'єктивна система потребує серйозної роботи з великою базою пацієнтів та записів їх голосів. Вона може бути побудована на базі мел-частотних кепстральних коефіцієнтів які, як показано, гарно корелюють з захворюваннями голосу. Для цього потрібно продовжувати роботу з їх аналізу та побудови системи прийняття рішення.

На базі обраних та систематизованих діагностичних ознак побудовано макет комп'ютерної системи, що є помічником лікаря в його роботі з пацієнтом. На робочому полі програми демонструються розраховані показники. В подальшому ця програма потребує апробації та тестів на можливі помилки, які легко вирішуються, оскільки виконані у простому та гнучкому середовищі.



## Літературні джерела

1. Україна входить до 30 країн за найбільшою часткою людей похилого віку — Мінсоцполітики [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: <https://hromadske.ua/posts/ukraina-vkhodyt-do-30-krain-za-naibilshoiu-chastkoiu-liudei-pokhyloho-viku-minsotspolityky>.
2. Шидловська Т. А. Функціональні порушення голосу / Т. А. Шидловська., 2011
3. Как на самом деле устроен речевой аппарат [Електронний ресурс] – Режим доступу до ресурсу: <https://yourspeech.ru/eloquence/talk/rechevoj-apparat.html>.
4. Л. Р. Рабінер, Р. В. Шафер «Цифровая обработка речевых сигналов».
5. Форманты [Електронний ресурс] // <https://corpuscul.net/shkola/postanovka-golosa/akkustika-golosovogo-apparata/formanty/>
6. How human control pitch [Електронний ресурс] – Режим доступу до ресурсу: <http://www.ncvs.org/ncvs/tutorials/voiceprod/tutorial/cover.html>.
7. L. Cveticanin. "Review on Mathematical and Mechanical Models of the Vocal Cord," Journal of Applied Mathematics, Vol. 2012, 2012. - P. 1136-1153.
8. Kotaro Fukui, Yuma Ishikawa, Eiji Shintaku, Keisuke Ohno, Nana Sakakibara, Atsuo Takanishi, Masaaki Honda. "Vocal Cord Model to Control Various Voices for Anthropomorphic Talking Robot," Proc. of 8th International Seminar on Speech Production, 2008. - P. 341-344.
9. J. L. Flanagan, "Source-system interaction in the vocal tract," Annals of the New York Academy of Sciences, vol. 155, pp. 9–17, 1968.
10. S. Adachi and J. Yu, "Two-dimensional model of vocal fold vibration for sound synthesis of voice and soprano singing," Journal of the Acoustical Society of America, vol. 117, no. 5, pp. 3213–3224, 2005.
11. Ishizaka K., Flanagan J. "Synthesis of Voiced Sound From a Two-Mass Model of the Vocal Cords," The Bell System Technical Journal, Vol. 51, No. 6, July-August, 1972. - P. 1233-1268.
12. Василенко Ю. С. Спектральный компьютерный анализ голоса – метод ранней и дифференциальной диагностики нарушений голосовой функции

- [Електронний ресурс] / Ю. С. Василенко – Режим доступу до ресурсу:  
<https://nikio.ru/спектральный-анализ-голоса/>.
- 13.Родькина В. А. Функциональные нарушения голоса в практике логопеда-фонопеда / Валентина Анатольевна Родькина. // / Научно-методический электронный журнал «Концепт». – 2019. – №12.
- 14.lingWAVES 3 - Acoustic Voice Assessment [Електронний ресурс] – Режим доступу до ресурсу:  
<https://www.wevosys.com/products/lingwaves/lingwaves.html>.
- 15.The Dysphonia Severity Index [Електронний ресурс] – Режим доступу до ресурсу: [https://www.wevosys.de/wissen/\\_data\\_wissen/4.pdf](https://www.wevosys.de/wissen/_data_wissen/4.pdf).
- 16.Diagnosis of Voice Disorders [Електронний ресурс] – Режим доступу до ресурсу: [https://www.med.or.jp/english/journal/pdf/2011\\_04/248\\_253.pdf](https://www.med.or.jp/english/journal/pdf/2011_04/248_253.pdf).
- 17.Maximum Phonation Time (MPT) [Електронний ресурс] – Режим доступу до ресурсу: <https://www.sltinfo.com/maximum-phonation-time/>.
18. LingWaves voice protocol [Електронний ресурс] – Режим доступу до ресурсу:  
<https://mmsp.com.au/mmsp/wp-content/uploads/2019/08/lingWAVES-Voice-Protocol-Cheat-Sheet-2017.pdf>.
- 19.Jitter and shimmer [Електронний ресурс] – Режим доступу до ресурсу:  
<https://wiki.aalto.fi/display/ITSP/Jitter+and+shimmer>.
- 20.Краснолуцький І. Є. Система діагностики стану вокального тракту людини на основі аналізу часових параметрів голосового сигналу / Краснолуцький Іван Євгенійович, 2021.
- 21.Кельман М. В. Психолого-педагогічне забезпечення збереження голосу дітей - вокалістів / Мар'яна Василівна Кельман. – 2020.
- 22.Куценко Є. Є. Оцінка частоти основного тону голосового сигналу користувача системи автентифікації / Є. Є. Куценко, М. С. Пастушенко. // Електронне наукове фахове видання – журнал «Проблеми телекомунікацій». – 2019. – №2.
- 23.Денисенко О. І. Спектральний аналіз вокальних звуків / Олександр Ігорович Денисенко. – 2020.

- 24.Денисенко О. І. Комп'ютерне моделювання алгоритмів діагностики голосової системи людини / Олександр Ігорович Денисенко. // Мікросистеми, Електроніка та Акустика.. – 2021.
- 25.Вашкевич М. И. Анализ акустических параметров голоса для выявления заболеваний гортани / М. И. Вашкевич. // ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ. – 2020.

## ДОДАТОК А. Скрипти програмних модулів

```
classdef FirstAttempt2 < matlab.apps.AppBase

% Properties that correspond to app components
properties (Access = public)
    UIFigure                matlab.ui.Figure
    FileMenu                matlab.ui.container.Menu
    OpenMenu                matlab.ui.container.Menu
    MenuMChKK               matlab.ui.container.Menu
    IlowdBEEditField        matlab.ui.control.NumericEditField
    IlowdBEEditFieldLabel   matlab.ui.control.Label
    ButtonCountDSI          matlab.ui.control.Button
    LabelDSI                matlab.ui.control.Label
    LabelF0                 matlab.ui.control.Label
    LabelShimmer            matlab.ui.control.Label
    LabelJitterPercent      matlab.ui.control.Label
    LabelJitter             matlab.ui.control.Label
    SliderPercent           matlab.ui.control.Slider
    Label_12                matlab.ui.control.Label
    LabelMPT                matlab.ui.control.Label
    LabelHNR                matlab.ui.control.Label
    ButtonBuildAcorr        matlab.ui.control.Button
    SliderTseg              matlab.ui.control.Slider
    Label_11                matlab.ui.control.Label
    LabelF0Delta            matlab.ui.control.Label
    ButtonCount             matlab.ui.control.Button
    LabelF0Low              matlab.ui.control.Label
    LabelF0High             matlab.ui.control.Label
    ButtonRecLow            matlab.ui.control.Button
    ButtonRecHigh           matlab.ui.control.Button
    SliderMPHdown           matlab.ui.control.Slider
    SliderMPHup             matlab.ui.control.Slider
    DropDownLetter          matlab.ui.control.DropDown
    YLoglin                 matlab.ui.control.DropDown
    XLoglin                 matlab.ui.control.DropDown
    DropDownSig             matlab.ui.control.DropDown
    DropDownSpec            matlab.ui.control.DropDown
    EditFieldName           matlab.ui.control.EditField
    Label_10                matlab.ui.control.Label
    ButtonSave              matlab.ui.control.Button
    EditFieldFreq           matlab.ui.control.NumericEditField
    Label_9                 matlab.ui.control.Label
    EditFieldSec            matlab.ui.control.NumericEditField
    Label_8                 matlab.ui.control.Label
    DropDownType            matlab.ui.control.DropDown
    ButtonRec               matlab.ui.control.Button
    ButtonStop              matlab.ui.control.Button
    ButtonPlay              matlab.ui.control.Button
    EditFieldFile           matlab.ui.control.EditField
    LabelFile               matlab.ui.control.Label
    EditFieldAmpl           matlab.ui.control.NumericEditField
    Label_6                 matlab.ui.control.Label
    SliderFreq              matlab.ui.control.Slider
    ButtonBuildTime         matlab.ui.control.Button
    EditRightSample         matlab.ui.control.NumericEditField
    LabelRightSample        matlab.ui.control.Label
end
```

EditLeftSample	matlab.ui.control.NumericEditField
LabelLeftSample	matlab.ui.control.Label
SliderOrder	matlab.ui.control.Slider
LabelOrder	matlab.ui.control.Label
ButtonBuildFreq	matlab.ui.control.Button
EditRightFreq	matlab.ui.control.NumericEditField
LabelRightFreq	matlab.ui.control.Label
EditLeftFreq	matlab.ui.control.NumericEditField
LabelLeftFreq	matlab.ui.control.Label
UIAxesDown	matlab.ui.control.UIAxes
UIAxesUp	matlab.ui.control.UIAxes
UIAxesAcor	matlab.ui.control.UIAxes
UIAxesFreq	matlab.ui.control.UIAxes
UIAxesTime	matlab.ui.control.UIAxes

end

properties (Access = public)

```

path2file % Description
SIG
SR
LeftFreq
RightFreq
order
LeftSample
RightSample
tvec
SlFreq
aplayer
arecorder
SIGis
cepstris
fs
tseg
tvechigh
vecF0high
f0vecEnvhigh
arecorderHigh
arecorderLow
tvecLow
vecF0low
f0vecEnvlow
F0Low
F0High
deltaF0
HNRdB
MPT
MPHup
MPHdown
Jitter
JitterPercent
ShimmerPercent
perc
ILow
high
low

```

end

```
% Callbacks that handle component events
```

```
methods (Access = private)
```

```
% Value changed function: EditLeftFreq
```

```
function EditLeftFreqValueChanged(app, event)
```

```
    app.LeftFreq = app.EditLeftFreq.Value;
```

```
    app.RightFreq = app.EditRightFreq.Value;
```

```
    if app.cepstris == 0
```

```
        value = app.YLogLin.Value;
```

```
        if app.SIGis == 1
```

```
            [fr,frdB,f] =
```

```
pburgfirst(app.SIG,app.SR,floor(app.order),app.LeftSample,app.RightSample,app.LeftFreq,app.RightFreq,app.DropDownType.Value);
```

```
            switch value
```

```
                case 'lin'
```

```
                    cla(app.UIAxesFreq);
```

```
                    plot(app.UIAxesFreq,f,fr),
```

```
                    [~,] =
```

```
fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
```

```
                    app.UIAxesFreq.YLim = [min(fr) max(fr)];
```

```
                case 'log'
```

```
cla(app.UIAxesFreq);
```

```
plot(app.UIAxesFreq,f,frdB),
```

```
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
```

```
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
```

```
end
```

```
end
```

```
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
```

```
end
```

```
app.SliderFreq.Limits = [app.EditLeftFreq.Value app.EditRightFreq.Value];
```

```
end
```

```
% Value changed function: EditRightFreq
```

```
function EditRightFreqValueChanged(app, event)
```

```
app.LeftFreq = app.EditLeftFreq.Value;
```

```
app.RightFreq = app.EditRightFreq.Value;
```

```
if app.cepstris == 0
```

```
    value = app.YLogLin.Value;
```

```
    if app.SIGis == 1
```

```
        [fr,frdB,f] =
```

```
pburgfirst(app.SIG,app.SR,floor(app.order),app.LeftSample,app.RightSample,app.LeftFreq,app.RightFreq,app.DropDownType.Value);
```

```
        switch value
```

```
            case 'lin'
```

```
                cla(app.UIAxesFreq);
```

```
                plot(app.UIAxesFreq,f,fr),
```

```
                [~,] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
```

```
                app.UIAxesFreq.YLim = [min(fr) max(fr)];
```

```
            case 'log'
```

```
                cla(app.UIAxesFreq);
```

```
                plot(app.UIAxesFreq,f,frdB),
```

```
                [~,] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
```

```
                app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
```

```

end
end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
end
app.SliderFreq.Limits = [app.EditLeftFreq.Value app.EditRightFreq.Value];
end

% Menu selected function: OpenMenu
function OpenMenuSelected(app, event)
[X,path] = uigetfile('*.wav','Оберіть WAV файл','G:\Мій диск\Навчання\Дисертація\');
if X ~= 0
app.path2file = [path X];
[app.SIG, app.SR] = audioread(app.path2file);
app.LeftSample = 1; app.EditLeftSample.Value = roundn(app.LeftSample/app.SR,-2);
app.RightSample = length(app.SIG); app.EditRightSample.Value =
roundn(app.RightSample/app.SR,-2);
app.tvec = (1:length(app.SIG))/app.SR;
plot(app.UIAxesTime,app.tvec(app.LeftSample:app.RightSample),normalizevec(app.SIG(app.LeftSa
mple:app.RightSample)))
app.EditRightFreq.Limits = [20 floor(app.SR/2)];
app.LeftFreq = 20; app.EditLeftFreq.Value = app.LeftFreq;
app.RightFreq = 8000; app.EditRightFreq.Value = app.RightFreq;
app.order = 50; app.SliderOrder.Value = app.order;
app.EditFieldFile.Value = app.path2file;
app.UIAxesTime.XLim = [app.LeftSample/app.SR app.RightSample/app.SR];
app.UIAxesTime.YLim = [-1 1];
app.SIGis = 1; app.cepstris = 0;
app.YLogLin.Value = 'log';
app.XLogLin.Value = 'lin'; app.UIAxesFreq.XScale = 'linear';
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFre
q,app.DropDownType.Value);
plot(app.UIAxesFreq, f,frdB)
app.SliderFreq.Limits = [app.LeftFreq app.RightFreq];
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
app.UIAxesFreq.YTickMode = 'auto';
[~] = fonems(app.DropDownLetter.Value,fr,frdB,'log',app.UIAxesFreq);
app.MPT = MPTtime(app.SIG,app.SR);
app.LabelMPT.Text = ['MPT = ' num2str(app.MPT) ' c'];
end
end

% Button pushed function: ButtonBuildFreq
function ButtonBuildFreqPushed(app, event)
if (app.SIGis == 1 && app.cepstris == 0)
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFre
q,app.DropDownType.Value);
plot(app.UIAxesFreq, f,frdB)
app.SliderFreq.Limits = [app.EditLeftFreq.Value app.EditRightFreq.Value];
% value = app.XLogLin.Value;
% [mflin,mflog,mttype] = MajorFreq (app.LeftFreq,app.RightFreq);
% switch value
% case 'lin'
% app.UIAxesFreq.XScale = 'linear';

```

```

% app.UIAxesFreq.XTickLabel = mflin;
% app.UIAxesFreq.XTick = mflin;
% case 'log'
% app.UIAxesFreq.XScale = 'log';
% app.UIAxesFreq.XTickLabel = mflog;
% app.UIAxesFreq.XTick = mflog;
% end
% if mtttype == 'auto'
% app.UIAxesFreq.XTickMode = mtttype;
% end
value = app.YLogLin.Value;
% [f1,f11,f2,f22,f3,f33] = fonems(app.DropDownLetter.Value);
switch value
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
app.UIAxesFreq.YTickMode = 'auto';
end
end

% Value changed function: SliderOrder
function SliderOrderValueChanged(app, event)
app.order = floor(app.SliderOrder.Value);
if (app.SIGis == 1 && app.cepstris == 0)
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFre
q,app.DropDownType.Value);
value = app.YLogLin.Value;
switch value
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
end
end

% Value changed function: EditLeftSample
function EditLeftSampleValueChanged(app, event)
app.LeftSample = floor(app.EditLeftSample.Value*app.SR);

```



```
end
```

```
% Value changed function: EditRightSample
```

```
function EditRightSampleValueChanged(app, event)
app.RightSample = floor(app.EditRightSample.Value*app.SR);
end
```

```
% Button pushed function: ButtonBuildTime
```

```
function ButtonBuildTimePushed(app, event)
if app.SIGis == 1
cla(app.UIAxesTime)
switch app.DropDownSig.Value
case '1'
plot(app.UIAxesTime,app.tvec(app.LeftSample:app.RightSample),normalizevec(app.SIG(app.LeftSample:app.RightSample)))
app.UIAxesTime.XLim = [app.LeftSample/app.SR app.RightSample/app.SR];
app.UIAxesTime.YLim = [-1 1];
app.MPT = MPTtime(app.SIG,app.SR);
app.LabelMPT = ['MPT = ' num2str(app.MPT) ' c'];
case '2'
[S,F,T] =
makespectrogram(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFreq);
imagesc(app.UIAxesTime,T+app.LeftSample/app.SR,F,log(1+abs(S)));
app.UIAxesTime.XLim = [app.LeftSample/app.SR app.RightSample/app.SR];
app.UIAxesTime.YLim = [app.LeftFreq app.RightFreq];
set(app.UIAxesTime,'YDir','normal');
end
end
end
```

```
% Value changed function: SliderFreq
```

```
function SliderFreqValueChanged(app, event)
if (app.SIGis == 1 && app.cepstris == 0)
value = app.YLogLin.Value;
app.SliderFreq.Limits = [app.EditLeftFreq.Value app.EditRightFreq.Value];
app.SlFreq = app.SliderFreq.Value;
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFreq,app.DropDownType.Value);
switch value
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
hold(app.UIAxesFreq,'on');
plot(app.UIAxesFreq,[app.SlFreq app.SlFreq],[min(fr) max(fr)]);
hold(app.UIAxesFreq,'off');
app.EditFieldAmpl.Value = fr(floor(app.SlFreq-app.LeftFreq+1));
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
```

```

hold(app.UIAxesFreq, 'on');
plot(app.UIAxesFreq,[app.SlFreq app.SlFreq],[min(frdB) max(frdB)]);
hold(app.UIAxesFreq, 'off');
app.EditFieldAmpl.Value = frdB(floor(app.SlFreq-app.LeftFreq+1));
end
app.EditFieldFreq.Value = app.SlFreq;
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
end
end

% Callback function
function ButtonPlayValueChanged(app, event)
if isequal(1,app.ButtonPlay.value)
a = audioplayer(app.SIG(app.LeftSample:app.RightSample),app.SR);
play(a);
end
end

% Button pushed function: ButtonPlay
function ButtonPlayPushed(app, event)
app.aplayer = audioplayer(app.SIG(app.LeftSample:app.RightSample),app.SR,16,-1);
while (1)
play(app.aplayer);
pause((app.RightSample-app.LeftSample)/app.SR);
break
end
end

% Button pushed function: ButtonStop
function ButtonStopPushed(app, event)
if app.aplayer ~= 0
stop(app.aplayer);
end
if app.arecorder ~= 0
stop(app.arecorder);
end
end

% Button pushed function: ButtonRec
function ButtonRecPushed(app, event)
app.SR=22050;
app.arecorder = audiorecorder (app.SR,16,1,1);
disp('start')
recordblocking(app.arecorder,app.EditFieldSec.Value);
disp('stop')
app.SIG = getaudiodata(app.arecorder);
app.LeftSample = 1; app.EditLeftSample.Value = roundn(app.LeftSample/app.SR,-2);
app.RightSample = length(app.SIG); app.EditRightSample.Value =
roundn(app.RightSample/app.SR,-2);
app.tvec = (1:length(app.SIG))/app.SR;
plot(app.UIAxesTime,app.tvec(app.LeftSample:app.RightSample),app.SIG(app.LeftSample:app.RightSample))
app.UIAxesTime.XLim = [app.LeftSample/app.SR app.RightSample/app.SR];
app.UIAxesTime.YLim = [min(app.SIG) max(app.SIG)];

```

```

app.EditRightFreq.Limits = [20 floor(app.SR/2)];
app.LeftFreq = 20; app.EditLeftFreq.Value = app.LeftFreq;
app.RightFreq = 8000; app.EditRightFreq.Value = app.RightFreq;
app.order = 50; app.SliderOrder.Value = app.order;
app.SIGis = 1; app.cepstris = 0;
app.SliderFreq.Limits = [app.LeftFreq app.RightFreq];
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFre
q,app.DropDownType.Value);
value = app.YLogLin.Value;
switch value
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
app.UIAxesFreq.Title.String = 'Графік спектру сигналу';
app.MPT = MPTime(app.SIG,app.SR);
app.LabelMPT = ['MPT = ' num2str(app.MPT) ' c'];
end

% Value changed function: DropDownType
function DropDownTypeValueChanged(app, event)
value = app.DropDownType.Value;
switch value
case {'pburg'}
app.LabelOrder.Text = 'Порядок';
app.SliderOrder.Limits = [2 100];
app.SliderOrder.MajorTicks = [2 10:10:100];
app.SliderOrder.Value = 50; app.order = 50;
case {'welch - bartlett','welch - gauswinn'}
app.LabelOrder.Text = 'NFFT';
app.SliderOrder.Limits = [32 1024];
app.SliderOrder.MajorTicks = 2.^(5:10);
app.SliderOrder.Value = 256; app.order = 256;
end
if (app.SIGis == 1 && app.cepstris == 0)
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFre
q,app.DropDownType.Value);
value1 = app.YLogLin.Value;
switch value1
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),

```

```

[~] = fonems(app.DropDownLetter.Value, fr, frdB, value, app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
end
end

% Button pushed function: ButtonSave
function ButtonSavePushed(app, event)
tt = clock;
defname = [[num2str(tt(1)) '.' num2str(tt(2)) '.' num2str(tt(3)) '.' num2str(tt(4)) '.'
num2str(tt(5))] '.' app.EditFieldName.Value ' ' app.DropDownLetter.Value];
[file, path] = uiputfile('*.wav', 'Оберіть шлях для збереження WAV
файлу', ['C:\Users\Пользователь\Google
Диск\Навчання\Диплом\190624_Денисенко_БакРаб\Гласные_звук' '\ defname]);
app.path2file = [path file];
if file ~= 0
audiowrite(app.path2file, app.SIG(app.LeftSample:app.RightSample), app.SR);
end
app.EditFieldFile.Value = app.path2file;
end

% Value changed function: DropDownSig
function DropDownSigValueChanged(app, event)
if app.SIGis == 1
cla(app.UIAxesTime)
switch app.DropDownSig.Value
case '1'
app.tvec = (1:length(app.SIG))/app.SR;
plot(app.UIAxesTime, app.tvec(app.LeftSample:app.RightSample), app.SIG(app.LeftSample:app.RightSample));
app.UIAxesTime.XLim = [app.LeftSample/app.SR app.RightSample/app.SR];
app.UIAxesTime.YLim = [min(app.SIG) max(app.SIG)];
app.UIAxesTime.YLabel.String = 'Amplitude';
app.UIAxesTime.Title.String = 'Графік сигналу';
case '2'
[S, F, T] =
makespectrogram(app.SIG, app.SR, app.order, app.LeftSample, app.RightSample, app.LeftFreq, app.RightFreq);
imagesc(app.UIAxesTime, T+app.LeftSample/app.SR, F, log10(1+abs(S)));
app.UIAxesTime.XLim = [app.LeftSample/app.SR app.RightSample/app.SR];
app.UIAxesTime.YLim = [app.LeftFreq app.RightFreq];
set(app.UIAxesTime, 'YDir', 'normal');
app.UIAxesTime.YLabel.String = 'Frequency(Hz)';
app.UIAxesTime.Title.String = 'Графік спектрограми';
end
end
end

% Value changed function: XLogLin
function XLogLinValueChanged(app, event)
if app.cepstris == 0
value = app.XLogLin.Value;
%[mflin,mflog] = MajorFreq (app.LeftFreq,app.RightFreq);
switch value

```

```

case 'lin'
app.UIAxesFreq.XScale = 'linear';
%app.UIAxesFreq.XTickLabel = mflin;
%app.UIAxesFreq.XTick = mflin;
case 'log'
app.UIAxesFreq.XScale = 'log';
%app.UIAxesFreq.XTickLabel = mflog;
%app.UIAxesFreq.XTick = mflog;
end
end
end

% Value changed function: YLogLin
function YLogLinValueChanged(app, event)
if (app.SIGis == 1 && app.cepstris == 0)
value = app.YLogLin.Value;
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFre
q,app.DropDownType.Value);
switch value
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
end
end

% Value changed function: DropDownSpec
function DropDownSpecValueChanged(app, event)
if app.SIGis == 1
value = app.DropDownSpec.Value;
switch value
case '1'
value1 = app.YLogLin.Value;
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFre
q,app.DropDownType.Value);
switch value1
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];

```

```

end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
app.cepstris = 0;
app.UIAxesFreq.Title.String = 'Графік спектру сигналу';
app.UIAxesFreq.XLabel.String = 'frequency(Hz)';
case '2'
[tvecc,ceps] = cepstr(app.SIG,app.SR,app.LeftSample,app.RightSample);
plot(app.UIAxesFreq,tvecc,ceps);
app.UIAxesFreq.YLim = [-0.1 0.5];
app.UIAxesFreq.XLim = [0 0.02];
app.cepstris = 1;
app.UIAxesFreq.Title.String = 'Графік кепстру сигналу';
app.UIAxesFreq.XLabel.String = 't,s';
end
end
end

% Value changed function: DropDownLetter
function DropDownLetterValueChanged(app, event)
if app.SIGis == 1
value = app.YLogLin.Value;
[fr,frdB,f] =
pburgfirst(app.SIG,app.SR,app.order,app.LeftSample,app.RightSample,app.LeftFreq,app.RightFreq,app.DropDownType.Value);
switch value
case 'lin'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,fr),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(fr) max(fr)];
case 'log'
cla(app.UIAxesFreq);
plot(app.UIAxesFreq,f,frdB),
[~] = fonems(app.DropDownLetter.Value,fr,frdB,value,app.UIAxesFreq);
app.UIAxesFreq.YLim = [min(frdB) max(frdB)];
end
end
app.UIAxesFreq.XLim = [app.LeftFreq app.RightFreq];
end

% Button pushed function: ButtonRecHigh
function ButtonRecHighPushed(app, event)
app.fs=44100;
% app.arecorderHigh = audiorecorder (app.fs,16,1,1);
% recordblocking(app.arecorderHigh,10);
% app.high = getaudiodata(app.arecorderHigh);
%
app.tseg = app.SliderTseg.Value;
app.high = audioread('G:\Мій диск\Навчання\Дисертація\SoundAHigh.wav');
[app.tvechhigh,app.vecF0high,app.f0vecEnvhigh] =
FundamentalFreqAcorrVec(app.high,app.fs,app.tseg);
plot(app.UIAxesUp,app.tvechhigh,app.vecF0high)
hold(app.UIAxesUp,"on")
plot(app.UIAxesUp,app.tvechhigh,app.f0vecEnvhigh)
hold(app.UIAxesUp,"off")
end

```

```

% Value changed function: SliderTseg
function SliderTsegValueChanged(app, event)
app.tseg = app.SliderTseg.Value;
end

% Button pushed function: ButtonRecLow
function ButtonRecLowPushed(app, event)
app.fs=44100;
% app.arecorderLow = audiorecorder (app.fs,16,1,1);
% recordblocking(app.arecorderLow,10);
% app.low = getaudiodata(app.arecorderLow);
%
app.tseg = app.SliderTseg.Value;
app.low = audioread('G:\Мій диск\Навчання\Дисертація\SoundALow.wav');
[app.tvecLow,app.vecF0Low,app.f0vecEnvLow] =
FundamentalFreqAcorrVec(app.low,app.fs,app.tseg);
plot(app.UIAxesDown,app.tvecLow,app.vecF0Low)
hold(app.UIAxesDown,"on")
plot(app.UIAxesDown,app.tvecLow,app.f0vecEnvLow)
hold(app.UIAxesDown,"off")
end

% Button pushed function: ButtonCount
function ButtonCountPushed(app, event)
[app.tvecHigh,app.vecF0High,app.f0vecEnvHigh] =
FundamentalFreqAcorrVec(app.high,app.fs,app.tseg);
[app.tvecLow,app.vecF0Low,app.f0vecEnvLow] =
FundamentalFreqAcorrVec(app.low,app.fs,app.tseg);
[app.F0Low,app.F0High,app.deltaF0,lmin,lmax] =
F0lowhigh(app.vecF0High,app.f0vecEnvHigh,app.vecF0Low,app.f0vecEnvLow);
app.LabelF0Low.Text = ['F0Low = ' num2str(floor(app.F0Low)) ' Гц'];
app.LabelF0High.Text = ['F0High = ' num2str(floor(app.F0High)) ' Гц'];
app.LabelF0Delta.Text = ['F0delta = ' num2str(floor(app.deltaF0)) ' Гц'];
hold(app.UIAxesUp,"on")
plot(app.UIAxesUp,app.tvecHigh(lmax),app.F0High,'s','LineWidth',2,'MarkerSize',10,'MarkerEdgeColor','b','MarkerFaceColor',[0.5,0.5,0.5]);
hold(app.UIAxesUp,"off")
hold(app.UIAxesDown,"on")
plot(app.UIAxesDown,app.tvecLow(lmin),app.F0Low,'s','LineWidth',2,'MarkerSize',10,'MarkerEdgeColor','b','MarkerFaceColor',[0.5,0.5,0.5]);
hold(app.UIAxesDown,"off")
plot(app.UIAxesUp,app.tvecHigh,app.vecF0High)
hold(app.UIAxesUp,"on")
plot(app.UIAxesUp,app.tvecHigh,app.f0vecEnvHigh)
hold(app.UIAxesUp,"off")
plot(app.UIAxesDown,app.tvecLow,app.vecF0Low)
hold(app.UIAxesDown,"on")
plot(app.UIAxesDown,app.tvecLow,app.f0vecEnvLow)
hold(app.UIAxesDown,"off")
end

% Button pushed function: ButtonBuildAcorr
function ButtonBuildAcorrPushed(app, event)

```

```

[app.HNRdB,Autocor,Peak,lc,N] = HNR(app.SIG(app.LeftSample:app.RightSample),app.SR);
tvecA = ((1:length(Autocor))-length(Autocor)/2)/app.SR;
plot(app.UIAxesAcor,tvecA,Autocor);
xlim(app.UIAxesAcor,[0.9*tvecA(lc(N-1)) 0.9*tvecA(lc(N+2))])
ylim(app.UIAxesAcor,[0.5*Peak(N+1) 1.2*Peak(N)])
hold(app.UIAxesAcor,"on");
plot(app.UIAxesAcor,tvecA(lc), Peak,'o');
hold(app.UIAxesAcor,"off");
app.LabelHNR.Text = ['HNRdB = ' num2str(app.HNRdB) ' dB'];
end

% Value changed function: SliderMPHup
function SliderMPHupValueChanged(app, event)
app.MPHup = app.SliderMPHup.Value;
app.MPHdown = -app.SliderMPHdown.Value;
app.perc = app.SliderPercent.Value/100;
[app.Jitter,app.JitterPercent,app.ShimmerPercent,peaksUp,nUp,peaksDown,nDown] =
ShimmerJitter(app.SIG(app.LeftSample:app.RightSample),app.SR,app.MPHdown,app.MPHup,app.perc)
;
plot(app.UIAxesTime,app.tvec(app.LeftSample:app.RightSample),normalizevec(app.SIG(app.LeftSa
mple:app.RightSample)));
hold(app.UIAxesTime,"on");
plot(app.UIAxesTime,nUp/app.SR,peaksUp,'r');
plot(app.UIAxesTime,nDown/app.SR,-peaksDown,'r');
hold(app.UIAxesTime,"off");
app.LabelJitter.Text = ['Jitter = ' num2str(app.Jitter)];
app.LabelJitterPercent.Text = ['Jitter(%) = ' num2str(app.JitterPercent) ' %'];
app.LabelShimmer.Text = ['Shimmer(%) = ' num2str(app.ShimmerPercent) ' %'];
end

% Value changed function: SliderMPHdown
function SliderMPHdownValueChanged(app, event)
app.MPHup = app.SliderMPHup.Value;
app.MPHdown = -app.SliderMPHdown.Value;
app.perc = app.SliderPercent.Value/100;
[app.Jitter,app.JitterPercent,app.ShimmerPercent,peaksUp,nUp,peaksDown,nDown] =
ShimmerJitter(app.SIG(app.LeftSample:app.RightSample),app.SR,app.MPHdown,app.MPHup,app.perc)
;
plot(app.UIAxesTime,app.tvec(app.LeftSample:app.RightSample),normalizevec(app.SIG(app.LeftSa
mple:app.RightSample)));
hold(app.UIAxesTime,"on");
plot(app.UIAxesTime,nUp/app.SR,peaksUp,'r');
plot(app.UIAxesTime,nDown/app.SR,-peaksDown,'r');
hold(app.UIAxesTime,"off");
app.LabelJitter.Text = ['Jitter = ' num2str(app.Jitter)];
app.LabelJitterPercent.Text = ['Jitter(%) = ' num2str(app.JitterPercent) ' %'];
app.LabelShimmer.Text = ['Shimmer(%) = ' num2str(app.ShimmerPercent) ' %'];
end

% Value changed function: SliderPercent
function SliderPercentValueChanged(app, event)
app.MPHup = app.SliderMPHup.Value;
app.MPHdown = -app.SliderMPHdown.Value;
app.perc = app.SliderPercent.Value/100;

```



```
[app.Jitter,app.JitterPercent,app.ShimmerPercent,peaksUp,nUp,peaksDown,nDown] =
ShimmerJitter(app.SIG(app.LeftSample:app.RightSample),app.SR,app.MPHdown,app.MPHup,app.perc)
;
plot(app.UIAxesTime,app.tvec(app.LeftSample:app.RightSample),normalizevec(app.SIG(app.LeftSa
mple:app.RightSample)));
hold(app.UIAxesTime,"on");
plot(app.UIAxesTime,nUp/app.SR,peaksUp,'r');
plot(app.UIAxesTime,nDown/app.SR,-peaksDown,'r');
hold(app.UIAxesTime,"off");
app.LabelJitter.Text = ['Jitter = ' num2str(app.Jitter)];
app.LabelJitterPercent.Text = ['Jitter(%) = ' num2str(app.JitterPercent) ' %'];
app.LabelShimmer.Text = ['Shimmer(%) = ' num2str(app.ShimmerPercent) ' %'];
end
```

**% Button pushed function: ButtonCountDSI**

```
function ButtonCountDSIPushed(app, event)
app.ILow = app.ILowBEditField.Value;
DSIndex = DSI(app.MPT,app.F0High,app.ILow,app.JitterPercent);
F0 = FundamentalFreqAcorr(app.SIG,app.SR);
app.LabelDSI.Text = ['DSI = ' num2str(DSIndex)];
app.LabelF0.Text = ['F0 = ' num2str(F0)];
end
```

**% Menu selected function: MenuMChKK**

```
function MenuMChKKSelected(app, event)
x=app.SIG; fs=app.SR;
% t=0:1/fs:(length(x)-1)/fs;
x=x/max(abs(x)); x=x-mean(x);
tseg=20e-3;
tnseg=800e-3;
N=floor(tseg*fs); N=N-mod(N,2);
M=19;
fl=50;
fh=5300;
%figure; plot(t,x)
%Крок перший: корекція сигналу з вирівнюванням енергій ВЧ та НЧ складових
s=zeros(1,length(x)-1);
for i=2:(length(x))
s(i)=x(i)-0.82*x(i-1);
end
% figure; plot(t,x), ylim([-1 1])
% w=zeros(1,N);
% for i=1:N
% w(i)=0.54+0.46*cos(2*pi*i/N);
% end
%Крок другий: розбиття сигналу на сегменти довжиною N з перекриттям N/2 та
%взважування їх вікному Хеммінга
w=hamming(N)'; w=1-w+min(w);
% figure, plot(w)
Nsize=length(s)-mod(length(s),N); %Округлення досліджуваного відрізка до кратності N
Nseg=2*Nsize/N; %кількість сегментів
sseg=zeros(Nseg-1,N);
for t=1:Nseg-1
sseg(t,:)=s(((t-1)*N/2+1):(((t-1)*N/2)+N));
end
%Крок третій: розрахунок короткочасного спектру за допомогою ШПФ для
```

```

%кожного сегменту, що аналізується
Sseg=zeros(Nseg-1,N);
for t=1:Nseg-1
% Sseg(t,:)=disFFT(sseg(t,:),w);
Sseg(t,:)=fft(sseg(t,:).*w);
end
%Крок четвертий: розрахунок набору з фільтрів трикутної форми
bfl=1127*log(1+f1/700); %менша частота досліджуваного діапазону в мелах
bfh=1127*log(1+fh/700); %більша частота досліджуваного діапазону в мелах
f=zeros(1,M);
for m=1:M
bf=bfl+m*(bfh-bfl)/(M+1); %центральна частота кожного фільтра в мелах
f(m)=floor((N/fs)*(exp(bf/1127)-1)*700); %перерахунок в Гц
end
fffh=floor((N/fs)*(exp(bfh/1127)-1)*700); %перерахунок верхньої частоти в
H=zeros(M,N);
% H(1,1:4)=[0 1 0.5 0];
for k=1:N %формування першого фільтра
if (k>=0)&&(k<=f(1))
H(1,k)=(k-1)/(f(1)-1);
elseif (k>=f(1))&&(k<=f(2))
H(1,k)=(f(2)-k)/(f(2)-f(1));
end
end
figure; plot(H(1,:)), hold on;
for m=2:M %формування фільтрів, що лишилися
for k=1:N
if (k>=f(m-1))&&(k<=f(m))
H(m,k)=(k-f(m-1))/(f(m)-f(m-1));
elseif (k>=f(m))&&(m<M)&&(k<=f(m+1))
H(m,k)=(f(m+1)-k)/(f(m+1)-f(m));
elseif (k>=f(m))&&(m==M)&&(k<=fffh)
H(m,k)=(fffh-k)/(fffh-f(m));
end
end
plot(H(m,:))
end
title('H_m(k)'), xlim([0 150])
Y=zeros(Nseg-1,M);
%Крок п'ятий: застосування набору фільтрів та розрахунок логарифму енергій
%на виході кожного фільтра
% figure;
for m=1:M
for t=1:Nseg-1
Y(t,m)=log(sum(abs(Sseg(t,:)).^2.*H(m,:)));
end
% plot(Y(:,m)), hold on
end
%Крок шостий: розрахунок мел-частотних кепстральних коефіцієнтів (МЧКК) за
%допомогою дискретно-конусного перетворення
for t=1:Nseg-1
for k=1:M
c(t,k)=sum(Y(t,:)*cos(pi*(k+1)*(m+1.5)/(M+1)));
end
end
%Крок сьомий: розрахунок кінцевих різниць МЧКК - додаткової інформаційної
%ознаки
for t=2:Nseg-1

```

```

dc(t-1,:)=c(t,:)-c(t-1,:);
end
%Крок восьмий: відбір сорока сегментів з отриманих, що відповідають
%0.02*40=0.8 с тривалість сегменту та розрахунок середніх значень та їх
%середньоквадратичних відхилень
figure;
cmu=mean(c(40:80,:)); subplot(221), plot(cmu), title('c^\mu')
dcmu=mean(dc(40:80,:)); subplot(222), plot(dcmu), title('\delta c^\mu')
cs=std(c(40:80,:)); subplot(223), plot(cs), title('c^\sigma')
dcs=std(dc(40:80,:)); subplot(224), plot(dcs), title('\delta c^\sigma')
end
end

% Component initialization
methods (Access = private)

% Create UIFigure and components
function createComponents(app)

% Create UIFigure and hide until all components are created
app.UIFigure = uifigure('Visible', 'off');
app.UIFigure.Position = [100 100 1374 837];
app.UIFigure.Name = 'Спектральний аналіз стану голосового тракту';
app.UIFigure.Scrollable = 'on';

% Create FileMenu
app.FileMenu = uimenu(app.UIFigure);
app.FileMenu.Text = 'File';

% Create OpenMenu
app.OpenMenu = uimenu(app.FileMenu);
app.OpenMenu.MenuSelectedFcn = createCallbackFcn(app, @OpenMenuSelected, true);
app.OpenMenu.Text = 'Open';

% Create MenuMChKK
app.MenuMChKK = uimenu(app.UIFigure);
app.MenuMChKK.MenuSelectedFcn = createCallbackFcn(app, @MenuMChKKSelected, true);
app.MenuMChKK.Text = 'МЧКК';

% Create UIAxesTime
app.UIAxesTime = uiaxes(app.UIFigure);
title(app.UIAxesTime, 'Графік сигналу')
xlabel(app.UIAxesTime, 'time(sec)')
ylabel(app.UIAxesTime, 'Amplitude')
app.UIAxesTime.XTick = [];
app.UIAxesTime.XTickLabel = '';
app.UIAxesTime.YTick = [-1 -0.9 -0.8 -0.7 -0.6 -0.5 -0.4 -0.3 -0.2 -0.1 0 0.1 0.2 0.3 0.4
0.5 0.6 0.7 0.8 0.9 1];
app.UIAxesTime.YTickLabel = {'-1'; '-0.9'; '-0.8'; '-0.7'; '-0.6'; '-0.5'; '-0.4'; '-0.3';
'-0.2'; '-0.1'; '0'; '0.1'; '0.2'; '0.3'; '0.4'; '0.5'; '0.6'; '0.7'; '0.8'; '0.9'; '1'};
app.UIAxesTime.Position = [54 437 585 355];

```

```

% Create UIAxesFreq
app.UIAxesFreq = uiaxes(app.UIFigure);
title(app.UIAxesFreq, 'Графік спектру сигналу')
xlabel(app.UIAxesFreq, 'frequency(Hz)')
ylabel(app.UIAxesFreq, 'dB/Hz')
app.UIAxesFreq.Position = [644 436 637 356];

% Create UIAxesAcor
app.UIAxesAcor = uiaxes(app.UIFigure);
title(app.UIAxesAcor, 'Автокореляційна функція')
xlabel(app.UIAxesAcor, 't,c')
zlabel(app.UIAxesAcor, 'Z')
app.UIAxesAcor.Position = [54 15 385 349];

% Create UIAxesUp
app.UIAxesUp = uiaxes(app.UIFigure);
title(app.UIAxesUp, 'ЧОТ, що зростає')
xlabel(app.UIAxesUp, 't,c')
ylabel(app.UIAxesUp, 'F0, Гц')
zlabel(app.UIAxesUp, 'Z')
app.UIAxesUp.Position = [648 46 324 318];

% Create UIAxesDown
app.UIAxesDown = uiaxes(app.UIFigure);
title(app.UIAxesDown, 'ЧОТ, що спадає')
xlabel(app.UIAxesDown, 't,c')
zlabel(app.UIAxesDown, 'Z')
app.UIAxesDown.Position = [956 46 324 318];

% Create LabelLeftFreq
app.LabelLeftFreq = uilabel(app.UIFigure);
app.LabelLeftFreq.HorizontalAlignment = 'right';
app.LabelLeftFreq.Position = [789 407 62 22];
app.LabelLeftFreq.Text = 'Ліва межа';

% Create EditLeftFreq
app.EditLeftFreq = uieditfield(app.UIFigure, 'numeric');
app.EditLeftFreq.ValueChangedFcn = createCallbackFcn(app, @EditLeftFreqValueChanged, true);
app.EditLeftFreq.Position = [858 405 34 22];
app.EditLeftFreq.Value = 20;

% Create LabelRightFreq
app.LabelRightFreq = uilabel(app.UIFigure);
app.LabelRightFreq.HorizontalAlignment = 'right';
app.LabelRightFreq.Position = [891 407 73 22];
app.LabelRightFreq.Text = 'Права межа';

% Create EditRightFreq
app.EditRightFreq = uieditfield(app.UIFigure, 'numeric');

```

```

app.EditRightFreq.ValueChangedFcn = createCallbackFcn(app, @EditRightFreqValueChanged,
true);
app.EditRightFreq.Position = [971 405 52 22];
app.EditRightFreq.Value = 2000;

% Create ButtonBuildFreq
app.ButtonBuildFreq = uibutton(app.UIFigure, 'push');
app.ButtonBuildFreq.ButtonPushedFcn = createCallbackFcn(app, @ButtonBuildFreqPushed, true);
app.ButtonBuildFreq.Position = [1260 405 100 22];
app.ButtonBuildFreq.Text = 'Побудова';

% Create LabelOrder
app.LabelOrder = uilabel(app.UIFigure);
app.LabelOrder.HorizontalAlignment = 'right';
app.LabelOrder.Position = [1281 428 52 22];
app.LabelOrder.Text = 'Порядок';

% Create SliderOrder
app.SliderOrder = uislider(app.UIFigure);
app.SliderOrder.Limits = [1 100];
app.SliderOrder.MajorTicks = [2 10 20 30 40 50 60 70 80 90 100];
app.SliderOrder.Orientation = 'vertical';
app.SliderOrder.ValueChangedFcn = createCallbackFcn(app, @SliderOrderValueChanged, true);
app.SliderOrder.Position = [1304 473 3 319];
app.SliderOrder.Value = 50;

% Create LabelLeftSample
app.LabelLeftSample = uilabel(app.UIFigure);
app.LabelLeftSample.HorizontalAlignment = 'right';
app.LabelLeftSample.Position = [377 407 62 22];
app.LabelLeftSample.Text = 'Ліва межа';

% Create EditLeftSample
app.EditLeftSample = uieditfield(app.UIFigure, 'numeric');
app.EditLeftSample.ValueChangedFcn = createCallbackFcn(app, @EditLeftSampleValueChanged,
true);
app.EditLeftSample.Position = [438 406 30 22];

% Create LabelRightSample
app.LabelRightSample = uilabel(app.UIFigure);
app.LabelRightSample.HorizontalAlignment = 'right';
app.LabelRightSample.Position = [467 407 73 22];
app.LabelRightSample.Text = 'Права межа';

% Create EditRightSample
app.EditRightSample = uieditfield(app.UIFigure, 'numeric');
app.EditRightSample.ValueChangedFcn = createCallbackFcn(app, @EditRightSampleValueChanged,
true);
app.EditRightSample.Position = [539 406 30 22];

```

```

% Create ButtonBuildTime
app.ButtonBuildTime = uibutton(app.UIFigure, 'push');
app.ButtonBuildTime.ButtonPushedFcn = createCallbackFcn(app, @ButtonBuildTimePushed, true);
app.ButtonBuildTime.Position = [573 406 67 22];
app.ButtonBuildTime.Text = 'Побудова';

% Create SliderFreq
app.SliderFreq = uislider(app.UIFigure);
app.SliderFreq.Limits = [20 2000];
app.SliderFreq.ValueChangedFcn = createCallbackFcn(app, @SliderFreqValueChanged, true);
app.SliderFreq.FontSize = 9;
app.SliderFreq.Position = [683 394 578 3];
app.SliderFreq.Value = 20;

% Create Label_6
app.Label_6 = uilabel(app.UIFigure);
app.Label_6.HorizontalAlignment = 'right';
app.Label_6.Position = [1130 405 68 22];
app.Label_6.Text = 'Амплітуда';

% Create EditFieldAmpl
app.EditFieldAmpl = uieditfield(app.UIFigure, 'numeric');
app.EditFieldAmpl.Eitable = 'off';
app.EditFieldAmpl.Position = [1205 405 44 22];

% Create LabelFile
app.LabelFile = uilabel(app.UIFigure);
app.LabelFile.HorizontalAlignment = 'right';
app.LabelFile.Position = [4 375 35 22];
app.LabelFile.Text = 'Файл';

% Create EditFieldFile
app.EditFieldFile = uieditfield(app.UIFigure, 'text');
app.EditFieldFile.Eitable = 'off';
app.EditFieldFile.FontSize = 10;
app.EditFieldFile.Position = [54 375 586 22];

% Create ButtonPlay
app.ButtonPlay = uibutton(app.UIFigure, 'push');
app.ButtonPlay.ButtonPushedFcn = createCallbackFcn(app, @ButtonPlayPushed, true);
app.ButtonPlay.Position = [11 407 67 22];
app.ButtonPlay.Text = 'Програти';

% Create ButtonStop
app.ButtonStop = uibutton(app.UIFigure, 'push');
app.ButtonStop.ButtonPushedFcn = createCallbackFcn(app, @ButtonStopPushed, true);
app.ButtonStop.Position = [77 407 67 22];
app.ButtonStop.Text = 'Зупинити';

% Create ButtonRec

```

```

app.ButtonRec = uibutton(app.UIFigure, 'push');
app.ButtonRec.ButtonPushedFcn = createCallbackFcn(app, @ButtonRecPushed, true);
app.ButtonRec.Position = [143 407 67 22];
app.ButtonRec.Text = 'Записати';

% Create DropDownType
app.DropDownType = uidropdown(app.UIFigure);
app.DropDownType.Items = {'pburg', 'welch - bartlett', 'welch - gausswin'};
app.DropDownType.ValueChangedFcn = createCallbackFcn(app, @DropDownTypeValueChanged, true);
app.DropDownType.Position = [683 405 100 22];
app.DropDownType.Value = 'pburg';

% Create Label_8
app.Label_8 = uilabel(app.UIFigure);
app.Label_8.HorizontalAlignment = 'right';
app.Label_8.Position = [280 408 66 22];
app.Label_8.Text = 'Тривалість';

% Create EditFieldSec
app.EditFieldSec = uieditfield(app.UIFigure, 'numeric');
app.EditFieldSec.Position = [349 407 30 22];
app.EditFieldSec.Value = 5;

% Create Label_9
app.Label_9 = uilabel(app.UIFigure);
app.Label_9.HorizontalAlignment = 'right';
app.Label_9.Position = [1022 405 50 22];
app.Label_9.Text = 'Частота';

% Create EditFieldFreq
app.EditFieldFreq = uieditfield(app.UIFigure, 'numeric');
app.EditFieldFreq.Editable = 'off';
app.EditFieldFreq.Position = [1079 405 52 22];

% Create ButtonSave
app.ButtonSave = uibutton(app.UIFigure, 'push');
app.ButtonSave.ButtonPushedFcn = createCallbackFcn(app, @ButtonSavePushed, true);
app.ButtonSave.Position = [209 407 67 22];
app.ButtonSave.Text = 'Зберегти';

% Create Label_10
app.Label_10 = uilabel(app.UIFigure);
app.Label_10.HorizontalAlignment = 'right';
app.Label_10.Position = [16 802 26 22];
app.Label_10.Text = 'Ім'я';

% Create EditFieldName
app.EditFieldName = uieditfield(app.UIFigure, 'text');
app.EditFieldName.Position = [74 802 202 22];

```

```

% Create DropDownSpec
app.DropDownSpec = uidropdown(app.UIFigure);
app.DropDownSpec.Items = {'Спектр', 'Кенстр'};
app.DropDownSpec.ItemsData = {'1', '2'};
app.DropDownSpec.ValueChangedFcn = createCallbackFcn(app, @DropDownSpecValueChanged, true);
app.DropDownSpec.Position = [682 802 100 22];
app.DropDownSpec.Value = '1';

% Create DropDownSig
app.DropDownSig = uidropdown(app.UIFigure);
app.DropDownSig.Items = {'Сигнал', 'Спектрограма'};
app.DropDownSig.ItemsData = {'1', '2'};
app.DropDownSig.ValueChangedFcn = createCallbackFcn(app, @DropDownSigValueChanged, true);
app.DropDownSig.Position = [539 802 100 22];
app.DropDownSig.Value = '1';

% Create XLogLin
app.XLogLin = uidropdown(app.UIFigure);
app.XLogLin.Items = {'lin', 'log'};
app.XLogLin.ValueChangedFcn = createCallbackFcn(app, @XLogLinValueChanged, true);
app.XLogLin.Position = [1269 375 39 22];
app.XLogLin.Value = 'lin';

% Create YLogLin
app.YLogLin = uidropdown(app.UIFigure);
app.YLogLin.Items = {'lin', 'log'};
app.YLogLin.ValueChangedFcn = createCallbackFcn(app, @YLogLinValueChanged, true);
app.YLogLin.Position = [1307 375 39 22];
app.YLogLin.Value = 'lin';

% Create DropDownLetter
app.DropDownLetter = uidropdown(app.UIFigure);
app.DropDownLetter.Items = {'A', 'E', 'H', 'I', 'O', 'Y'};
app.DropDownLetter.ValueChangedFcn = createCallbackFcn(app, @DropDownLetterValueChanged, true);
app.DropDownLetter.Position = [285 802 50 22];
app.DropDownLetter.Value = 'A';

% Create SliderMPHup
app.SliderMPHup = uislider(app.UIFigure);
app.SliderMPHup.Limits = [0 0.99];
app.SliderMPHup.MajorTicks = [0 0.25 0.5 0.75 1];
app.SliderMPHup.Orientation = 'vertical';
app.SliderMPHup.ValueChangedFcn = createCallbackFcn(app, @SliderMPHupValueChanged, true);
app.SliderMPHup.Position = [4 618 7 155];
app.SliderMPHup.Value = 0.3;

% Create SliderMPHdown
app.SliderMPHdown = uislider(app.UIFigure);
app.SliderMPHdown.Limits = [-0.99 0];
app.SliderMPHdown.MajorTicks = [-1 -0.75 -0.5 -0.25 0];
app.SliderMPHdown.Orientation = 'vertical';

```



```

app.SliderMPHdown.ValueChangedFcn = createCallbackFcn(app, @SliderMPHdownValueChanged,
true);
app.SliderMPHdown.Position = [4 464 7 155];
app.SliderMPHdown.Value = -0.25;

% Create ButtonRecHigh
app.ButtonRecHigh = uibutton(app.UIFigure, 'push');
app.ButtonRecHigh.ButtonPushedFcn = createCallbackFcn(app, @ButtonRecHighPushed, true);
app.ButtonRecHigh.Position = [682 22 136 24];
app.ButtonRecHigh.Text = 'Записати зростаючий';

% Create ButtonRecLow
app.ButtonRecLow = uibutton(app.UIFigure, 'push');
app.ButtonRecLow.ButtonPushedFcn = createCallbackFcn(app, @ButtonRecLowPushed, true);
app.ButtonRecLow.Position = [996 22 113 24];
app.ButtonRecLow.Text = 'Записати спадний';

% Create LabelF0High
app.LabelF0High = uilabel(app.UIFigure);
app.LabelF0High.Position = [817 22 73 22];
app.LabelF0High.Text = '';

% Create LabelF0Low
app.LabelF0Low = uilabel(app.UIFigure);
app.LabelF0Low.Position = [1115 22 91 22];
app.LabelF0Low.Text = '';

% Create ButtonCount
app.ButtonCount = uibutton(app.UIFigure, 'push');
app.ButtonCount.ButtonPushedFcn = createCallbackFcn(app, @ButtonCountPushed, true);
app.ButtonCount.Position = [905 22 80 24];
app.ButtonCount.Text = 'Вирахувати';

% Create LabelF0Delta
app.LabelF0Delta = uilabel(app.UIFigure);
app.LabelF0Delta.Position = [1210 23 123 20];
app.LabelF0Delta.Text = '';

% Create Label_11
app.Label_11 = uilabel(app.UIFigure);
app.Label_11.Position = [1286 199 66 28];
app.Label_11.Text = {'Тривалість'; 'сегменту'};

% Create SliderTseg
app.SliderTseg = uislidder(app.UIFigure);
app.SliderTseg.Limits = [0.03 0.15];
app.SliderTseg.Orientation = 'vertical';
app.SliderTseg.ValueChangedFcn = createCallbackFcn(app, @SliderTsegValueChanged, true);
app.SliderTseg.Position = [1292 235 7 106];
app.SliderTseg.Value = 0.05;

```

```

% Create ButtonBuildAcorr
app.ButtonBuildAcorr = uibutton(app.UIFigure, 'push');
app.ButtonBuildAcorr.ButtonPushedFcn = createCallbackFcn(app, @ButtonBuildAcorrPushed,
true);
app.ButtonBuildAcorr.Position = [454 319 100 22];
app.ButtonBuildAcorr.Text = 'Побудова';

% Create LabelHNR
app.LabelHNR = uilabel(app.UIFigure);
app.LabelHNR.Position = [438 290 159 22];
app.LabelHNR.Text = '';

% Create LabelMPT
app.LabelMPT = uilabel(app.UIFigure);
app.LabelMPT.Position = [438 252 159 28];
app.LabelMPT.Text = '';

% Create Label_12
app.Label_12 = uilabel(app.UIFigure);
app.Label_12.HorizontalAlignment = 'right';
app.Label_12.Position = [557 25 100 22];
app.Label_12.Text = 'Відсоток від ЧОТ';

% Create SliderPercent
app.SliderPercent = uislider(app.UIFigure);
app.SliderPercent.Orientation = 'vertical';
app.SliderPercent.ValueChangedFcn = createCallbackFcn(app, @SliderPercentValueChanged,
true);
app.SliderPercent.Position = [596 64 7 283];
app.SliderPercent.Value = 100;

% Create LabelJitter
app.LabelJitter = uilabel(app.UIFigure);
app.LabelJitter.Position = [438 219 159 28];
app.LabelJitter.Text = '';

% Create LabelJitterPercent
app.LabelJitterPercent = uilabel(app.UIFigure);
app.LabelJitterPercent.Position = [438 183 159 25];
app.LabelJitterPercent.Text = '';

% Create LabelShimmer
app.LabelShimmer = uilabel(app.UIFigure);
app.LabelShimmer.Position = [438 149 159 26];
app.LabelShimmer.Text = '';

% Create LabelF0
app.LabelF0 = uilabel(app.UIFigure);

```

```

app.LabelF0.Position = [438 114 159 26];
app.LabelF0.Text = '';

% Create LabelDSI
app.LabelDSI = uilabel(app.UIFigure);
app.LabelDSI.Position = [438 81 159 26];
app.LabelDSI.Text = '';

% Create ButtonCountDSI
app.ButtonCountDSI = uibutton(app.UIFigure, 'push');
app.ButtonCountDSI.ButtonPushedFcn = createCallbackFcn(app, @ButtonCountDSIPushed, true);
app.ButtonCountDSI.Position = [440 43 100 22];
app.ButtonCountDSI.Text = 'Вирахувати';

% Create IlowdBEditFieldLabel
app.IlowdBEditFieldLabel = uilabel(app.UIFigure);
app.IlowdBEditFieldLabel.HorizontalAlignment = 'right';
app.IlowdBEditFieldLabel.Position = [442 4 50 22];
app.IlowdBEditFieldLabel.Text = 'Ilow(dB)';

% Create IlowdBEditField
app.IlowdBEditField = uieditfield(app.UIFigure, 'numeric');
app.IlowdBEditField.Position = [507 4 30 22];
app.IlowdBEditField.Value = 60;

% Show the figure after all components are created
app.UIFigure.Visible = 'on';
end
end

% App creation and deletion
methods (Access = public)

% Construct app
function app = FirstAttempt2

% Create UIFigure and components
createComponents(app)

% Register the app with App Designer
registerApp(app, app.UIFigure)

if nargin == 0
clear app
end
end

% Code that executes before app deletion

```

```

function delete(app)

% Delete UIFigure when app is deleted
delete(app.UIFigure)
end
end
end

function DSI = DSI(MPT,F0High,ILow,JitterPercent)
DSI=0.13*MPT+0.0053*F0High+0.26*ILow-1.18*JitterPercent+12.4;
end

function [MPT] = MPTime(vec,fs)
vec=vec-mean(vec); vec=vec/abs(max(vec));
env=envelope(vec,150,'rms');
envdB=10*log10(env/max(env));
mpn=0;
for i=1:length(envdB)
    if envdB(i) > -6
        mpn=mpn+1;
    end
end
MPT=mpn/fs;
end

function [tvec,vecF0,f0vecEnv] = FundamentalFreqAcorrVec(vec,fs,tseg)
N=floor(tseg*fs); %довжина сегменту
Nseg=floor(length(vec)/N); %кількість сегментів
vecF0=zeros(1,Nseg);
for t=1:Nseg
    seg=vec(((t-1)*N+1):(t*N));%розбиття на сегменти
    vecF0(t)=FundamentalFreqAcorr(seg,fs); %розрахунок частоти основного тону
    if isnan(vecF0(t)) == 1
        vecF0(t)=vecF0(t-1);
    end
end
[up,lo]=envelope(vecF0,5,'Peak');
f0vecEnv=(up+lo)/2;
tvec=tseg*(0:Nseg-1);
end

function F0 = FundamentalFreqAcorr(vec,fs)
[corr,~]=xcorr(vec,vec); %автокореляційна функція
[~,lc]=max(corr); corr(lc)=0;
[maxc,~]=max(corr); corr(lc)=maxc;
[~,lcm] = findpeaks(corr,'MinPeakHeight',0.5*max(corr));
lcmd=zeros(1,length(lcm)-1);
for i=2:length(lcm)
    lcmd(i-1)=lcm(i)-lcm(i-1);
end
F0=fs/mean(lcmd);%розрахунок частоти основного тону як обернене до
% середнього значення відстані між максимумами кореляційної функції
end

```

```

function [Jitter,JitterPercent,ShimmerPercent,peaksUp,nUp,peaksDown,nDown] =
ShimmerJitter(vec,fs,MPHD,MPHU,perc)
vec=vec-mean(vec); vec=vec/max(abs(vec));
F0 = FundamentalFreqAcorr(vec,fs);
F0=floor(F0);

[peaksUp,nUp] = findpeaks(vec,'MinPeakHeight',MPHU,'MinPeakDistance',perc*(fs/F0));
[peaksDown,nDown] = findpeaks(-vec,'MinPeakHeight',MPHD,'MinPeakDistance',perc*(fs/F0));

N=min(length(nUp),length(nDown));
Tup(1) = nUp(2)-nUp(1);
Tdown(1) = nDown(2)-nDown(1);
Aup(1) = peaksUp(2)-peaksUp(1);
Adown(1) = peaksDown(2)-peaksDown(1);
for i=2:N-1
    Tup(i) = nUp(i+1)-nUp(i);
    Tdown(i) = nDown(i+1)-nDown(i);
    Tdifdown(i-1) = abs(Tdown(i)-Tdown(i-1));
    Tdifup(i-1) = abs(Tup(i)-Tup(i-1));
    Adifup(i) = abs(peaksUp(i)-peaksUp(i-1));
    Adifdown(i) = abs(peaksDown(i)-peaksDown(i-1));
end

Tmean = mean([Tup Tdown]);
Tdiffmean = mean([Tdifdown Tdifup]);
Amean = mean([peaksUp' peaksDown']);
Adiffmean = mean([Adifdown Adifup]);
Jitter=Tdiffmean/fs;
JitterPercent = 100*Tdiffmean/Tmean;
ShimmerPercent = 100*Adiffmean/Amean;
end

```

```

function [HNRdB,Autocor,Peak,lc,N] = HNR(vec,fs)
vec=vec-mean(vec); vec=vec/max(abs(vec));
F0 = FundamentalFreqAcorr(vec,fs);
F0=floor(F0);

Autocor=xcorr(vec,vec);
[Peak,lc] = findpeaks (Autocor,'MinPeakDistance',0.5*(fs/F0));

[MaxPeak,N] = max(Peak);
HNRdB = 10*log10(Peak(N+1)/(MaxPeak-Peak(N+1)));
end

```

```

function [mtlin,mtlog,mttype] = MajorFreq(leftfreq, rightfreq)
%%
if (leftfreq < 1000) && (rightfreq <= 1000)
    mtlog = [20 50 100 200 400 800 1000];
    mtlin = [20 100 400 500 600 700 800 900 1000];
end
if (leftfreq < 1000) && (rightfreq <= 2000) && (rightfreq > 1000)
    mtlog = [20 100 200 400 800 1000 2000];
    mtlin = [20 100 400 500 600 800 1000 2000];
end
if (leftfreq < 1000) && (rightfreq <= 4000) && (rightfreq > 2000)
    mtlog = [20 100 200 400 800 2000 4000];
    mtlin = [20 400 600 800 1000 2000 4000];
end
if (leftfreq < 1000) && (rightfreq <= 8000) && (rightfreq > 4000)

```

```

        mtlog = [20 100 200 800 2000 4000 8000];
        mtlin = [20 400 1000 2000 4000 6000 8000];
    end
    if (leftfreq < 1000) && (rightfreq <= 11000) && (rightfreq > 8000)
        mtlog = [20 100 200 800 2000 4000 8000 11000];
        mtlin = [20 400 1000 2000 4000 8000 11000];
    end
    %%
    if (leftfreq >= 1000) && (rightfreq <= 2000) && (rightfreq > 1000)
        mtlog = [1000 1100 1300 1600 1800 2000];
        mtlin = [1000 1200 1400 1600 1800 2000];
    end
    if (leftfreq >= 1000) && (rightfreq <= 4000) && (rightfreq > 2000)
        mtlog = [1000 1200 1600 2000 3000 4000];
        mtlin = [1000 2000 2500 3000 3500 4000];
    end
    if (leftfreq >= 1000) && (rightfreq <= 8000) && (rightfreq > 4000)
        mtlog = [1000 2000 3000 4000 6000 8000];
        mtlin = [1000 2000 4000 6000 7000 8000];
    end
    if (leftfreq >= 1000) && (rightfreq <= 11000) && (rightfreq > 8000)
        mtlog = [1000 2000 4000 8000 10000 11000];
        mtlin = [1000 2000 3000 5000 8000 11000];
    end
    %%
    if rightfreq-leftfreq < 1000
        mttype = 'auto';
    else
        mttype = 0;
    end
end
end

```

```

function [s,f,t] = makespectrogram(SIG,SR,order,LeftSample,RightSample,LeftFreq,RightFreq)
win1 = bartlett(order);
[s,f,t] =
spectrogram(SIG(LeftSample:RightSample),win1,floor(order/10),LeftFreq:RightFreq,SR,'yaxis');
end

```

```

function [fr,frdB,f] =
pburgfirst(SIGtall,SR,order,LeftSample,RightSample,LeftFreq,RightFreq,type)
SIGshort=SIGtall(LeftSample:RightSample); SIGshort=SIGshort-mean(SIGshort);
SIGshort=SIGshort/max(abs(SIGshort));
[SIGenv,~]=envelope(SIGshort,100,'rms');
ii=1; SIGnew = 1;
for i=1:length(SIGshort)
    if SIGenv(i)>0.1
        SIGnew(ii)=SIGshort(i);
        ii=ii+1;
    end
end
switch type
case 'pburg'
    [fr,f]=pburg(SIGnew,order,LeftFreq:RightFreq,SR);
    fr=fr/max(fr);
    frdB=10*log10(fr);
case 'welch - gausswin'
    win = gausswin(order);
    [fr,f] = pwelch(SIGnew,win,floor(order/10),[LeftFreq:RightFreq],SR);
end

```

```

        fr=fr/max(fr);
        frdB = 10*log10(fr);
    case 'welch - bartlett'
        win = bartlett(order);
        [fr,f] = pwelch(SIGnew,win,floor(order/10),[LeftFreq:RightFreq],SR);
        fr=fr/max(fr);
        frdB = 10*log10(fr);
end
end

function [vecout] = normalizevec(vec)
vec=vec-mean(vec); vecout=vec/(max(abs(vec)));
end

function [a] = fonems(letter,fr,frdB,type,UIAxesFreq)
switch letter
    case 'A'
        f1 = 700-30; f11 = 700+30;    f2 = 1080-36;    f22 = 1080+36;    f3 = 2060-65;    f33 =
2060+65;
    case 'E'
        f1 = 440-20;  f11 = 440+20;    f2 = 1800-48;  f22 = 1800+48;  f3 = 2550-65;  f33 =
2550+65;
    case 'U'
        f1 = 300-22;  f11 = 300+22;    f2 = 1480-60;  f22 = 1480+60;  f3 = 2330-40;  f33 =
2330+40;
    case 'I'
        f1 = 240-30;  f11 = 240+30;    f2 = 2250-38;  f22 = 2250+38;  f3 = 3200-120;  f33 =
3200+120;
    case 'O'
        f1 = 535-27;  f11 = 535+27;    f2 = 780-32;    f22 = 780+32;    f3 = 2500-50;    f33 =
2500+50;
    case 'Y'
        f1 = 300-35;  f11 = 300+35;    f2 = 625-25;    f22 = 625+25;    f3 = 2500-55;    f33 =
2500+55;
end
hold(UIAxesFreq,'on');
switch type
    case 'lin'
        plot(UIAxesFreq, [f1 f11],[max(fr)-0.05 max(fr)-0.05], 'r'), plot(UIAxesFreq, [f1
f1],[max(fr)-0.05 min(fr)+0.05], 'r'), plot(UIAxesFreq, [f11 f11],[max(fr)-0.05
min(fr)+0.05], 'r'), plot(UIAxesFreq, [f1 f11],[min(fr)+0.05 min(fr)+0.05], 'r'),
        text(UIAxesFreq, f1+1, max(fr)*0.8, 'F1');
        plot(UIAxesFreq, [f2 f22],[max(fr)-0.05 max(fr)-0.05], 'r'), plot(UIAxesFreq, [f2
f2],[max(fr)-0.05 min(fr)+0.05], 'r'), plot(UIAxesFreq, [f22 f22],[max(fr)-0.05
min(fr)+0.05], 'r'), plot(UIAxesFreq, [f2 f22],[min(fr)+0.05 min(fr)+0.05], 'r'),
        text(UIAxesFreq, f2+1, max(fr)*0.8, 'F2');
        plot(UIAxesFreq, [f3 f33],[max(fr)-0.05 max(fr)-0.05], 'r'), plot(UIAxesFreq, [f3
f3],[max(fr)-0.05 min(fr)+0.05], 'r'), plot(UIAxesFreq, [f33 f33],[max(fr)-0.05
min(fr)+0.05], 'r'), plot(UIAxesFreq, [f3 f33],[min(fr)+0.05 min(fr)+0.05], 'r'),
        text(UIAxesFreq, f3+1, max(fr)*0.8, 'F3');
    case 'log'
        plot(UIAxesFreq, [f1 f11],[max(frdB)-1 max(frdB)-1], 'r'), plot(UIAxesFreq, [f1
f1],[max(frdB)-1 min(frdB)+1], 'r'), plot(UIAxesFreq, [f11 f11],[max(frdB)-1
min(frdB)+1], 'r'), plot(UIAxesFreq, [f1 f11],[min(frdB)+1 min(frdB)+1], 'r'),
        text(UIAxesFreq, f1+1, max(frdB)-10, 'F1');
        plot(UIAxesFreq, [f2 f22],[max(frdB)-1 max(frdB)-1], 'r'), plot(UIAxesFreq, [f2
f2],[max(frdB)-1 min(frdB)+1], 'r'), plot(UIAxesFreq, [f22 f22],[max(frdB)-1
min(frdB)+1], 'r'), plot(UIAxesFreq, [f2 f22],[min(frdB)+1 min(frdB)+1], 'r'),
        text(UIAxesFreq, f2+1, max(frdB)-10, 'F2');

```

```

        plot(UIAxesFreq, [f3 f33],[max(frdB)-1 max(frdB)-1], 'r'), plot(UIAxesFreq, [f3
f3],[max(frdB)-1 min(frdB)+1], 'r'), plot(UIAxesFreq, [f33 f33],[max(frdB)-1
min(frdB)+1], 'r'), plot(UIAxesFreq, [f3 f33],[min(frdB)+1 min(frdB)+1], 'r'),
        text(UIAxesFreq, f3+1, max(frdB)-10, 'F3');
end
hold(UIAxesFreq, 'off');
a=1;
end

```