# Порівняння методів обчислення тремтіння голосу для україномовних дикторів

*Джичка Н. І. аспірант*

Кафедра біотехнічних систем
Тернопільський національний технічний університет імені Івана Пулюя
magnanimous@ukr.net

## Abstract

This article is devoted to the problem of jitter estimation of Ukrainian speakers' voices. The most popular method of jitter estimation and its modification were used in the experiment to compare voices of two groups of Ukrainian speakers: with normal voices and with hoarseness on the other hand.

## 1. Вступ

Проблема вчасного діагностування раптового пошкодження голосу особливо важлива для людей голосомовних професій: лекторів, дикторів радіо та телебачення, співаків, вчителів, телефоністів, диспетчерів, акторів, оскільки характер їхньої роботи спричиняє значні навантаження голосового апарату. Постійні ненормовані навантаження на голос викликають утворення поліпів на голосових зв‘язках, розвиток яких призводить до ускладнення змикання голосової щілини, дисфонії, афонії і втрати працездатності. Будь-які порушення структури голосових зв‘язок впливають на характер їхньої вібрації, що впливає на якість голосу і відображається в голосовому сигналі. Основним симптомом, що свідчить про пошкодження тканин голосових зв‘язок, є хриплість голосу. Стадія розвитку поліпів визначає ступінь важкості захворювання.

Всі професійні захворювання голосу поділяють на три групи. До першої групи відноситься фонастенія – порушення голосу без очевидних змін в гортані, котра проявляється тільки суб‘єктивними відчуттями пацієнта. До другої групи професійних захворювань належать слабо виражені порушення в тканинах голосових зв‘язок – стійке почервоніння країв голосових зв‘язок, розширення кровоносних судин на їхній поверхні. До третьої групи професійних захворювань відносять захворювання із значними змінами в тканинах голосових зв‘язок: «вузлики співаків», фіброми голосових зв‘язок.

Вчасне діагностування захворювання другої групи дозволяє запобігти серйознішим порушенням в тканинах голосових зв‘язок та швидше відновити працездатність пацієнта.

Діагностування захворювань гортані та її внутрішніх органів здійснюється поетапно і включає в себе фізичну оцінку голосу, неінвазивну оцінку голосу (в тому числі і акустичну оцінку записів голосу за допомогою спеціалізованих комп‘ютерних систем), інвазивні тести та фотографування.

Перші два етапи є поверховим обстеженням, проте акустична оцінка записів дозволяє оцінити величину хриплості голосу та виявити порушення, невидимі при дзеркальному огляді. При цьому не потребуються подальші інвазивні тести, які можуть бути неприємними та небезпечними для пацієнта.

Розвиток комп‘ютерних технологій та вдосконалення методів акустичної оцінки записів звуків дозволяє об‘єктивно оцінювати хриплість голосу та коректувати діагноз.

Найбільш відомими серед систем оцінювання звукових записів є Multidimensional Voice Program (MDVP) (виробник KayPENTAX, MEEI, USA) та lingWAVES Voice Clinic Suite Pro (WEVOSYS, Germany). Ці програмні продукти використовуються лише в країнах-виробниках і орієнтовані на особливості вимови звуків американської, англійської та німецької мов.

Тому важливою задачею для акустичної оцінки записів голосу в Україні є розроблення відповідного програмного забезпечення, орієнтованого на особливості вимови українських звуків, для неінвазивного діагностування другої групи пошкодження голосових зв‘язок.

## 2. Методи вимірювання хриплості

Об‘єктивним параметром вимірювання хриплості є тремтіння – варіація тривалості періоду основного тону голосового сигналу [1]. Найбільш поширений метод вимірювання тремтіння – Jitter (Jitt), який використовується в системах аналізу голосу lingWAVES Voice Clinic Suite Pro та Multidimensional Voice Program (MDVP) [2, 3]. Цей метод допускає періодичність сигналу протягом тривалого часу і характеризує весь сигнал одним числовим значенням.

Пропонуються два підходи до визначення тремтіння голосу:

– Тремтіння – це звичайна варіація тривалості періоду основного тону, яка вимірюється відніманням кожного періоду з послідовності періодів основного тону від сусіднього періоду або комбінації сусідніх періодів.

$$Jitta = \frac{1}{N-1}\sum_{k=1}^{N-1}\left|P_0(n+1) - P_0(n)\right|, \qquad (1)$$

тут $P_0(n)$ – послідовність тривалості періодів основного тону (мкс).

– Тремтіння – відношення середньої абсолютної різниці між сусідніми періодами (з послідовності періодів) до середнього періоду (%)

$$Jitt = 100\frac{(1/(N-1))\sum_{k=1}^{N-1}\left|P_0(n+1) - P_0(n)\right|}{(1/N)\sum_{k=1}^{N}P_0(n)}. \qquad (2)$$

Дарсіо Дж. Сілва, Луіс С. Олівейра та Маріо Андреа запропонували обчислювати локальне тремтіння для вікна

аналізу, що включає чотири послідовні періоди основного тону, поступово зсуваючи вікно аналізу на один період основного тону [4]. В такому випадку з‘являється можливість відділити постійно існуючу складову варіації періоду основного тону $\Delta p$, характерну для нормального голосу, від тремтіння $\varepsilon$, викликаного патологією голосових зв‘язок.

Монотонна складова варіації періоду основного тону визначається за формулою:

$$\Delta p = \frac{P_0(n) - P_0(1) - (-1)^n \varepsilon}{(n-1)}. \qquad (3)$$

Тремтіння, викликане патологією голосових зв‘язок, для вікна аналізу з чотирьох періодів основного тону:

$$\varepsilon = \frac{1}{4}\left[\left(P_0(2) - P_0(1)\right) - \left(P_0(3) - P_0(2)\right)\right]. \qquad (4)$$

Аналогічно до MDVP, оцінка локального тремтіння (як відношення середньої абсолютної різниці між сусідніми періодами до середнього періоду), виражена у відсотках, позначається *LocJitt*, а оцінка абсолютного значення локального тремтіння, виражена у мікросекундах, позначається *LocJita* [4].

## 3. Постановка експерименту

Для проведення експерименту по вимірюванню тремтіння голосу була створена база даних записів протяжних голосних звуків /a/, /o/, /e/, /i/ україномовних дикторів чоловічої та жіночої статі. Диктори відтворювали кожен звук спочатку на звичних для них частотах, а потім на високих частотах (фальцет). Для чоловічих голосів частота основного тону при фальцеті збільшувалася на 1 октаву, для жіночих голосів – на 0,5 октави. Умови проведення запису: мікрофон розміщений на відстані 10 – 15 см від рота, тривалість фонації кожного звуку 4 с. Програмне забезпечення, яке використовувалося для запису, – Microsoft Sound Recorder (Wave PCM, частота дискретизації 32 кГц, моно, 16 біт).

Записи дикторів були розділені на дві групи. В першу групу ввійшли записи п‘яти дикторів без скарг на голос. В другу – записи чотирьох дикторів з хриплістю в голосі.

Діапазон для пошуку частоти основного тону голосу з сигналу приймався у межах від 80 Гц (нижня межа основного тону для басів) до 600 Гц (верхня межа основного тону для тенорів). Після знаходження домінуючої частоти основного тону, подальший аналіз проводиться в діапазоні ($F_0 \pm F_0/2$), який включає всі частоти коливань голосових зв‘язок. Оскільки нормальні голосові зв‘язки в процесі фонації можуть змінювати частоту коливань на октаву нижче або вище від домінуючої частоти основного тону [5].

В отриманому сигналі за допомогою обчислення першої похідної знаходилися піки та ями. Відстань між двома сусідніми піками є тривалістю періоду основного тону. Тривалості всіх періодів утворюють послідовність $P_0(n)$, де $n$ – кількість періодів у сигналі.

Для кожної групи записів в середовищі Matlab проводились обчислення таких параметрів:

*Jitta* *(мкс)*
*Jitt* *(%)*
*LocJitta* *(мкс)*
*LocJitt* *(%)*
$\Delta p$ *(мкс)*

Отримані в цьому експерименті значення *Jitt* порівнювалися з результатами експерименту англійських вчених та лікарів Джонса Т., Треболда М., Плента Ф., Чісема Б., Іріса Дж.: середня кількість тремтіння в групі пацієнтів склала 9.8%, в групі з штучно створеною хриплістю (введення гістаміну в тканини голосових зв‘язок) склала 2.64%, а в групі з нормальним голосом склала 1.04 % [1].

## 4. Результати експерименту

Для записів голосних звуків /a/, /o/, /e/, /i/ кожного з дев‘яти дикторів проводилося обчислення параметрів *Jitta*, *LocJitta*, $\Delta p$ (на рисунках позначено *dp*), *Jitt* та *LocJitt* згідно методів, описаних у пункті 2. Середні значення величин *Jitta*, *LocJitta*, $\Delta p$ для першої групи дикторів (без скарг на голос) для звуків /a/, /o/, /e/, /i/ показано на рисунках 1 – 4 відповідно.

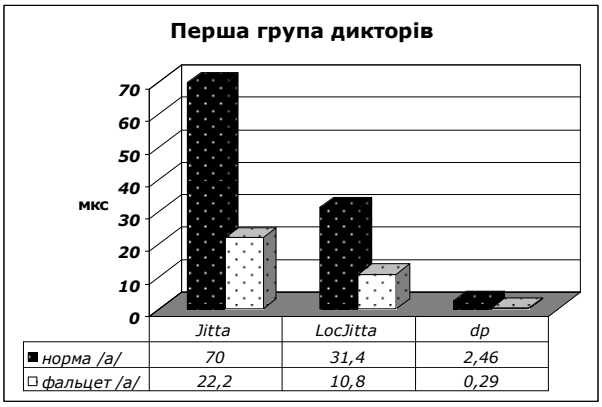


*Рисунок 1*: Гістограма варіації тривалості періоду основного тону для звуку /a/ для групи дикторів без скарг на голос.

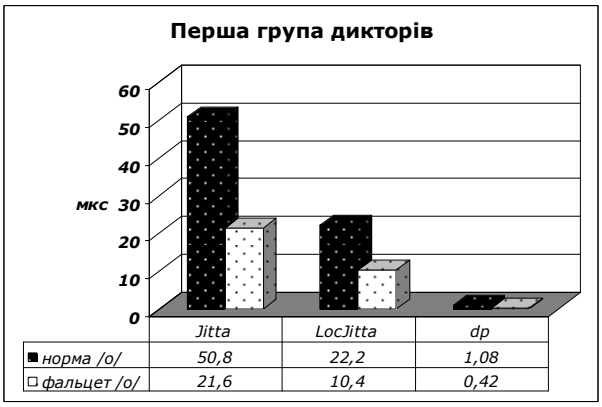


*Рисунок 2*: Гістограма варіації тривалості періоду основного тону для звуку /o/ для групи дикторів без скарг на голос.

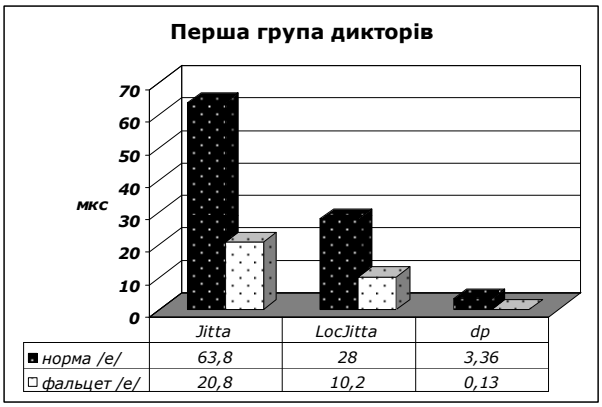


*Рисунок 3*: Гістограма варіації тривалості періоду основного тону для звуку /е/ для групи дикторів без скарг на голос.

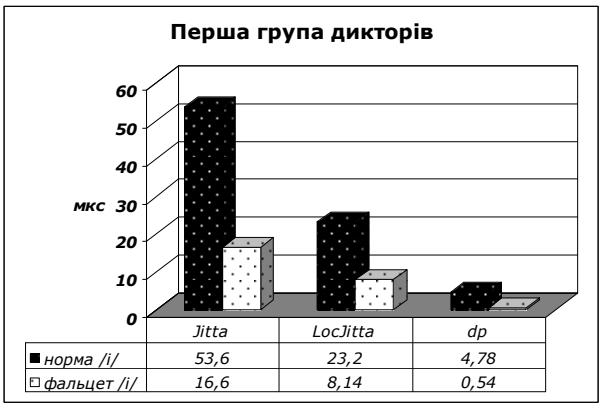


*Рисунок 4*: Гістограма варіації тривалості періоду основного тону для звуку /і/ для групи дикторів без скарг на голос.

Як видно з рисунків 1 – 4 величина *LocJitta* майже вдвічі менша ніж *Jitta*, але при цьому постійна складова варіації періоду основного тону $\Delta p$ дуже мала, як для нормальної так і фальцетної вимови голосних звуків.

Середні значення величин *Jitt* та *LocJitt* для першої групи дикторів (без скарг на голос) для звуків /а/, /о/, /е/, /і/ показано на рисунку 5.

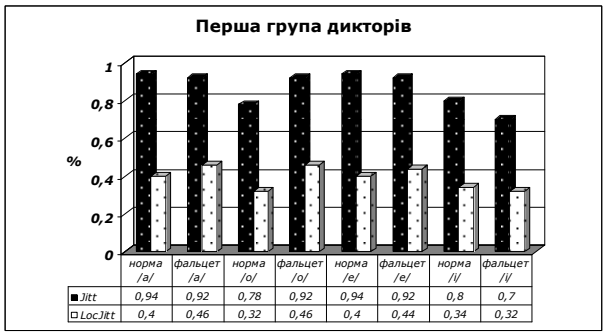


*Рисунок 5*: Гістограма відносної варіації тривалості періоду основного тону для звуків /а/, /о/, /е/, /і/для групи дикторів без скарг на голос.

На рисунках 6 – 9 показано середні значення величин *Jitta*, *LocJitta*, $\Delta p$ для другої групи дикторів (з хриплістю в голосі) для звуків /а/, /о/, /е/, /і/ відповідно.

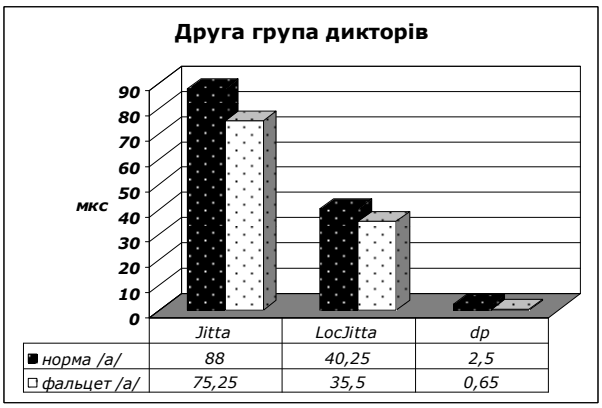


*Рисунок 6*: Гістограма варіації тривалості періоду основного тону для звуку /а/ для групи дикторів з хриплістю в голосі.

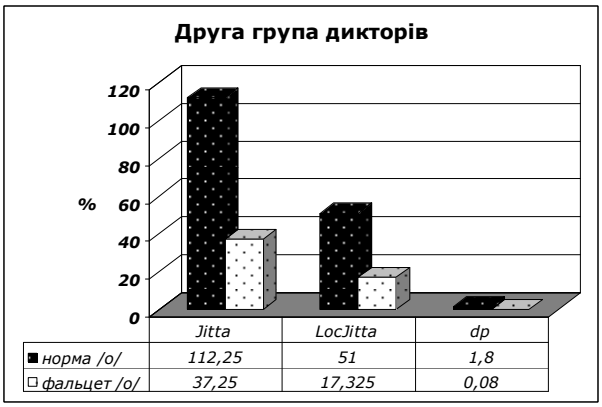


*Рисунок 7*: Гістограма варіації тривалості періоду основного тону для звуку /о/ для групи дикторів з хриплістю в голосі.

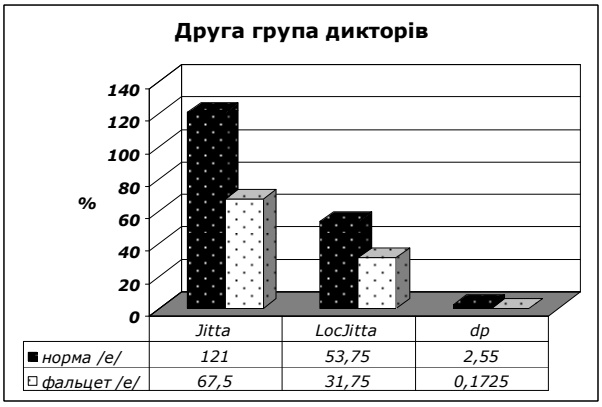


*Рисунок 8*: Гістограма варіації тривалості періоду основного тону для звуку /е/ для групи дикторів з хриплістю в голосі.

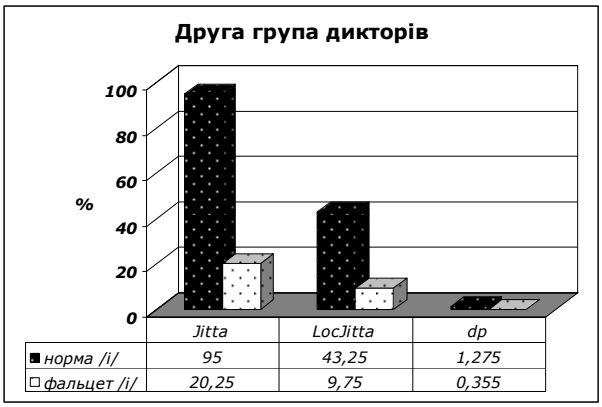


*Рисунок 9*: Гістограма варіації тривалості періоду основного тону для звуку /i/ для групи дикторів з хриплістю в голосі.

На рисунку 10 показано середні значення величин *Jitt* та *LocJitt* для другої групи дикторів (з хриплістю в голосі) для звуків /a/, /o/, /e/, /i/.

*Рисунок 10*: Гістограма відносної варіації тривалості періоду основного тону для звуків /a/, /o/, /e/, /i/для групи дикторів з хриплістю в голосі.

Як видно з рисунків 5 та 10 величина *LocJitt* майже вдвічі менша ніж *Jitt* як для нормальних, так і для патологічних голосів. При цьому значення *Jitt* для групи з нормальними голосами знаходяться в тому ж діапазоні, що і в експерименті англійських вчених та лікарів для групи з нормальними голосами [1].

## 5. Висновки

Помічено відмінність між величиною тремтіння при вимові звуків /a/ та /i/: для одного і того ж диктора величина *Jitt* для звуку /a/ більша ніж для звуку /i/. Це може пояснюватися тим, що, чим вища частота коливань голосових зв‘язок, тим менше збоїв в їхній роботі, а, як відомо, частота основного тону звуку /i/ завжди вища, ніж звуку /a/.

Для нормальних голосових зв‘язок (перша група дикторів) величина тремтіння *Jitt* склала 0,7 – 0,94 %; *LocJitt* – відповідно 0,32 – 0,46 %.

Для голосових зв‘язок з підозрою на запальний процес (друга група дикторів) величина тремтіння *Jitt* склала 1,2 – 3,7 %; *LocJitt* – відповідно 0,6 – 1,73 %.

Оскільки величини *Jitt* та *Jitta*, а також *LocJitt* та *LocJitta* прямо пропорційно пов‘язані між собою, то значення *LocJitta* завжди менше, ніж *Jitta*.

Постійна складова варіації періоду основного тону $\Delta p$ для фальцетної вимови завжди менша, ніж для звичної вимови голосних звуків як для нормальних, так і для патологічних голосів.

Фальцетна вимова голосних звуків передбачає зростання частоти коливань голосових зв‘язок, що сприяє зменшенню збоїв в їхній роботі, тому, навіть при хриплості в голосі, величина *Jitt* при нормальній фонації вища, ніж при фальцеті.

Для людей з хриплістю в голосі при фонації може спостерігатися нехарактерне збільшення частоти основного тону з метою підвищення якості голосу та розбірливості вимови. Така особливість може пояснюватися тим, що під час запальних процесів в гортані, в крові збільшується кількість гістаміну, що призводить до напухання органів гортані (в тому числі і голосових зв’язок), розмір голосової щілини зменшується, а розмір голосових зв’язок збільшується. Відомо що частота коливань голосових зв’язок залежить від їхнього розміру та форми. Тому для приведення збільшених голосових зв’язок у вібрацію, потрібний більший тиск потоку повітря з легень, що спричиняє збільшення частоти основного тону. Комбінація цих факторів призводить до того, що протягом деякого часу після пошкодження голосових зв’язок, голос людини підвищується.

## 6. Подяки

Особливу подяку хочу виразити аспірантам кафедри біотехнічних систем Тернопільського національного технічного університету імені Івана Пулюя та учасникам літньої школи-семінару “Мовленнєві інформаційні технології” Жукин’2010, які люб‘язно погодилися взяти участь в експерименті для формування бази даних протяжних голосних звуків /a/, /o/, /e/, /i/ української мови на звичних та фальцетних частотах.

## 7. Посилання

[1] Jones TM, Trabold M, Plante F, Cheetham BM, Earis JE, “Objective assesment of hoarseness by measuring jitter.”, Clin Otolaryngol Allied Sci., 2001 Feb; 26(1):29-32.

[2] http://www.wevosys.com/products/lingwaves/

[3] Multidimensional Voice Program Model 5105 http://www.kaypentax.com/

[4] Darcio G. Silva, Luis C. Oliveira and Mario Andrea, “Jitter estimation algorithms for detection of patological voices.”, EURASIP Journal on Advances in Signal Processing, vol. 2009, Article ID 567875, 9 pages.

[5] Jan G. Svec, Harm K. Shuttle, Donald G. Miller, “A subharmonic vibratory pattern in normal vocal folds.”, Journal of speech and hearing research, vol. 39 135-143 February 1996.