

ВИЯВЛЕННЯ ПАТОЛОГІЇ ГОЛОСУ НА ОСНОВІ СТАТИСТИЧНОЇ ОБРОБКИ ГОЛОСНИХ ЗВУКІВ УКРАЇНОМОВНИХ ДИКТОРІВ

© Драган Я., Джичка Н., 2010

Розглянута проблема коректної оцінки тремтіння голосу для україномовних дикторів. Проаналізовано методи обчислення тремтіння, запропоновані різними науковцями з різних країн. Виконано експериментальні дослідження застосування цих методів для обчислення тремтіння голосу для україномовних дикторів без скарг на хриплість, а також для дикторів з підозрою на запальний процес у гортані.

Ключові слова: патологія голосу, україномовний диктор, обробка голосу, оцінка тремтіння, хриплість.

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.

Keywords: pathological voice, Ukrainian speaker, voice processing, jitter estimation, hoarseness.

Постановка проблеми

За даними ВООЗ, щорічно у 5 % жителів планети виникають захворювання голосового апарату, викликані неправильним володінням голосу. Найбільший ризик виникнення цих захворювань у представників голосомовних професій – лекторів, дикторів радіо та телебачення, співаків, вчителів, телефоністів, диспетчерів, акторів, оскільки характер їхньої роботи спричиняє значні ненормовані навантаження голосового апарату. За важкістю розвитку захворювання поділяють на три групи. До першої групи належить фонастенія – порушення голосу без очевидних змін у гортані, котра проявляється тільки суб'єктивними відчуттями пацієнта. До другої групи професійних захворювань належать слабо виражені порушення у тканинах голосових зв'язок – стійке почервоніння країв голосових зв'язок, розширення кровоносних судин на їхній поверхні. До третьої групи професійних захворювань зараховують захворювання зі значними змінами в тканинах голосових зв'язок: «вузлики співаків», фіброми голосових зв'язок тощо. Переважно перша стадія захворювання проявляється протягом перших двох тижнів після прецеденту, друга – упродовж наступних двох тижнів, і третя (за відсутності будь-якого лікування) – через місяць. Отже, протягом місяця пацієнт, якому вчасно не діагностували захворювання та не розпочали лікування, продовжуватиме свою діяльність, погіршуючи стан органів голосового апарату, що призведе до утворення поліпів на голосових зв'язках, ускладнення змикання голосової щілини, дисфонії, афонії і втрати працездатності. Вчасне діагностування захворювання другої групи дає змогу запобігти серйознішим порушенням у тканинах голосових зв'язок та швидше відновити працездатність пацієнта.

Аналіз стану проблеми

Діагностування захворювань гортані та її внутрішніх органів здійснюється поетапно і передбачає фізичну оцінку голосу, неінвазивну оцінку голосу (зокрема і акустичну оцінку записів голосу за допомогою спеціалізованих комп'ютерних систем), інвазивні тести та фотографування.

Перші два етапи є поверхневим обстеженням, проте акустична оцінка записів дає змогу оцінити якість голосу та виявити порушення, невидимі під час дзеркального огляду. 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. Акустична оцінка дає змогу оцінити хриплість голосу та поставити попередній діагноз ще на початковій стадії захворювання, призначити лікування та відновити нормальний голос пацієнта за порівняно невеликий проміжок часу. При цьому не потрібні подальші інвазивні тести, які можуть бути неприємними та небезпечними для пацієнта.

Розвиток комп'ютерних технологій та вдосконалення методів акустичної оцінки записів звуків уможливорює об'єктивне оцінювання хриплості голосу та коректування діагнозу.

Найвідомішими 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. В Україні не здійснюється комп'ютерна оцінка записів голосу пацієнтів зі скаргами на порушення вимови, оскільки немає відповідного програмного забезпечення, орієнтованого на особливості вимови українських звуків.

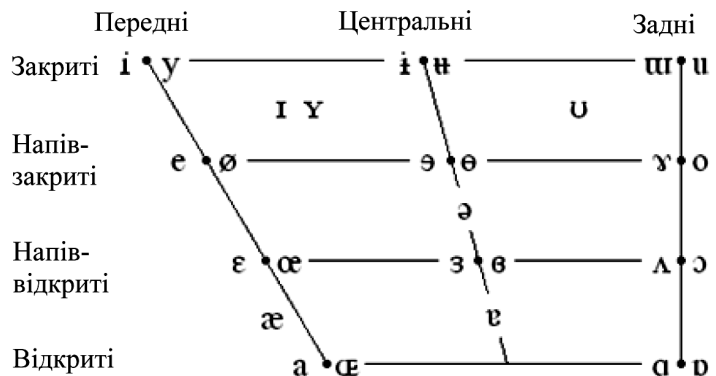


Рис. 1. Міжнародна класифікація голосних звуків [1]

Міжнародна класифікація голосних звуків налічує кілька варіантів вимови (рис. 1) одного і того самого звуку /a/ (для різних мовних груп вимова цих звуків відрізняється). Так, наприклад, німці та українці вимовляють звук /a/ по-різному. Те саме стосується й інших голосних звуків.

Тому важливим завданням для неінвазивного діагностування початкової стадії пошкодження голосових зв'язок в Україні є використання відповідного математичного апарату для розроблення програмного забезпечення, орієнтованого на особливості вимови українських звуків.

Постановка завдання

Метою роботи є застосування відомих методів обчислення тремтіння голосу до записів голосних звуків україномовних дикторів як з нормальними голосовими зв'язками, так і з підозрою на запальний процес у гортані.

Методи вимірювання хриплості

Об'єктивним параметром вимірювання хриплості є тремтіння – варіація тривалості хвилі основного тону голосового сигналу [2]. Найпоширеніший метод вимірювання тремтіння – Jitter (Jitt), який використовується в системах аналізу голосу lingWAVES Voice Clinic Suite Pro та Multidimensional Voice Program (MDVP) [3, 4]. Цей метод допускає періодичність сигналу протягом тривалого часу і характеризує увесь сигнал одним числовим значенням.

Пропонуються два підходи до визначення тремтіння голосу [5]:

- Тремтіння – це звичайна варіація тривалості хвилі основного тону, яка визначається як різниця тривалостей сусідніх хвиль основного тону

$$Jitta = \frac{1}{N-1} \sum_{n=1}^{N-1} |P_0(n+1) - P_0(n)|, \quad (1)$$

де $P_0(n)$ – послідовність тривалостей хвиль основного тону, мкс.

- Тремтіння – відношення середньої абсолютної різниці тривалостей сусідніх хвиль (з послідовності тривалостей) до середнього значення тривалості хвилі (%)

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

Дарсіо Дж. Сілва, Луїс С. Олівейра та Маріо Андреа запропонували обчислювати локальне тремтіння для вікна аналізу, що містить чотири послідовні хвилі основного тону, поступово зсуваючи вікно аналізу на одну хвилю основного тону. В такому разі з'являється можливість відділити постійну складову варіації тривалості хвилі основного тону Δp , характерну для нормального голосу, від тремтіння ϵ , спричиненого патологією голосових зв'язок [5].

Монотонна складова варіації тривалості хвилі основного тону визначається за формулою:

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

Тремтіння, викликане патологією голосових зв'язок, для вікна аналізу з чотирьох хвиль основного тону:

$$e = \frac{1}{4} [(P_0(2) - P_0(1)) - (P_0(3) - P_0(2))]. \quad (4)$$

Аналогічно до MDVP, оцінка локального тремтіння (як відношення середньої абсолютної різниці тривалостей сусідніх хвиль до середнього значення тривалості хвилі), виражена у відсотках, позначається *LocJitt*, а оцінка абсолютного значення локального тремтіння, виражена у мікросекундах, – *LocJita* [5].

Постановка експерименту

Для здійснення експерименту з вимірювання 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 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 бітів).

Записи дикторів поділено на дві групи. В першу групу увійшли записи п'яти дикторів без скарг на голос. У другу – записи чотирьох дикторів з хрипкістю в голосі.

Діапазон для пошуку частоти основного тону голосу з сигналу вибрано у межах від 80 Гц (нижня межа основного тону для басів) до 600 Гц (верхня межа основного тону для тенорів). Після виявлення домінуючої частоти основного тону подальший аналіз ведеться у діапазоні $(F_0 \pm F_0/2)$, який охоплює всі частоти коливань голосових зв'язок. Оскільки нормальні голосові зв'язки в процесі фонації можуть змінювати частоту коливань на октаву нижче або вище від частоти основного тону [6], що домінує.

В отриманому сигналі за допомогою обчислення першої похідної виявлено піки та ями. Відстань між двома сусідніми піками є тривалістю хвилі основного тону. Значення тривалостей всіх хвиль утворюють послідовність $P_0(n)$, де n – кількість хвиль основного тону в сигналі.

Для кожної групи записів в середовищі Matlab обчислювали такі параметри:

1. Середня варіація тривалості хвилі основного тону *Jitta* (мкс).
2. Відносна варіація тривалості хвилі основного тону *Jitt* (%).
3. Середня варіація тривалості хвилі основного тону *LocJitta* (мкс).
4. Відносна варіація тривалості хвилі основного тону *LocJitt* (%).
5. Монотонна складова варіації тривалості хвилі основного тону Δp (мкс).

Отримані в цьому експерименті значення *Jitt* порівнювалися з результатами експерименту англійських вчених та лікарів Т. Джонса, М. Треболда, Ф. Плента, Б. Чісема, Дж. Іріса: середня кількість тремтіння в групі пацієнтів становить 9.8 %, у групі з штучно створеною хрипкістю (введення гістаміну в тканини голосових зв'язок) становила 2.64 %, а в групі з нормальним голосом – 1.04 % [2].

Результати експерименту

Для записів голосних звуків /a/, /o/, /e/, /i/ кожного з дев'яти дикторів обчислювали параметри *Jitta* та *Jitt*, що пропонується в системах MDVP та lingWAVES; а також *LocJitta*, Δp , *LocJitt*, що пропонують дослідники Дарсіо Дж. Сілва, Луїса С. Олівейра та Маріо Андреа згідно з методами, описаними вище.

Значення параметрів *Jitta* (мкс), *Jitt* (%), *LocJitta* (мкс), *LocJitt* (%), Δp (мкс), обчислених для звуків /a/, /o/, /e/, /i/, вимовлених першою групою дикторів (без скарг на голос), наведено у табл. 1.

Таблиця 1

Параметри голосових сигналів першої групи дикторів (без скарг на голос)

Звук	Характер вимови	Параметри методу,		Параметри методу,		
		який використовують в MDVP та lingWAVES		запропонованого Дарсіо Дж. Сілва, Луїса С. Олівейра та Маріо Андреа		
		<i>Jitta</i> , мкс	<i>Jitt</i> , %	<i>LocJitta</i> , мкс	Δp , мкс	<i>LocJitt</i> , %
/a/	Нормальна вимова	70	0,94	31,4	2,46	0,4
	Фальцетна вимова	22,2	0,92	10,8	0,29	0,46
/o/	Нормальна вимова	50,8	0,78	22,2	1,08	0,32
	Фальцетна вимова	21,6	0,92	10,4	0,42	0,46
/e/	Нормальна вимова	63,8	0,94	28	3,36	0,4
	Фальцетна вимова	20,8	0,92	10,2	0,13	0,44
/i/	Нормальна вимова	53,6	0,8	23,2	4,78	0,34
	Фальцетна вимова	16,6	0,7	8,14	0,54	0,32

Як видно з табл. 1, величина *LocJitta* майже вдвічі менша, ніж *Jitta*, але постійна складова варіації тривалості хвилі основного тону Δp дуже мала як для нормальної, так для і фальцетної вимови голосних звуків.

Значення параметрів *Jitta* (мкс), *Jitt* (%), *LocJitta* (мкс), *LocJitt* (%), Δp (мкс), обчислених для звуків /a/, /o/, /e/, /i/, вимовлених другою групою дикторів (з хрипкістю в голосі), подано у табл. 2.

Таблиця 2

Параметри голосових сигналів другої групи дикторів (з хрипкістю в голосі)

Звук	Характер вимови	Параметри методу,		Параметри методу,		
		який використовують в MDVP та lingWAVES		запропонованого Дарсіо Дж. Сілва, Луїса С. Олівейра та Маріо Андреа		
		<i>Jitta</i> , мкс	<i>Jitt</i> , %	<i>LocJitta</i> , мкс	Δp , мкс	<i>LocJitt</i> , %
/a/	Нормальна вимова	88	1,7	40,25	2,5	0,775
	Фальцетна вимова	75,25	3,625	35,5	0,65	1,725
/o/	Нормальна вимова	112,25	2,025	51	1,8	0,925
	Фальцетна вимова	37,25	1,625	17,325	0,08	0,8
/e/	Нормальна вимова	121	2,075	53,75	2,55	0,925
	Фальцетна вимова	67,5	3,25	31,75	0,1725	1,525
/i/	Нормальна вимова	95	1,675	43,25	1,275	1,2
	Фальцетна вимова	20,25	0,75	9,75	0,355	0,6

Як видно з табл. 1 і 2, згідно з методом вимірювання хрипкості, запропонованим Дарсіо Дж. Сілва, Луїса С. Олівейра та Маріо Андреа, значення середньої та відносної варіації тривалості хвилі основного тону (*LocJitta* та *LocJitt*) майже вдвічі менші, ніж значення цих величин, обчислених за допомогою методу, що використовують в MDVP та lingWAVES (*Jitta* та *Jitt*).

Для нормальної вимови голосних звуків першої групи дикторів значення *Jitt* повинні міститися в тому самому діапазоні, що і в експерименті англійських вчених та лікарів Т. Джонса,

М. Треболда, Ф. Плента, Б. Чісема, Дж. Іріса для групи з нормальними голосами. А для нормальної вимови голосних звуків другої групи дикторів значення *Jitt* менші, ніж у тому самому експерименті для групи зі штучно створеною хриплистю [2].

Висновки

Встановлено відмінність між величиною тремтіння під час вимови звуків /a/ та /i/: для одного і того самого 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* становить 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* становить 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 тривалості хвилі основного тону Δp для фальцетної вимови завжди менша, ніж для нормальної вимови голосних звуків як для голосу в нормі, так і в разі патології.

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

Для людей з хриплистю 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 у гортані в крові збільшується кількість гістаміну, що призводить до напухання органів гортані (зокрема 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.

1. *International Phonetic Alphabet (IPA) [Електронний ресурс] : IPA Help / SIL International. 2008. – Current version: 2.1. – Режим доступу до ресурсу: <http://www.sil.org/computing/ipa-help/index.htm>* 2. *Objective assesment of hoarseness by measuring jitter / T.M. Jones, M. Trabold, F. Plante et al // Clin Otolaryngol Allied Sci. – 2001. – Vol. 26 (1). – P. 29–32.* 3. *lingWAVES : A leading tested product for voice and speech analysis, biofeedback and documentation [Електронний ресурс] / WEVOSYS Forchhein. – 2008. – Режим доступу до ресурсу: <http://www.wevosys.com/products/lingwaves/lingwaves.html>* 4. *Multidimensional Voice Program: Model 5105 / Division of PENTAX medical company. – NJ.: KayPENTAX. – USA, 2008. – 2 p.* 5. *Darcio G. Silva Jitter estimation algorithms for detection of pathological voices [Електронний ресурс] / Darcio G. Silva, Luis C. Oliveira, Mario Andrea // EURASIP Journal on Advances in Signal Processing. – 2009. – Vol. 2009. – 9 p. – Режим доступу до журн. : <http://www.hindawi.com/journals/asp/2009/567875.html>* 6. *Jan G. Švec. A subharmonic vibratory pattern in normal vocal folds / Jan G. Švec, Harm K. Shuttle, Donald G. Miller // Journal of speech and hearing research, 1996. – Vol. 39. – P. 135–143.*