Драґан Я. Обґрунтування математичної моделі фрикативного звуку у вигляді періодично корельованого випадкового процесу / Драґан Я., Яворська Є., Дозорський В. // Вісник ТНТУ. — 2010. — Том 15. — № 3. — С. 159-164. — (математичне моделювання. математика. фізика).

УДК 612.78:319.216

Я. Драґан, докт.фіз.-мат.наук, проф.;
Є. Яворська, канд.техн.наук, доцент; В. Дозорський

Тернопільський національний технічний університет імені Івана Пулюя

ОБҐРУНТУВАННЯ МАТЕМАТИЧНОЇ МОДЕЛІ ФРИКАТИВНОГО ЗВУКУ У ВИГЛЯДІ ПЕРІОДИЧНО КОРЕЛЬОВАНОГО ВИПАДКОВОГО ПРОЦЕСУ

Резюме. *Проаналізовано можливості моделювання фрикативного звуку як детермінованого та стаціонарного випадкового процесу. Виявлено неадекватність цих моделей задачі діагностики органів голосового апарату на ранніх етапах їх захворювання, бо вони не враховують повторюваність у сигналі, випадковість форм порушення вимови, пов'язану з функціональним станом органів голосового апарату. Обґрунтовано математичну модель фрикативного звуку у вигляді періодично корельованого випадкового процесу із використанням енергетичної теорії стохастичних сигналів, що враховує часову структуру, поєднання випадковості й періодичності сигналу.*

Ключові слова: *фрикативний звук, періодично корельований випадковий процес.*

Ya. Dragan, E. Yavorska, V. Dozorsky

GROUND OF MATHEMATICAL MODEL OF FRICATIVE SOUND AS THE PERIODICALLY CORRELATED RANDOM PROCESS

The summary. *The analysis of possibility of fricative sound modeling as the determined and stationary random process is conducted. The inadequacy of these models is found out for the task of vocal organs diagnostics on the early stages of disease, because they do not take into account the signal periodicity, chance of forms of pronunciation violation, related to the functional state of vocal organs. Grounded mathematical model of fricative sound as the periodically correlated random process with the use of power theory of stochastic signals, which takes into account a time structure, combination of chance and periodicity of signal.*

Key words: *fricative sound, periodically correlated random process.*

Постановка проблеми. Опрацювання голосових звуків є одним із методів дослідження стану голосового апарату людини [1] та проводиться для діагностики захворювань органів апарату на ранніх стадіях, виявлення місця і важкості патологічного процесу, контролю за лікуванням, визначення доцільності оперативних втручань. Для виявлення порушень у роботі органів голосового апарату потрібно проводити аналіз вокалізованих фрикативних звуків (ФЗ), що належать до класу приголосних і творяться при некогерентному збудженні порожнин рота потоком видихуваного повітря. В творенні цього типу звуків беруть участь органи джерела звуку та органи артикуляційного апарату людини [2, 3, 4].

Опис ФЗ за допомогою їхніх моделей виходить із необхідності відобразити суттєві, для даного типу задач, закономірності досліджуваних об'єктів і явищ, втілити їх у конструктивній математичній формі. Це забезпечує можливість теоретичного апріорного дослідження, обґрунтування алгоритмів та інтерпретацій отриманих результатів обчислень опрацювання ФЗ та формування із них інформативних ознак.

Аналіз останніх досліджень. На сьогодні можна виділити два підходи щодо побудови математичних моделей ФЗ – детермінований і стохастичний. Серед детермінованих відомими є моделі у вигляді суміші гармонічних сигналів, що виникають при збудженні порожнин рота [5]. Однак такі моделі не враховують випадкового характеру сигналу, пов'язаного з функціональним станом органів голосового апарату, та форм порушення вимови.

У випадку стохастичного підходу відомою є стохастична стаціонарна модель [5]. Вона відображає складність ФЗ у спектральному розподілі потужності, але не відображає його часової структури, що обов'язково повинна бути врахована, бо відображає роботу артикуляційного апарату [2-5].

Мета роботи. Побудувати математичну модель ФЗ, адекватну задачу діагностики захворювань органів голосового апарату людини, яка б враховувала часову структуру сигналу, присутню в нім повторюваність і випадковість.

Постановка задачі. На підставі аналізу моделей ФЗ можна сказати, що жодна модель не враховує його фізичну природу та випадковий характер порушень його вимови. З цього формується задача побудови адекватної моделі. Оскільки ФЗ є досить складним, за своєю природою, акустичним сигналом, що містить в собі випадковість і повторюваність. Тому для його опрацювання необхідно використати математичний апарат, який би враховував фізичну природу цього сигналу, поєднуючи в собі ці властивості.

Результати дослідження. ФЗ твориться в результаті некогерентного збудження мовного тракту шумовим джерелом [3,4]. Шум генерується турбулентним потоком повітря в місці звуження чи щілини. У випадку вокалізованих ФЗ повітряний потік проходить через голосові складки, що збуджуються квазіперіодичною послідовністю імпульсів, створюючи тим самим характерну періодичність у звукові, що називається основним тоном. Для досліджень використано фрикативний звук [л], реєстрограма якого зображена на рисунку 1.

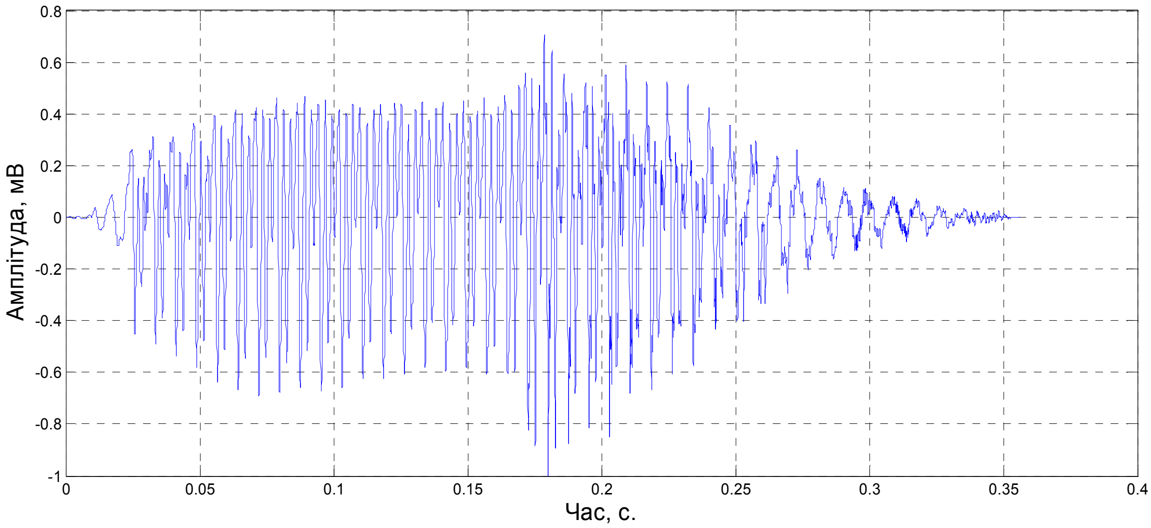

Рис. 1. Приклад реєстрограми фрикативного звуку [л]

При моделюванні ФЗ методами детермінованих моделей коливань виявлено, що спектри амплітуд залежать від вибору початкового моменту відліку часу, відстань між формантними лініями для ФЗ однієї і тієї ж людини є змінною. На рисунку 2 зображено амплітудні спектри вибірок однакового об'єму, взятих з реєстрограми ФЗ [л] через рівні проміжки часу.

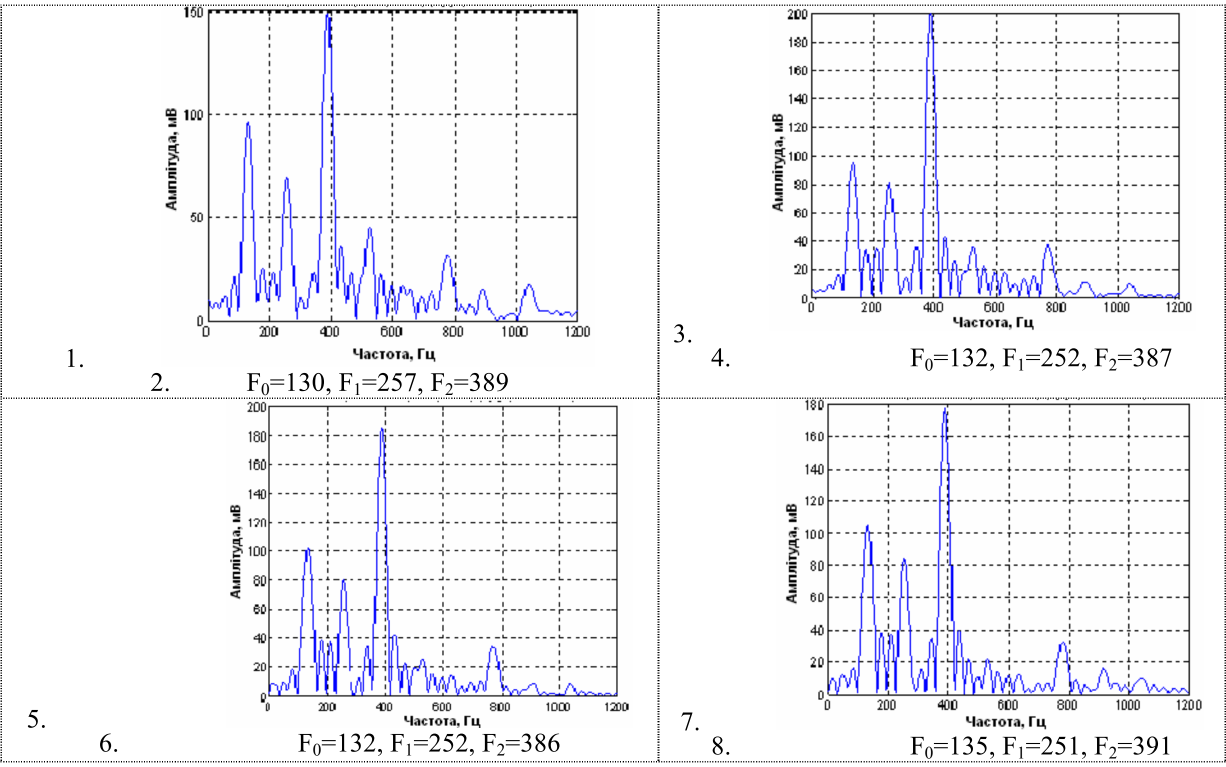

Рис. 2. Амплітудні спектри вибірок з реєстрограми звуку [л]. Тривалість вибірки – 0,03 с. F_0, F_1, F_2 – частоти, на яких розміщені перші три форманти

У рамках детермінованої моделі період основного тону є змінною неінваріантною характеристикою. Така модель не адекватна ні природі сигналу, ні задачі оцінювання стану голосового апарату. Можна припустити що в якості моделі доцільно використати випадковий стаціонарний процес.

На рисунку 3 зображено оцінювання автокореляційної функції, де бачимо, що в сигналі є присутня повторюваність.

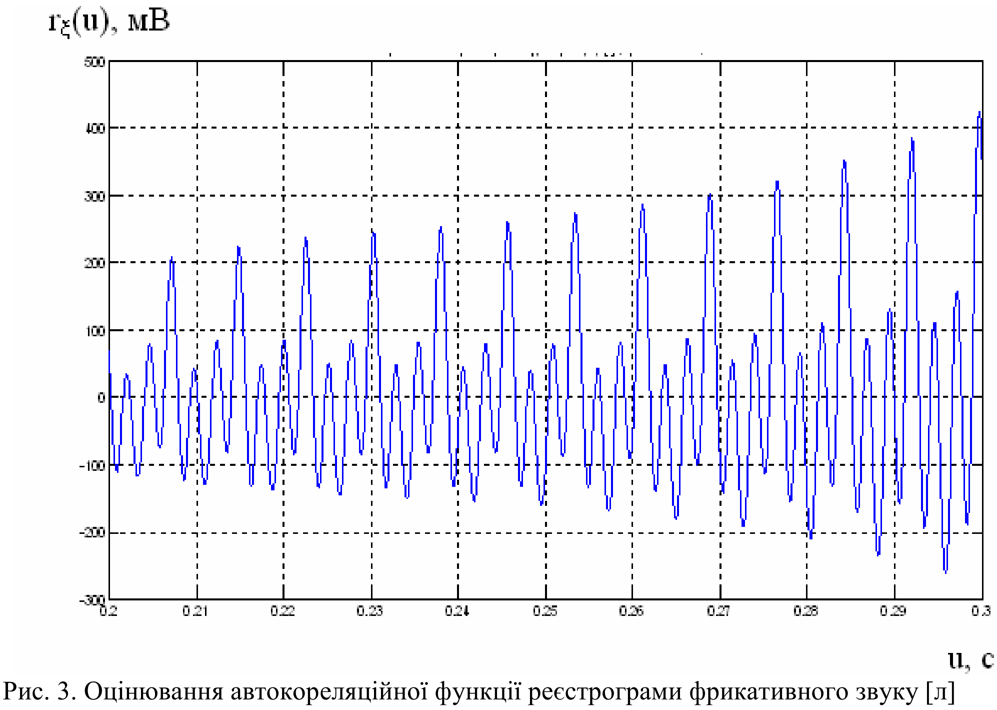

Рис. 3. Оцінювання автокореляційної функції реєстрограми фрикативного звуку [л]

Відповідно до [5], ФЗ на коротких проміжках часу, рівних кільком періодам основного тону, можна розглядати як стохастичний стаціонарний процес, однак в цілому він є нестаціонарним змінним упродовж періоду процесу.

Враховуючи сказане, можна припустити, що в якості моделі доцільно використати нестаціонарний випадковий процес. Відповідно до енергетичної теорії стохастичних сигналів (ЕТСС), враховуючи скінченність потужності сигналу в межах періоду, можна припустити, що моделлю в даному разі є нестаціонарний випадковий процес скінченої середньої потужності (класу π^{T}). Тоді ЕТСС обґрунтовує алгоритм обчислення оцінок характеристик стаціонарного наближення для випадкових процесів класу π, які будуть дорівнювати оцінкам стаціонарного процесу, що складається з таких самих як і процес класу π гармонік і з такими самими потужностями, але вже некорельованих (бо стаціонарність і корельованість гармонік для процесів класу π рівносильні) [6, 7].

Енергетична теорія стохастичних сигналів обґрунтовує зображення такого типу сигналів із законом збереження середньої потужності при цьому та вказанням типу їхньої корельованості у часовій області або у часі, повторюваністю ймовірнісних характеристик, зображення через стаціонарні компоненти і стаціонарні послідовності відліків [6, 7].

Відповідно до ЕТСС з цього випливає, що адекватною моделлю фрикативного звуку буде математична модель у вигляді періодично корельованого випадкового процесу класу π^T, яка поєднує випадковість значень з повторюваністю, трактуючи її як періодичність імовірнісних характеристик. Тоді задача зведеться до того, щоб на підставі апріорного теоретичного аналізу структури цієї моделі виявити можливі інваріанти, що їх може дати опрацювання емпіричних даних, та обґрунтувати алгоритм цього опрацювання [6, 7].

Математична модель у вигляді періодично корельованого випадкового процесу. Періодично корельовані випадкові процеси (ПКВП) – це моделі стохастичних коливань з періодичною зміною ймовірнісних характеристик. Вони творять підклас π^T [6, 7].

Означення класу π мало однією із причин вивчення ПКВП у спектральній області – їхньої гармонізованості, вигляду спектру (типу корельованості гармонічних складових), закону збереження (потужності в узагальненій теоремі Вінера-Хінчина) та зображення таких процесів через спектральні компоненти. Очевидно, що умова (1) при обчисленні середніх характеристик призводить до того, що усереднення по всій осі переходить в усереднення по відрізку довжини T, що (внаслідок інваріантності усереднення зсувів) можна вважати як усереднення на відрізку $[0,T)$, тобто середні величини характеристики процесу записати виразами [6, 7]

$$m = M_t\{m(t)\} = \frac{1}{T}\int_0^T m(t)dt \text{ та} \tag{1}$$

$$B(u) \overset{\Delta}{=} M_t\{r(t+u,t)\} = \frac{1}{T}\int_0^T r(t+u,t)dt = \frac{1}{T}\int_0^T b(t,u)dt\,, \tag{2}$$

де M_t – символ усереднення по всій осі; T – період корельованості ПКВП.

Тому характеристики мають розклади у ряди Фур'є:

$$m(t) = \sum_{k\in Z} m_k e^{ik\Lambda t} \text{ та} \tag{3}$$

$$b(t,u) = \sum_{t\in Z} B_k(u) e^{ik\Lambda t}\,, \tag{4}$$

де $\Lambda \stackrel{\Delta}{=} \frac{2\pi}{T}$,

як слід розуміти у сенсі теорії узагальнених функцій Шварца [6, 7], коли розглядати ПКВП скінченої середньої потужності. Справді, оскільки у випадку періодичної функції

$$M_t\{f(t)\} = \lim_{\Theta\to\infty} \frac{1}{2\Theta} \int_{-\infty}^{\infty} f(t)dt = \lim_{\Theta\to\infty} \frac{1}{2NT + 2\Delta}\left[N\int_{o}^{T} + N\int_{-\Theta}^{-NT} + N\int_{NT}^{\Theta} \right] f(t)dt = \frac{1}{T}\int_0^T f(t)dt, \quad (5)$$

де $N = E\left(\frac{\Theta}{T}\right)$, $E(\bullet)$ – ціла частина числа, $\Delta = \Theta - NT$, то звідси середня потужність ПКВП визначається усередненням на періоді корельованості, тобто на відрізку $[0,T]$. Тоді

$$P^T{}_\xi = \frac{1}{T}\int_0^T E\left|{}^0\xi(t)\right|^2 dt = \frac{1}{T}\int_0^T r(t,t)dt, \quad (6)$$

а умова належності до класу π набуде вигляду

$$P^T{}_\xi = \frac{1}{T}\int_0^T r(t,t)dt < \infty. \quad (7)$$

Тому клас ПКВП, для якого виконується ця умова, коли його трактувати як підклас у класі π, був названий класом π^T. На цей клас поширюється з відповідними видозмінами теорія класу π [6, 7]. Зокрема замість простору $\hbar^\pi$ вводиться простір $\hbar^T = L^2\left([0,T); \frac{1}{T}, K\right)$, тобто простір інтегрованих на $[0,T)$ з квадратом за мірою dt/T функцій над колгоморівським гільбертовим простором K випадкових величин скінченої дисперсії, і норма у цьому просторі $\|\xi_\bullet\|_{\hbar^T} = \sqrt{P_\xi^T}$.

Кореляційні компоненти

$$B_k(u) \stackrel{\Delta}{=} \frac{1}{T}\int_0^T r(t+u,t)e^{-ik\Lambda t}dt \quad (8)$$

у міру їхньої обмеженості $B_k(u) \le B_0(u) \le B_0 = P_\xi^T$ належить до класу B^2, тому мають зображення у вигляді ряду Фур'є за мірами (загалом комплексно значними)

$$B_k(u) = \int_R e^{iu\lambda} F_k(d\lambda), \quad (9)$$

де R – множина дійсних чисел, F – спектральна біміра.

Коваріація ПКВП має зображення у вигляді

$$r(t,s) = \iint_{R2} e^{i(d\lambda, d\mu)} F(d\lambda, d\mu). \quad (10)$$

Встановлені властивості коваріації ПКВП дають підставу вивести вираз самого процесу через його стаціонарні складові, тобто встановити структуру цього класу процесів [5].

ПКВП належить до класу π^T тоді і тільки тоді, коли він має зображення

$$\xi(t) = \sum_{k \in Z} \xi_k(t) e^{ik\Lambda t} , \tag{11}$$

де, $\xi_k(t)$ – стаціонарні компоненти ПКВП, Z – множина цілих чисел.

На основі вибраної математичної моделі реалізовано алгоритми опрацювання ФЗ засобами ЕТСС (синфазний, компонентний і фільтровий) для отримання статистичних оцінок їхніх імовірнісних характеристик, які є показниками стану органів голосового апарату [6, 7].

Висновки. З аналізу структури звуку [л] та описаних властивостей періодично корельованих випадкових процесів випливає, що математична модель процесу такого класу дає змогу адекватно описати сигнал, а саме врахувати поєднання випадковості та періодичності сигналу, а тому й розробити методи визначення інваріантних інформативних ознак звуку, виходячи зі статистики таких сигналів для завдань діагностики захворювань органів голосового апарату на ранніх стадіях.

Література

1. Джафек Брюс. Секреты оториноларингологии: пер. с англ. / Б. Джафек, Е. Старк [под ред. Н. И. Новикова, А. Ю. Овчинникова]. – М.: СПб: БИОНОМ: Невский диалект, 2001. – 624 с. – ISBN 5-79870197-1.
2. Фант Гунер. Акустическая теория речеобразования: пер. с англ. / Гунер Фант [под ред. В. С. Григорьева]. – М.: Наука, 1964. – 284 с.
3. Фланаган Джеймс. Анализ, синтез и восприятие речи: пер. с англ. / Джеймс Фланаган [под ред. А. А. Пирогова]. – М.: Связь, 1968. – 396 с.
4. Sadaoki Furui. Digital speech. Processing, synthesis and recognition / Sadaoki Furui. – Tokyo: Tokyo institute of technology, 2000. – 439 c. – ISBN 0-8247-0452-5.
5. Рабинер Лоренс. Цифровая обработка речевых сигналов: пер. с англ. / Л. Рабинер, Р. Шафер [под ред. М. В. Назарова, Ю. Н. Прохорова]. – М.: Радио и связь, 1981. – 496 с.
6. Драган Ярослав Петрович. Енергетична теорія лінійних моделей стохастичних сигналів: монографія / Я. П. Драган. – Львів: Центр стратегічних досліджень еко-біотехнічних систем, 1997. –XVI+333 с. – ISBN 5-12-003724-0.
7. Драган Ярослав Петрович. Ритмика морского волнения и подводные акустические сигналы / Я. П. Драган, И. М. Яворский. – К.: Наукова думка, 1982. – 246 с.

Отримано 02.09.2010 р.