

# **Automatyczna ocena zaburzeń realizacji głosek w dysartrii**

## **Automatic assessment of distortions in isolated sounds in dysarthria disease**

Tomasz Orzechowski, Katarzyna Chmurzyńska,  
Piotr Radkowski, Andrzej Izworski

Akademia Górniczo-Hutnicza, Instytut Automatyki, Kraków  
fando@agh.edu.pl, kate@agh.edu.pl  
piotr.radkowski@agh.edu.pl  
izwa@agh.edu.pl

### STRESZCZENIE

Przeprowadzone badania stanowią kolejny etap mający na celu opracowanie najskuteczniejszej metody pozwalającej na automatyczną ocenę postępu choroby neurodegeneracyjnej na podstawie stopnia zmian w głosie. Badania przeprowadzono na pacjentach kliniki CM-UJ Kraków, u których stwierdzono uszkodzenia układu pozapiramidowego. Materiał lingwistyczny został celowo dobrany tak, aby wychwycić zmiany głosu charakterystyczne dla pacjentów z dysartrią występującą w tego typu chorobach neurodegeneracyjnych. W czasie pilotażowego badania zarejestrowano kilkanaście zróżnicowanych nagrań dla każdej z 18 osób składającą się z pojedynczych głosek, wyrazów oraz fraz. Wstępne badania pokazały, że najskuteczniejszą metodą jest ocena stabilności realizacji pojedynczych, wyizolowanych głosek różnymi metodami. Materiał dla którego uzyskano zadowalające wyniki klasyfikacji poddano kolejnej analizie, mającej na celu poprawę skuteczności oceny postępu zmian w głosie i lepsze przygotowanie materiału lingwistycznego do planowanych, właściwych badań.

### ABSTRACT

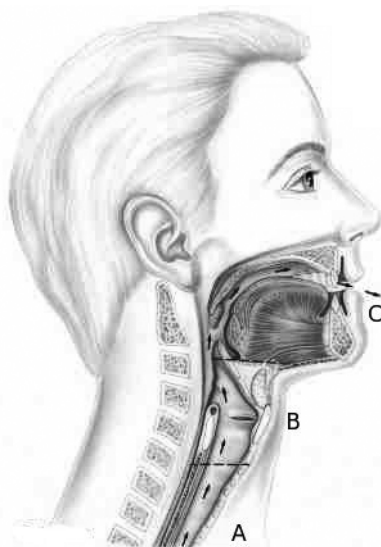
Research presented in this paper are next phase of survey which main objective is to propose a procedure allowing automatic assessment of neurodegenerative disease progression. Analysis are based on pathological changes in speech caused by dysarthria. The examination was carried out on group of patients of CM-UJ clinic in Krakow who suffered from neurodegenerative disease with the damage of the extrapyramidal system. During initial research number of voice samples was recorded from each patient. The elements which realization was tested had been chosen based on the linguistic knowledge and were targeted on identifying anomalies typical for patients with dysarthria. Previous studies showed that analysis of stability of isolated sounds in speech is most effective method. It allows to parameterize the research results, making complex classification feasible. Classified samples were subjected to further analysis. Current research are concentrated on increasing accuracy of disease progression assessment and improvement of linguistic material for next phase of studies.

## 1. Wstęp

Prezentowana metoda jest pierwszym etapem konstrukcji systemu do automatycznego wykrywania jak i oceny zmian w głosie chorego. W szczególności położono nacisk na wykrycie i ocenę drżenia biomedycznego, obecnego w głosie w postaci zmian o charakterze dyzartrii czyli upośledzenia zdolności wytwarzania artykułowanych dźwięków mowy w skutek zakłócenia nerwowych mechanizmów tworzenia głosek [2].

Wytwarzanie głosu i mowy wymaga ścisłego współdziałania ze sobą wielu ośrodków, które z foniatrycznego punktu widzenia można by przypisać do narządów (ryc. 1):

- wytwarzających niezbędny dla fonacji strumień powietrza wydechowego: płuca, oskrzela, tchawica (A),
- wytwarzających ton podstawowy: krtań (B),
- formujących obwiednię widma i tworzące głoski mowy: nasada, gardło, jama nosowa, jama ustna (C).



Rycina 1. Narząd głosu człowieka. Źródło: Medical Illustrations ([www.origin8.nl/medical](http://www.origin8.nl/medical)).

Warunkiem prawidłowej mowy i śpiewu jest harmonijna współpraca wszystkich wymienionych wyżej narządów, a nieprawidłowy przebieg jednej czynności pociąga za sobą zaburzenie pozostałych. Obok typowych zmian wynikających z choroby neurodegeneracyjnej (np. drżenie ciała, kończyn, sztywność mięśni) obserwuje się zmiany w głosie [3].

## 2. Zmiany w głosie

U pacjentów z chorobami neurodegeneracyjnymi występują zmiany w mowie o typie dyzartrii czyli zaburzeń opuszkowo-fonacyjno-artykulacyjnych spowodowanych uszkodzeniem ośrodków i dróg unerwiających aparat mowy, występujących mimo zachowa-

nego planu mowy [4]. Inne definicje charakteryzują dyzartrię jako upośledzenie zdolności wytwarzania artykułowanych dźwięków mowy wskutek zakłócenia nerwowych mechanizmów tworzenia głosu oraz modulacji, natężenia, kształtu i rezonansu [2].

Typem dyzartrii, który poddano powtórnej analizie i który występuje w chorobach neurodegeneracyjnych układu pozapiramidowego jest dyzartria hiperkinetyczna. Charakteryzuje się on między innymi:

- zniekształceniem fonacji przez zmiany głośności,
- obecnością nagłych przerw w wypowiedzi,
- nieprawidłową artykulacją oraz nieregularnością przerw w artykulacji,
- przedłużaniem i powtarzaniem głosek na skutek nieprawidłowego napięcia mięśni,
- występowaniem nadmiernej nosowości oraz skracaniem fraz,
- występowaniem wahań głośności; głos jest drżący, słaby, pojawiają się przerwy w emisji.

### 3. Badania pilotażowe

Głos osób został zarejestrowany przy pomocy profesjonalnego sprzętu cyfrowego, w wytłumionym pomieszczeniu w celu eliminacji wszelkich niepożądanych czynników mogących wpłynąć negatywnie na końcowy wynik. Grupa badawcza stanowiła 18 osób chorych, w wieku 20–80 lat. Grupa porównawcza była złożona z osób zdrowych w wieku 20–50 lat.

Wstępna analiza nagrań składających się z pojedynczych głosek, wyrazów oraz fraz pokazała, że najskuteczniejszą i najprostszą w realizacji metodą oceny jest analiza pojedynczych głosek.

Badani proszeni byli o wypowiedzianie głosek w izolacji. Analiza ta pozwoliła na automatyczne rozpoznanie kilku cech charakterystycznych dla dyzartrii. Przede wszystkim chodzi o zmianę w postaci zakłócenia mowy polegającą na nieoczekiwanych przerwach w fonacji, jak i również wykrycie wahań głośności oraz drżenia.

Grupę samogłosek stanowiły [a], [e] oraz [i]. Grupę spółgłosek stanowiły [s], [x], [p], [k], [g]. Głoski [s], [x] są przykładami z grupy głosek szczelinowych natomiast głoski [p], [k], [g] zwarto-wybuchowych.

Badanych proszono o wypowiedzenie każdej głoski na jednym oddechu. Długi czas emisji pozwala na szczegółową analizę.

Analiza pozwoliła wyodrębnić trzy podstawowe algorytmy oceny zmian realizacji głosek.

#### 3.1. Analiza widma realizacji głoski

Do porównania realizacji głosek osób zdrowych i chorych wykorzystano analizę widma otrzymaną za pomocą Dyskretnej Transformaty Fouriera. Analiza widma w skali Barkhausena pozwoliła na lepszą klasyfikację zarejestrowanych danych. W celu efektywnej klasyfikacji danych wyznaczono macierz energii, a następnie zmniejszono wymiar macierzy przez zsumowanie wartości znajdujących się w kolejnych przedziałach odpowiadających skali Barkhausena.

U osób chorych widoczne są zmiany w wymawianych spółgłoskach – dokładnie przejście w inną głoskę w czasie realizacji. Jest ono zarówno słyszalne, jak i możliwe

do automatycznego wykrycia na podstawie porównania macierzy widma. Zmiany te obserwowalne były szczególnie dla podanych niżej spółgłosek, dla których zaznaczono także występującą zmianę.

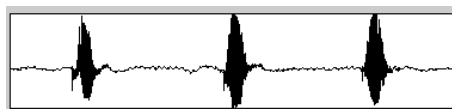
<b>głoska</b>	<b>zmiana realizacji</b>
[k]	[k] → [a] oraz [k] → [y]
[g]	[g] → [y] oraz [g] → [e]
[s]	[s] → [y]
[x]	[x] → [a]

Takie zaburzenia spowodowane są osłabieniem sprężystości mięśni krtani. Dlatego spółgłosce towarzyszy samogłoska, która nie wymaga takiego napięcia.

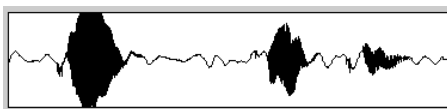
Końcowy efekt tej analizy pokazał, że mimo widocznej zmiany realizacji głosek świadczącej o występowaniu choroby, nie można obiektywnie i miarodajnie ocenić stopnia zaawansowania choroby.

### 3.2. Analiza natężenia głosek wymawianych wielokrotnie w izolacji

Badani byli poproszeni o powtórzenie cztery razy tej samej głoski zwarto-wybuchowej. Sugestia wynikała z wiedzy, iż w czasie realizacji sekwencji tych samych głosek spodziewano się zaburzenia sekwencji dźwięku oraz pojawienia się mamrotania. Analizując przebieg czasowy i amplitudę kolejnych głosek, obserwujemy u większości badanych, że amplituda końcowych głosek była odpowiednio niższa od pierwszej (ryc. 2a i 2b). Spowodowane jest to sztywnością mięśni, która jest cechą charakterystyczną dla choroby Parkinsona.



Rycina 2a. Amplituda ostatnich trzech głosek osoby zdrowej.



Rycina 2b. Amplituda ostatnich trzech głosek osoby chorej.

Analiza wyników potwierdziła, że w grupie kontrolnej zjawisko spadku amplitudy kolejnych głosek było znacznie rzadsze. Pozwala to, podobnie jak analiza widma, na stwierdzenie, że badany ma symptomy chorobowe jednak ocena stopnia zaawansowania choroby jest bardzo przypadkowa. Analiza kilku nagrań zarejestrowanych dla tej samej osoby wykazała dużą przypadkowość oraz silny związek ze zmęczeniem pacjenta, co wykluczyło tą metodę jako obiektywną ocenę zaawansowania choroby.

### 3.3. Analiza stabilności dźwięków ciągłych

Badanych poproszono o wypowiedzenie na jednym oddechu głoski: [a], [e], [i], [s], [x]. Dźwięk dłużej emitowany pozwolił na analizę energii w celu zaobserwowania przeniesienia zmian częstotliwości wynikającej z patologicznego drżenia. Uzyskane wartości porównano z wartościami grupy kontrolnej.

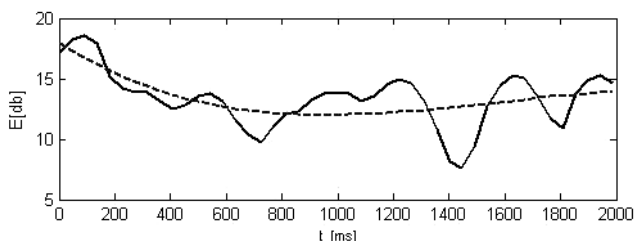
W celu uśrednienia wyników każdy zarejestrowany sygnał głoski został podzielony na ramki 160 ms, z nałożeniem ramek 40 ms. Moc poszczególnych ramek  $P_2$  została wyznaczona przez zsumowanie wartości znajdujących się w zakresie od  $x(t)$  do  $x(t+m)$ , gdzie  $m$  jest długością ramki.

$$P_2(t) = \sum_{k=t}^{t+m} |x(k)|^2 \quad (1)$$

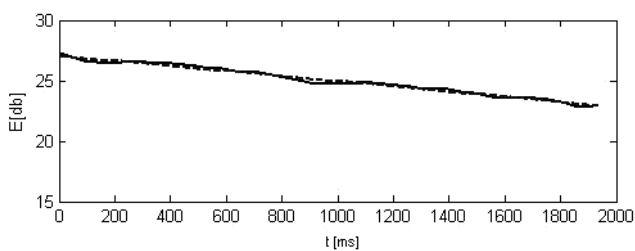
W celu wyznaczenia stabilności realizacji głosek dopasowano do otrzymanego wyżej przebiegu wielomian  $p(t)$  stopnia 4, a następnie określono stabilność jako współczynnik  $Stab$ .

$$Stab = \sum_{k=1}^t |P_2(t) - p(t)| \quad (2)$$

Poniżej podano przykładowy wynik badania osoby chorej i zdrowej.



Rycina 3a. Głoska [a] z grupy osób chorych,  $Stab=1,43$ .



Rycina 3b. Głoska [a] z grupy osób kontrolnych,  $Stab=0,14$ .

Wartość współczynnika  $Stab$  był znacznie większy dla osób chorych, co pozwoliło na wyznaczenie kolejnego, znaczącego parametru klasyfikującego osoby ze zmianami patologicznymi w mowie.

Najlepszy efekt uzyskano dzięki analizie zmian mocy sygnału w zadanych odcinkach czasowych. Na tej podstawie wyznaczono współczynnik stabilności realizacji głosek, który wyraźnie określał stopień nasilenia choroby.

## 4. Przegląd metod do oceny stabilności realizacji głóska

Badania pilotażowe wykazały, że najskuteczniejszą analizą była ocena stabilności realizacji głóska na podstawie analizy energii sygnału. Dlatego też zdecydowano się podać cały materiał powtórnej, bardziej szczegółowej analizie. Do dalszych badań wybrano trzy metody analizy mowy.

### 4.1. Noise-to-harmonic ratio (NHR)

NHR jest parametrem określającym stosunek szumu do sygnału. Wartość NHR określa stosunek części nieharmonicznej (hałasu) w paśmie powyżej 1500 Hz do części harmonicznej poniżej 1500 Hz. Wartość NHR jest zatem odzwierciedleniem zawartości szumu w sygnale.

### 4.2. Jitter

Parametr pozwala określić względne zmiany częstotliwości podstawowej z okresu na okres w analizowanej próbce dźwięku.

Jeżeli uśredniona częstotliwość podstawowa tonu zostanie określona jako:

$$\bar{f}_0 = \frac{\sum_{i=1}^N f_i}{N} \quad (3)$$

gdzie  $f_i$  jest częstotliwością w  $i$ -tym okresie, wówczas współczynnik Jitter może być zdefiniowany jako:

$$J = \frac{\sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (f_i - f_{i-1})^2}}{\frac{1}{N} \sum_{i=1}^N f_i} \cdot 100\% \quad (4)$$

Wartość parametru Jitter jest odzwierciedleniem częstotliwości drgań strun głóskowych.

### 4.3. Średnia częstotliwość podstawowa

Nagrania są dzielone na ramki i dla każdej z nich jest określana częstotliwość podstawowa F0. W czasie analizy jest badana stabilność częstotliwości podstawowej pomiędzy kolejnymi ramkami, zdefiniowana jako odchylenie standardowe F0 dla wszystkich ramek w próbce.

## 5. Podsumowanie

Celem pracy było przedstawienie metod opisu zniekształconej mowy pozwalających na uzyskanie lepszych niż dotychczas [5] wyników automatycznej klasyfikacji stadiów choroby neurodegeneracyjnej. Materiał dla którego uzyskano zadowalające wyniki klasyfikacji poddano kolejnej analizie, mającej na celu poprawę skuteczności oceny postępu zmian w głósie i lepsze przygotowanie materiału lingwistycznego do planowanych, właściwych badań. Przetestowane i wybrane metody oceny stabilności głóska zostaną użyte do badań na całym dostępnym materiale fonetycznym.

## BIBLIOGRAFIA

- [1] Ryszard Tadeusiewicz (1998), *Sygnal mowy*, Wydawnictwa Komunikacji i Łączności, Warszawa.
- [2] Johns D. F. (1985), *Clinical management of neurogenic communicative disorders*, College Hill, Boston.
- [3] Ramig L., Sapir S., Countryman S. (2001), *Intensive voice treatment for patients with Parkinson's disease*, *Journal Neurol Neurosurg Psychiatry* 71, p. 439–498.
- [4] Lamig L., Sapir S., Countryman S. (2001), *Intensive voice treatment (LSVT®) for patients with Parkinson's disease: a 2 year follow up*, *Journal Neurol Neurosurg Psychiatry* 71, p. 439–498.
- [5] Izworski A., Augustyniak P., Orzechowski T. (2006), *Processing and analysis of voice anomalies in course of Parkinson's diseases*, Proceedings of the eighth IASTED International Conference on Signal and Image Processing, August 14–16, 2006, Honolulu, Hawaii, USA, ACTA Press, p. 354–357.

