

Ewa Niebudek-Bogusz<sup>1</sup>Paweł Strumiłło<sup>2</sup>Justyna Wiktorowicz<sup>3</sup>Mariola Śliwińska-Kowalska<sup>1</sup>

## PORÓWNANIE WYNIKÓW ANALIZY CEPSTRALNEJ Z INNYMI PARAMETRAMI OCENY GŁOSU U PACJENTÓW Z DYSFONIAMI ZAWODOWYMI

COMPARISON OF CEPSTRAL COEFFICIENTS TO OTHER VOICE EVALUATION PARAMETERS  
IN PATIENTS WITH OCCUPATIONAL DYSPHONIA

<sup>1</sup> Instytut Medycyny Pracy w Łodzi / Nofer Institute of Occupational Medicine, Łódź, Poland  
Klinika Audiologii i Foniatrii / Audiology and Phoniatrics Clinic

<sup>2</sup> Politechnika Łódzka / Lodz University of Technology, Łódź, Poland  
Instytut Elektroniki / Institute of Electronics

<sup>3</sup> Uniwersytet Łódzki / University of Lodz, Łódź, Poland  
Wydział Ekonomiczno-Socjologiczny, Katedra Statystyki Ekonomicznej i Społecznej / Faculty of Economics and Sociology,  
Chair of Economic and Social Statics

### STRESZCZENIE

**Wprowadzenie:** W ostatnim czasie wśród obiektywnych metod oceny głosu uznaniem cieszy się analiza akustyczna oparta na wyznaczeniu współczynników cepstralnych MFCC (mel-frequency cepstral coefficients). Celem badania była ocena ich zastosowania w diagnozowaniu dysfonii zawodowych w porównaniu z innymi subiektywnymi i obiektywnymi parametrami diagnostycznymi zaburzeń głosu. **Materiał i metody:** W badaniu wzięły udział 2 grupy kobiet: grupa badana – 55 nauczycielek (średni wiek: 45 lat) z dysfoniami o podłożu zawodowym, potwierdzonymi badaniem laryngowidestroboskopowym, oraz grupa porównawcza – 40 kobiet z głosem prawidłowym (średni wiek: 43 lata). Próbkę dźwiękową (samogłoska 'a' oraz 4 znormalizowane fonetycznie zdania) poddano analizie MFCC. Wyniki porównano z parametrami akustycznymi (z grupy jittera, z grupy shimmera, parametrem oceny szumów NHR i współczynnikiem chrypki Yanagihary), parametrem aerodynamicznym (maksymalnym czasem fonacji) i parametrami subiektywnymi (skalą percepcyjną GRBAS i wskaźnikiem niepełnosprawności głosowej VHI). **Wyniki:** Analiza cepstralna wykazała znaczące różnice między grupą badaną a porównawczą, istotne dla współczynników MFCC2, MFCC3, MFCC5, MFCC6, MFCC8, MFCC10, szczególnie dla MFCC6 ( $p < 0,001$ ) oraz dla MFCC8 ( $p < 0,009$ ), co może sugerować ich przydatność kliniczną. Z kolei w grupie badanej MFCC4, MFCC8 i MFCC10 istotnie korelowały z większością zastosowanych parametrów obiektywnych oceny głosu. Ponadto współczynnik MFCC8, który u badanych nauczycielek korelował istotnie z wszystkimi ww. 8 parametrami obiektywnymi, wykazał też istotną zależność z cechą dystynktywną A (asthenity) subiektywnej skali GRBAS, cechującej głos słaby, zmęczony. **Wnioski:** Analiza cepstralna, oparta na wyznaczeniu współczynników MFCC, jest dobrze rokującym narzędziem do obiektywnej diagnostyki dysfonii zawodowych, które bardziej niż inne metody analizy akustycznej odzwierciedla cechy percepcyjne głosu. Med. Pr. 2013;64(6):805–816

**Słowa kluczowe:** kompleksowa ocena głosu, współczynniki cepstralne MFCC, zawodowe zaburzenia głosu

### ABSTRACT

**Background:** Special consideration has recently been given to cepstral analysis with mel-frequency cepstral coefficients (MFCCs). The aim of this study was to assess the applicability of MFCCs in acoustic analysis for diagnosing occupational dysphonia in comparison to subjective and objective parameters of voice evaluation. **Materials and Methods:** The study comprised 2 groups, one of 55 female teachers (mean age: 45 years) with occupational dysphonia confirmed by videostroboscopy and 40 female controls with normal voice (mean age: 43 years). The acoustic samples involving sustained vowels "a" and four standardized sentences were analyzed by computed analysis of MFCCs. The results were compared to acoustic parameters of jitter and shimmer groups, noise to harmonic ratio, Yanagihara index evaluating the grade of hoarseness, the aerodynamic parameter: maximum phonation time and also subjective parameters: GRBAS perceptual scale and Voice Handicap Index (VHI). **Results:** The compared results revealed differences between the study and control groups, significant for MFCC2, MFCC3, MFCC5, MFCC6, MFCC8, MFCC10, particularly for MFCC6 ( $p < 0.001$ ) and MFCC8 ( $p < 0.009$ ), which may suggest their clinical applicability. In the study group, MFCC4, MFCC8 and MFCC10 correlated significantly with the major objective parameters of voice assessment. Moreover, MFCC8 coefficient, which in the female teachers correlated with all eight objective parameters, also showed the significant relation with perceptual voice feature A (asthenity) of subjective scale GRBAS, characteristic of weak tired voice. **Conclusions:** The cepstral analysis with mel

frequency cepstral coefficients is a promising tool for evaluating occupational voice disorders, capable of reflecting the perceptual voice features better than other methods of acoustic analysis. *Med Pr* 2013;64(6):805–816

**Key words:** complex voice assessment, mel-cepstral coefficients, MFCCs, occupational voice disorders

Autorka do korespondencji / Corresponding author: Ewa Niebudek-Bogusz, Klinika Audiologii i Foniatrii, Instytut Medycyny Pracy w Łodzi, ul. św. Teresy 8, 91-348 Łódź, e-mail: ebogusz@imp.lodz.pl  
Nadesłano: 8 października 2013, zatwierdzono: 4 listopada 2013

## WSTĘP

Diagnozowanie zaburzeń głosu, także o podłożu zawodowym, może powodować trudności ze względu na indywidualne właściwości osobnicze fenomenu akustycznego, jakim jest głos – nośnik mowy (1). Prawidłowa funkcja głosowa ma zasadnicze znaczenie we współczesnym społeczeństwie opartym na komunikacji. Jakość głosu odzwierciedla nie tylko stan funkcjonalny krtani, istotny w rozpoznawaniu zawodowych zaburzeń głosu, ale również ogólny stan zdrowia, emocje czy aktualny nastrój (2).

Wielu klinicystów i naukowców zajmujących się głosem podkreśla, że czynność fonacyjna krtani nie może być oceniana za pomocą tylko jednej metody, nawet jeśli zastosujemy nowoczesną metodę instrumentalną z użyciem najnowszych zdobyczy technik przetwarzania cyfrowego dźwięku i obrazu.

Konsensusem międzynarodowym grup badaczy jest stwierdzenie, że głos wymaga też subiektywnej oceny percepcyjnej, tzw. odsłuchowej (3). Jest to najstarsza metoda nieinstrumentalna, od dawna stosowana do określania jakości głosu. Mimo rozwoju zaawansowanych technik diagnostycznych nadal jest powszechnie stosowana w codziennej praktyce laryngologicznej, foniatrycznej i logopedycznej. Ocena percepcyjna głosu dokonywana jest przez badającego lekarza foniatrę czy logopedę-terapeutę głosu na podstawie brzmienia głosu badanego pacjenta.

Obecnie używa się do niej usystematyzowanych skal – najczęściej stosowana jest skala GRBAS (G – grade of hoarseness, stopień chrypki; R – roughness, szorstkość głosu; B – breathiness, głos chuchający; A – asthenic, głos słaby; S – strained, głos napięty) (4).

Inną subiektywną metodą diagnostyczną jest ocena jakości głosu dokonywana przez pacjenta. Wielu autorów podkreśla, że samoocena głosu jest istotnym elementem diagnozy i ma ważny wpływ na dalsze postępowanie terapeutyczne (5,6). Najpowszechniej stosowanym narzędziem samooceny głosu jest kwestionariusz określający wskaźnik niepełnosprawności głosowej (Voice Handicap Index – VHI).

Z innej strony klinicyści zajmujący się zaburzeniami głosu wiedzą, że konieczna jest ewaluacja metod badawczych według wytycznych medycyny opartej na faktach (Evidence Based Medicine) (7) oraz obiektywizacja badań klinicznych głosu. Metodą, która daje możliwość parametryzacji czynności fonacyjnej, jest akustyczna analiza głosu przeprowadzana w oparciu o nowoczesne techniki komputerowe (8,9).

W innym artykule (10) autorzy niniejszej publikacji przedstawili zastosowanie nowej nieliniowej metody analizy akustycznej z wykorzystaniem współczynników cepstralnych (mel-frequency cepstral coefficients – MFCC) w diagnozowaniu zaburzeń głosu o podłożu zawodowym (10). Po klasyfikacji współczynników MFCC głos patologiczny został wykryty z dużą trafnością: z 91-procentową czułością i 83-procentową specyficznością dla głoski ‘a’ oraz dla badanych zdań z czułością i specyficznością w granicach 87–100%.

Celem niniejszej pracy jest ocena współczynników MFCC u pacjentów z dysfoniami zawodowymi w porównaniu z innymi metodami oceny głosu, w tym: oceną percepcyjną głosu, wskaźnikiem niepełnosprawności głosowej według VHI, maksymalnym czasem fonacji i parametrami konwencjonalnej analizy akustycznej.

## MATERIAŁ I METODY

W badaniu wzięły udział 2 grupy kobiet: jako grupa badana – 55 nauczycielek z dysfoniami zawodowymi (średni wiek: 45 lat) oraz jako grupa porównawcza – 40 kobiet z głosem prawidłowym (średni wiek: 43 lata). Grupa badana została wyłoniona na podstawie badania foniatrycznego z laryngowideostroboskopią. Zakwalifikowano do niej nauczycielki z niedomykalnością głośni o etiologii zawodowej. U części z nich (25 osób) występowały też – jako objaw patologicznej kompensacji niewydolności głośni – małe zmiany obrzękowo-przerostowe fałdów głosowych (guzki głosowe, polipy).

W obydwu grupach przeprowadzono badania akustyczne, nagrywając samogłoskę ‘a’ podczas przedłużonej fonacji oraz 4 znormalizowane fonetycznie zda-

nia. Zarejestrowane próbki głosowe poddano analizie akustycznej za pomocą niezależnych metod – analizy cepstralnej w wyznaczaniu współczynników MFCC oraz konwencjonalnej analizy akustycznej. Metodę analizy cepstralnej w wyznaczaniu współczynników MFCC1–MFCC12 oraz ich pochodnych 1- i 2-rzędowych dokładnie opisano w innym artykule (10). Z kolei konwencjonalnej analizy wieloparametrycznej tonu krtaniowego dokonywano za pomocą programu Dia-gnoScope (11) i wyznaczano następujące parametry:

- parametry oceniające względną zmianę częstotliwości (grupa jittera) – jitter, RAP (relative average perturbation – iloraz średniej zmiany częstotliwości) i PPQ (pitch period perturbation quotient – iloraz zmiany okresu częstotliwości);
- parametry oceniające względną zmianę amplitudy (grupa shimmera) – shimmer i APQ (amplitude perturbation quotient – względna zmiana amplitudy ze stałym współczynnikiem);
- parametr względnych pomiarów szumu i zakłóceń – NHR (noise to harmonic rate);
- współczynnik określający nasilenie chrypki według Yanagihary (Yg), uwzględniający zakres komponentu szumowego głosu i jego wpływ na ukształtowanie składowych harmoniczných.

W programie Dia-gnoScope wartość współczynnika chrypki Yanagihary dla każdego okresu podstawowego fonacji wyznaczana jest jako:

$$Yg = 4,0 - 2,5 \times NHQ(F_1, F_0) - 1,25 \times NHQ(F_2, F_0) \quad [1]$$

gdzie:

$F_0$  – chwilowa częstotliwość podstawowa, uśredniona

po 4 najbliższych okresach podstawowych,

$F_1$  – chwilowa częstotliwość formantu pierwszego,

wyznaczona w danym okresie podstawowym,

$F_2$  – chwilowa częstotliwość formantu drugiego,

wyznaczona w danym okresie podstawowym,

$NHQ(F_n, F_0)$  – znormalizowany do przedziału [0;1] stosunek

wartości składowych nieharmonicznych  $(k+1/4) \times F_0$ ,

$(k+1/2) \times F_0$ ,  $(k+3/4) \times F_0$  do harmonicznej  $k \times F_0$  w najbliższym

sąsiedztwie częstotliwości formantu  $F_n$  ( $k = [F_n/F_0]$ ).

Wyznaczany w ten sposób współczynnik Yg przyjmuje wartości rzeczywiste z przedziału 0,25–4,0 – im współczynnik jest wyższy, tym głos bardziej zachrypnięty.

Ponadto w grupie badanej oceniono również głos za pomocą metod subiektywnych – wskaźnika niepełnosprawności głosowej według VHI i skali percepcyjnej GRBAS. Każda nauczycielka dokonała samooceny

własnego głosu, wypełniając kwestionariusz VHI, który wskazuje na stopień zaburzeń głosu według następującej punktacji:

- 0–30 pkt – niewielka niesprawność głosu,
- 31–60 pkt – średnia niesprawność głosu,
- 61–120 pkt – duża niesprawność głosu.

Z kolei foniatra i logopeda oceniali głosy badanych nauczycielek percepcyjnie za pomocą skali GRBAS. Określa ona zaburzenia głosu w stopniu od 0 do 3 przy pomocy 5 parametrów:

- G (grade of hoarseness) – stopień chrypki,
- R (roughness) – szorstkość głosu wynikająca z nieregularności drgań fałdów głosowych,
- B (breathiness) – głos chuchający, będący wynikiem wydobywania się powietrza w czasie fonacji przez niezwartą głośnię,
- A (asthenic) – głos słaby, asteniczny, męczliwy,
- S (strained) – głos napięty, hyperfunkcjonalny.

U badanych nauczycielek mierzono także maksymalny czas fonacji (maximum phonation time – MPT) – parametr aerodynamiczny, który jest prostym, ale wartościowym obiektywnym miernikiem wydolności głosowej. Jego pomiar polegał na określeniu czasu fonowania głoski ‘a’ podczas jak najdłuższego wydechu. Wynik to średnia 3 pomiarów, a norma to 20 s.

W analizie statystycznej, którą przeprowadzono, porównano w grupie badanej i porównawczej parametry oceny obiektywnej głosu, tj. parametr aerodynamiczny MPT, parametry akustyczne z grup jittera i shimmera, NHR, współczynnik Yg oraz nowo zastosowane klinicznie współczynniki MFCC. Wykorzystano do tego celu klasyczne podejście, tj. test t-Studenta, po zbadaniu normalności rozkładu danych testem Shapiro-Wilka. Następnie dokonano oceny korelacji między współczynnikami MFCC a parametrami subiektywnymi i obiektywnymi kompleksowej oceny głosu, stosowanymi do tej pory w Klinice Audiologii i Foniatrii Instytutu Medycyny Pracy im. prof. J. Nofera w Łodzi. Analizę przeprowadzono z wykorzystaniem współczynników korelacji – liniowej Pearsona lub rho Spearmana (konsekwentnie zastosowano odpowiednie testy istotności dla tych współczynników).

W dalszym etapie jako metody statystycznej do oceny związków między MFCC a obiektywnymi parametrami oceny głosu użyto analizy regresji liniowej. Wyznaczony współczynnik kierunkowy prostej potwierdził kierunek związku między analizowanymi parametrami głosu, pozwolił też na ocenę zakresu zmian zmiennej zależnej (konkretnego parametru akustycznej oceny głosu) wraz z jednostkowym wzrostem zmiennej niezależnej (MFCC).

## WYNIKI

Analiza wieloparametryczna tonu podstawowego wykazała, że u badanych nauczycielek z dysfonią zawodową niektóre parametry były istotnie gorsze niż u kobiet z prawidłowym głosem z grupy porównawczej (tab. 1). Wartości różniły się:

- dla shimmera – z istotnością statystyczną:  $p = 0,027$  (8,17 vs 4,06),
- dla współczynnika chrypki Yg – z istotnością statystyczną:  $p = 0,042$  (1,60 vs 1,27),
- dla jittera – na granicy istotności:  $p = 0,057$  (0,973 vs 0,57).

Porównując obiektywny parametr dynamiczny MPT wykazano, że średni maksymalny czas fonacji był także istotnie ( $p = 0,000$ ) krótszy (gorszy) w grupie badanej niż w porównawczej (9,8 s vs 15,6 s). Pozostałe parametry akustyczne nie różniły się istotnie.

Ponadto wynik kwestionariusza samoceny głosu według VHI w badanej grupie wynosił średnio 43 punkty, co wskazuje, że nauczycielki określały swoją niepełnosprawność głosową przeważnie jako średnią. Z kolei mediana wyników skali GRBAS dla całej grupy przyjęła wartości G1R1B2A2S1. Cechę dystynktywną głosu B (głos chuchający) w stopniu 2. określono u 49,1%, czyli prawie połowy nauczycielek, a w stopniu 3. u 7,3%. Głos jako A (asteniczny) w stopniu 2. został określony u 54,5% nauczycielek, a w stopniu 3. u 3,6% nauczycielek.

Następnie porównano wartości średnie poszczególnych współczynników MFCC dla grupy badanej i kontrolnej. Największe różnice wykazano dla współczynników MFCC1–MFCC12 (tab. 2). Porównując wartości poszczególnych współczynników w tej grupie, wykazano istotne różnice dla MFCC2, MFCC3, MFCC5, MFCC6, MFCC8 i MFCC10, zwłaszcza dla MFCC6 ( $p < 0,001$ ) i MFCC8 ( $p < 0,009$ ), co może sugerować ich przydatność kliniczną.

W grupie 1-rzędowych współczynników cepstralnych, tj. MFCC13–MFCC24, tylko współczynnik MFCC19 różnił się istotnie w grupie badanej i porównawczej. Z kolei w grupie pochodnych 2-rzędowych, tj. MFCC25–MFCC36 współczynniki MFCC27, MFCC30, MFCC32 i MFCC36 wykazywały wartości istotnie różne ( $p < 0,05$ ) między grupami. Z kolei dostępna literatura przedmiotu (12) przedstawia pochodne 2-rzędowe MFCC jako parametry o małej wartości klinicznej ze względu na ich duże zindywidualizowanie osobnicze. Z tego powodu do dalszej analizy statystycznej wzięto pod uwagę wartości współczynników MFCC1–MFCC12, dokonując oceny ich korelacji z wynikami innych obiektywnych metod diagnostycznych –

**Tabela 1.** Wartości MPT, grup jittera i shimmera, NHR i Yg w grupie porównawczej i badanej  
**Table 1.** Values of MPT, groups of jitter and shimmer, NHR and Yg, in the control and study groups

Parametr Parameter	M	SD	p
MPT			0,000*
grupa porównawcza / control group	15,584	4,420	
grupa badana / study group	9,812	4,520	
Jitter			0,057
grupa porównawcza / control group	0,566	0,271	
grupa badana / study group	0,973	1,319	
RAP			0,061
grupa porównawcza / control group	0,303	0,163	
grupa badana / study group	0,547	0,806	
PPQ			0,206
grupa porównawcza / control group	0,344	0,158	
grupa badana / study group	0,507	0,807	
Shimmer			0,027*
grupa porównawcza / control group	4,057	1,495	
grupa badana / study group	8,168	11,497	
APQ			0,093
grupa porównawcza / control group	3,650	1,682	
grupa badana / study group	5,381	6,315	
NHR			0,144
grupa porównawcza / control group	2,410	1,183	
grupa badana / study group	3,250	3,411	
Yg			0,042*
grupa porównawcza / control group	1,274	0,609	
grupa badana / study group	1,601	0,807	

MPT – maksymalny czas fonacji / maximum phonation time.

RAP – względna średnia zaburzeń / relative average perturbation.

PPQ – iloraz zmiany okresu częstotliwości / pitch period perturbation quotient.

APQ – względna zmiana amplitudy ze stałym współczynnikiem / amplitude perturbation quotient.

NHR – parametr względnych pomiarów szumu i zakłóceń / noise to harmonic rate.

Yg – współczynnik chrypki Yanagihary / Yanagihara factor.

M – mean / średnia; S – odchylenie standardowe / standard deviation.

p – prawdopodobieństwo w teście t-Studenta / p-value in t-Student test.

\* Różnice istotne statystycznie przy  $\alpha = 0,05$  / Statistically significant differences;  $\alpha = 0,05$ .

parametru aerodynamicznego MPT, parametrów akustycznych – jittera, RAP, PPQ, shimmera, APQ – parametru NHR i współczynnika Yg.

Tabela 3 przedstawia ww. korelacje dla grupy kontrolnej porównawczej – najczęściej korelował z porównywanymi parametrami obiektywnymi współczynnik MFCC6. Z kolei w tabeli 4. przedstawiono korelacje dla grupy badanej. Na podstawie wyników analizy korelacji w grupie nauczycielek należy podkreślić, że współczynnik MFCC8 wykazywał istotną zależność ze wszystkimi 8 wymienionymi parametrami obiektywnymi oceny głosu. Stwierdzono o dwie mniej istotne zależności dla współczynnika MFCC10, natomiast współczynnik MFCC4 istotnie korelował z 5 badanymi parametrami obiektywnymi.

**Tabela 2.** Wartości MFCC w grupie porównawczej i badanej  
**Table 2.** Values of MFCC in the control and study groups

Parametr Parameter	M	SD	p
MFCC1			0,897
grupa porównawcza / control group	2,696	0,229	
grupa badana / study group	2,742	0,266	
MFCC2			0,041*
grupa porównawcza / control group	2,598	0,244	
grupa badana / study group	1,924	0,215	
MFCC3			0,012*
grupa porównawcza / control group	1,276	0,216	
grupa badana / study group	0,489	0,218	
MFCC4			0,319
grupa porównawcza / control group	1,884	0,135	
grupa badana / study group	2,098	0,164	
MFCC5			0,034*
grupa porównawcza / control group	0,137	0,066	
grupa badana / study group	0,100	0,086	
MFCC6			0,001*
grupa porównawcza / control group	0,066	0,133	
grupa badana / study group	0,601	0,086	
MFCC7			0,237
grupa porównawcza / control group	1,446	0,124	
grupa badana / study group	1,625	0,084	
MFCC8			0,009*
grupa porównawcza / control group	0,570	0,075	
grupa badana / study group	1,289	0,087	
MFCC9			0,606
grupa porównawcza / control group	0,389	0,081	
grupa badana / study group	0,327	0,087	
MFCC10			0,025*
grupa porównawcza / control group	0,397	0,068	
grupa badana / study group	0,097	0,079	
MFCC11			0,211
grupa porównawcza / control group	0,686	0,080	
grupa badana / study group	0,817	0,104	
MFCC12			0,158
grupa porównawcza / control group	0,158	0,099	
grupa badana / study group	0,351	0,093	

MFCC – współczynniki cepstralne / mel-frequency cepstral coefficients.  
 Inne objaśnienia jak w tabeli 1 / Other abbreviations as in Table 1.

Należy przypomnieć, że wartości MFCC8 i MFCC10 różniły się istotnie między grupą badaną a porównawczą, natomiast dla MFCC4 takiej różnicy nie zaobserwowano. Istotna z punktu widzenia klinicznego jest ujemna korelacja MFCC8 i MFCC10 z maksymalnym czasem fonacji MPT – im krótszy (gorszy) czas fonacji, tym średnio wyższe (gorsze) te współczynniki.

**Tabela 3.** Współczynniki korelacji dla MFCC i maksymalnego czasu fonacji, parametrów grup jittera i shimmera, NHR i współczynnika Yg w grupie porównawczej\*  
**Table 3.** Coefficients of correlation between MFCC and maximum phonation time, groups of jitter and shimmer, NHR and Yg in the control group\*

Parametr Parameter	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC10	MFCC11	MFCC12
MPT	0,094	0,189	-0,216	-0,011	-0,060	0,137	0,220	-0,125	-0,086	-0,023	-0,159	-0,085
Jitter	-0,230	<b>0,531</b> ↑	<b>0,431</b> ↑	0,281	0,076	<b>0,482</b> ↑	<b>0,615</b> ↑	0,190	-0,021	0,054	-0,293	<b>-0,503</b> ↓
RAP	-0,209	<b>0,491</b> ↑	<b>0,361</b> ↑	0,240	0,133	<b>0,426</b> ↑	<b>0,554</b> ↑	0,160	-0,028	0,118	-0,223	<b>-0,430</b> ↓
PPQ	-0,231	<b>0,558</b> ↑	<b>0,422</b> ↑	0,276	0,034	<b>0,467</b> ↑	<b>0,580</b> ↑	0,123	-0,048	0,043	<b>-0,364</b> ↓	<b>-0,499</b> ↓
Shimmer	-0,217	0,080	0,080	0,308	0,003	<b>0,435</b> ↑	<b>0,424</b> ↑	0,283	0,120	0,068	<b>-0,324</b> ↓	<b>-0,315</b> ↓
APQ	-0,245	0,092	0,238	<b>0,363</b> ↑	0,041	<b>0,543</b> ↑	<b>0,563</b> ↑	0,289	0,278	0,020	-0,268	<b>-0,349</b> ↓
NHR	-0,115	-0,175	0,088	0,271	-0,148	<b>0,319</b> ↑	0,295	<b>0,385</b> ↑	0,077	0,000	-0,142	-0,053
Yg	<b>-0,352</b> ↓	<b>0,357</b> ↑	<b>0,535</b> ↑	<b>0,467</b> ↑	0,078	<b>0,445</b> ↑	<b>0,607</b> ↑	0,213	-0,019	0,031	<b>-0,434</b> ↓	<b>-0,468</b> ↓

MFCC – współczynniki cepstralne / mel-frequency cepstral coefficients.

\* Zastosowano współczynnik korelacji liniowej Pearsona. Pogrubienia oznaczają, że zależność jest statystycznie istotna / Pearson's linear correlation coefficient was used. Bold means significant correlation.

↑ Korelacja dodatnia / positive correlation, ↓ korelacja ujemna / negative correlation.

Skróty jak w tabeli 1 i 2 / Other abbreviations as in Tables 1 and 2.

**Tabela 4.** Współczynniki korelacji dla parametrów MFCC i maksymalnego czasu fonacji, parametrów grup jittera i shimmera, NHR i współczynnika Yg w grupie badanej  
**Table 4.** Coefficients of correlation between MFCC and maximum phonation time, groups of jitter and shimmer, NHR and Yg in the study group

Parametr Parameter	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC10	MFCC11	MFCC12
MPT	0,063	0,275	-0,217	-0,234	-0,057	-0,034	-0,162	-0,414 ↓	-0,199	-0,364 ↓	-0,279	-0,174
Jitter	-0,023	-0,183	0,033	<b>0,305 ↑</b>	0,132	0,265	0,146	<b>0,389 ↑</b>	0,140	<b>0,346 ↑</b>	0,039	-0,055
RAP	-0,047	-0,188	-0,002	<b>0,311 ↑</b>	0,156	0,270	0,150	<b>0,376 ↑</b>	0,128	<b>0,350 ↑</b>	0,030	-0,043
PPQ	-0,073	-0,299	-0,086	<b>0,334 ↑</b>	0,152	<b>0,329 ↑</b>	0,102	<b>0,368 ↑</b>	0,215	<b>0,406 ↑</b>	0,140	0,003
Shimmer	0,085	-0,275	-0,130	0,224	0,033	0,172	-0,012	<b>0,344 ↑</b>	0,282	<b>0,316 ↑</b>	0,138	0,081
APQ	0,060	<b>-0,315 ↓</b>	-0,095	0,292	0,055	0,229	0,032	<b>0,391 ↑</b>	0,299	<b>0,350 ↑</b>	0,157	0,047
NHR	0,072	<b>-0,551 ↓</b>	0,044	<b>0,366 ↑</b>	-0,079	0,181	0,021	<b>0,490 ↑</b>	<b>0,428 ↑</b>	0,287	0,252	0,091
Yg	-0,229	-0,211	0,290	<b>0,612 ↑</b>	<b>0,445 ↑</b>	0,202	0,177	<b>0,440 ↑</b>	0,072	0,225	0,052	-0,197

Objaśnienia jak w tabeli 3 / Abbreviations as in Table 3.

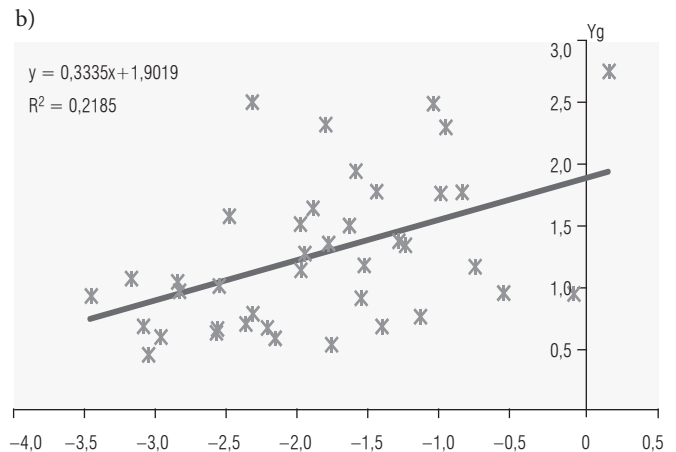
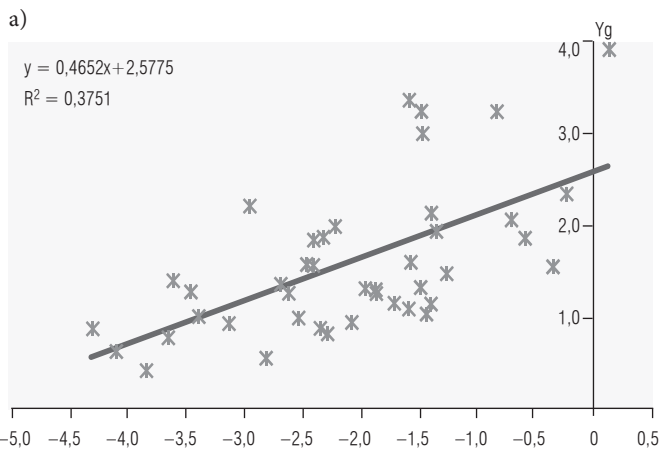
Zaobserwowane korelacje między zmiennymi zostały potwierdzone także za pomocą analizy regresji liniowej. Rycina 1. przedstawia przebieg zależności między parametrem akustycznym określającym nasilenie chrypki – tj. współczynnikiem Yanagihary (Yg) – a tymi współczynnikami MFCC, dla których odnotowano istotne współczynniki regresji. Jak wykazano wcześniej (wykorzystując współczynnik korelacji liniowej Pearsona), w przypadku grupy badanej Yg pozostaje w istotnej, a przy tym stosunkowo dość silnej relacji z MFCC4, MFCC5 i MFCC8 – im wyższa wartość wymienionych MFCC, tym średnio wyższy poziom Yg. Potwierdzają to dane zaprezentowane na rycinie 1.

Stopień wyjaśnienia zmienności Yg zmiennością MFCC jest stosunkowo wysoki – sięga ok. 20% dla relacji Yg z MFCC5 i MFCC8 oraz prawie 40% dla MFCC4. W przypadku grupy porównawczej kierunek korelacji jest taki sam, jednak jej siła jest znacznie mniejsza – stopień wyjaśnienia Yg zmianami MFCC5 i MFCC8 sięga odpowiednio: 0,6% i 4%. Lepiej jest w przypadku MFCC4 – współczynnik determinacji sięga ok. 22%. Dla osób zdrowych wpływ MFCC5 i MFCC8 na Yg nie jest statystycznie istotny, co potwierdza zarówno test istotności współczynnika korelacji, jak i współczynnika regresji.

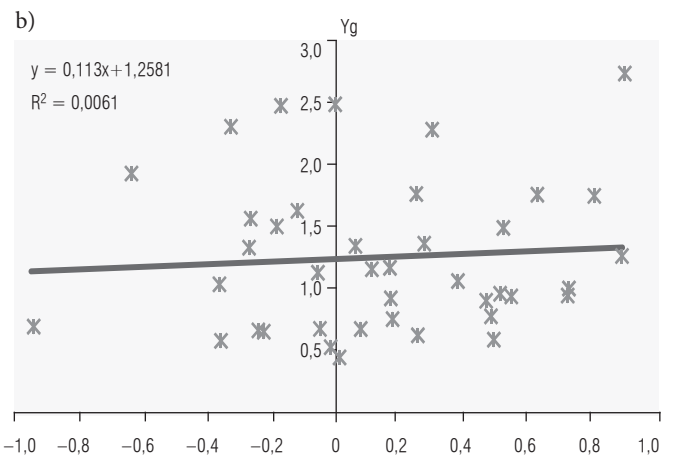
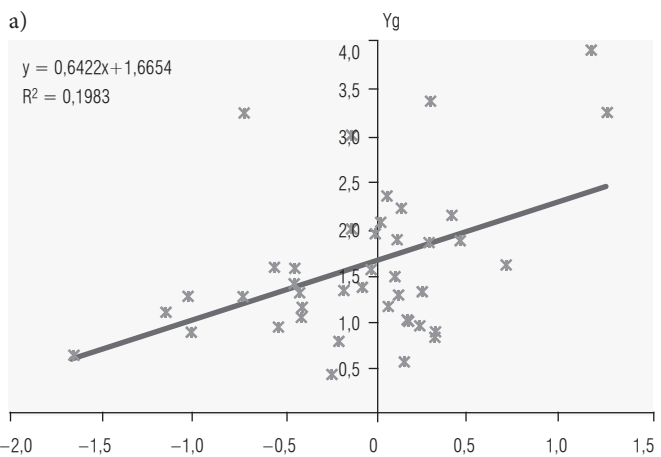
Analogicznie, rycina 2. prezentuje graficznie ocenę regresji liniowej dla parametru jittera. W grupie badanej zaobserwowano istotną zależność dla jittera oraz MFCC4, MFCC8 i MFCC10. Natomiast dla parametru NHR, oceniającego względne pomiary szumu i zakłóceń, współczynnik regresji wartości wskazywał na zależność z MFCC2, MFCC4, MFCC8 i MFCC9 (ryc. 3). Stwierdzono przy tym istotność współczynnika regresji dla omawianych par zmiennych.

Oceniając korelację współczynników MFCC1–MFCC12 z parametrami subiektywnymi oceny głosu, nie zaobserwowano tak istotnych korelacji jak dla badanych parametrów obiektywnych opisanych powyżej. W tabeli 5. przedstawiono korelacje współczynników MFCC1–MFCC12 z wynikami samooceny głosu VHI, a tabela 6. obrazuje korelacje MFCC1–MFCC12 z wynikami skali percepcyjnej GRBAS. Ciekawą obserwacją jest to, że współczynnik MFCC8, który istotnie koreluje ze wszystkim badanymi parametrami obiektywnymi, wykazuje też istotną zależność z parametrem A (asthenity – głos słaby) subiektywnej skali GRBAS, która określa „męczliwość” głosu. Ten parametr, określający głos jako asteniczny, zaobserwowano w stopniu 2. lub 3. w ponad połowie grupy badanej.

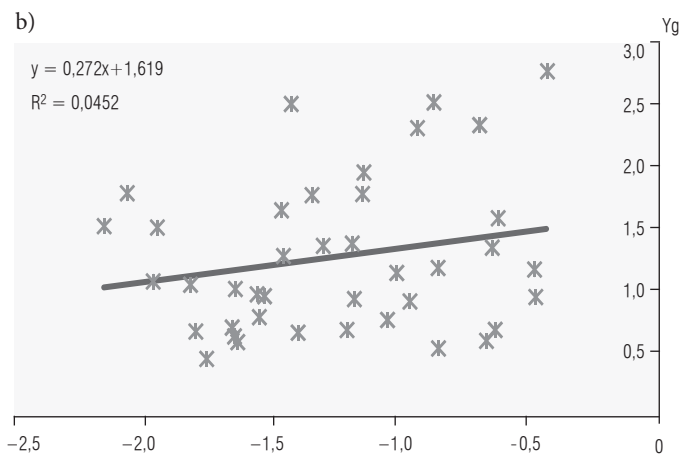
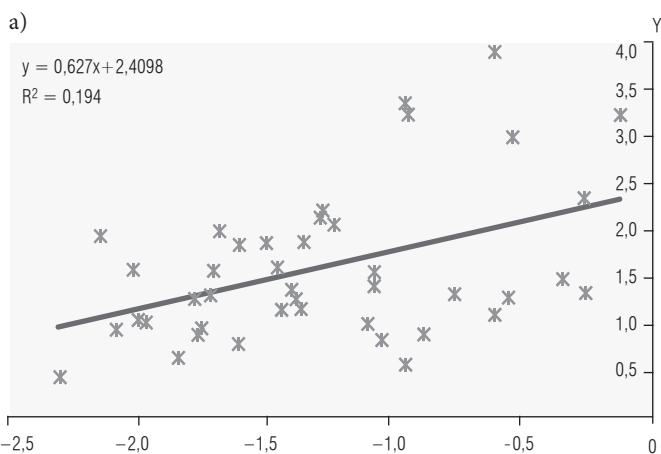
## MFCC4



## MFCC5



## MFCC8



a) grupa badana / study group.

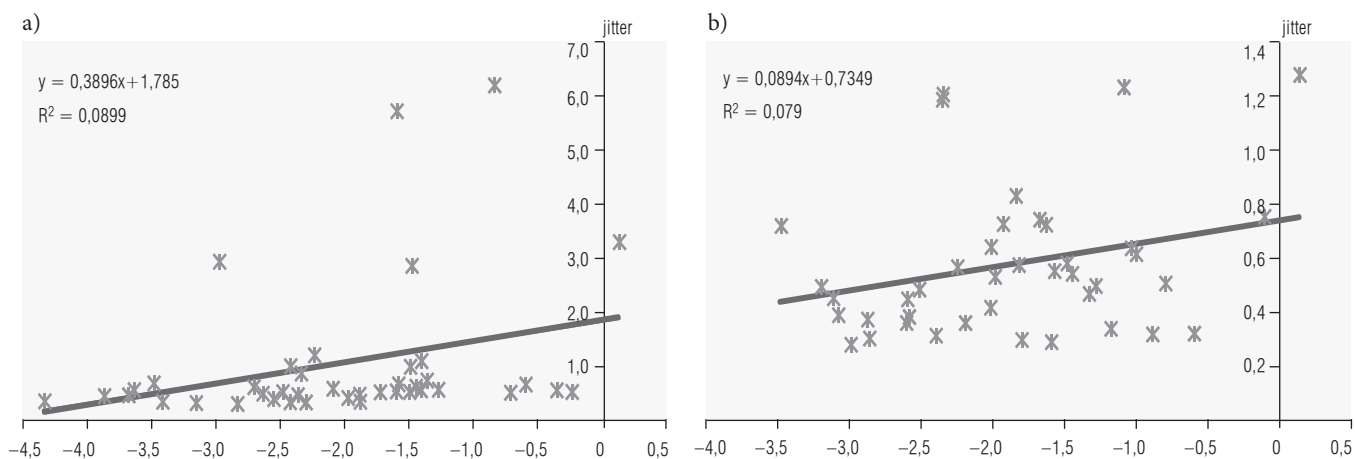
b) grupa porównawcza / control group.

y – zmienna zależna / dependent variable: Yg.

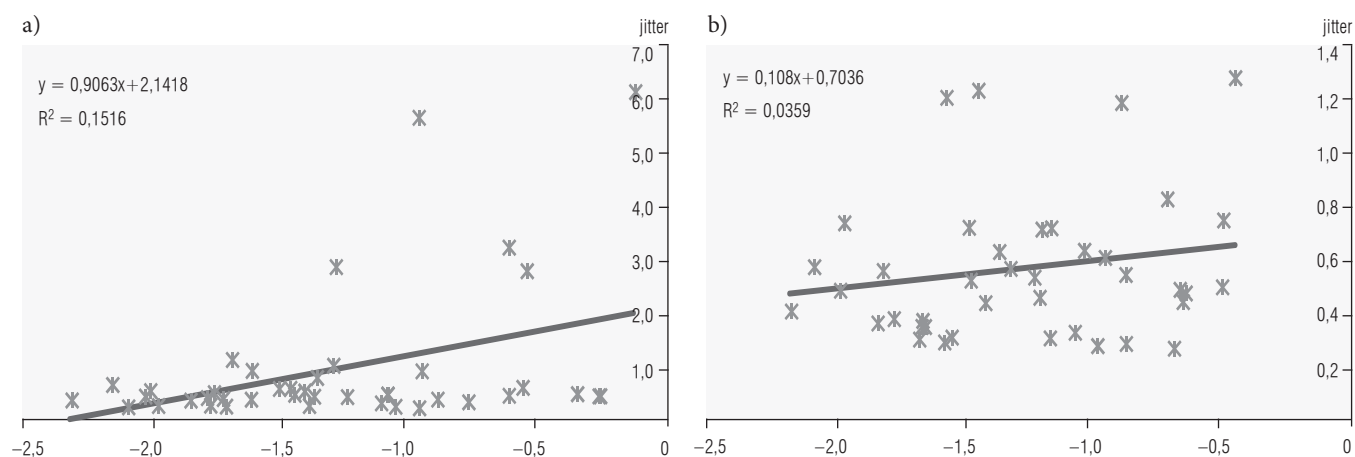
$R^2$  – współczynnik determinacji / coefficient of determination.

**Ryc. 1.** Ocena zależności między parametrem Yg a współczynnikami MFCC – istotne statystycznie wyniki analizy regresji liniowej  
**Fig. 1.** Relations between Yg parameter and MFCC coefficients – statistically significant results of linear regression analysis

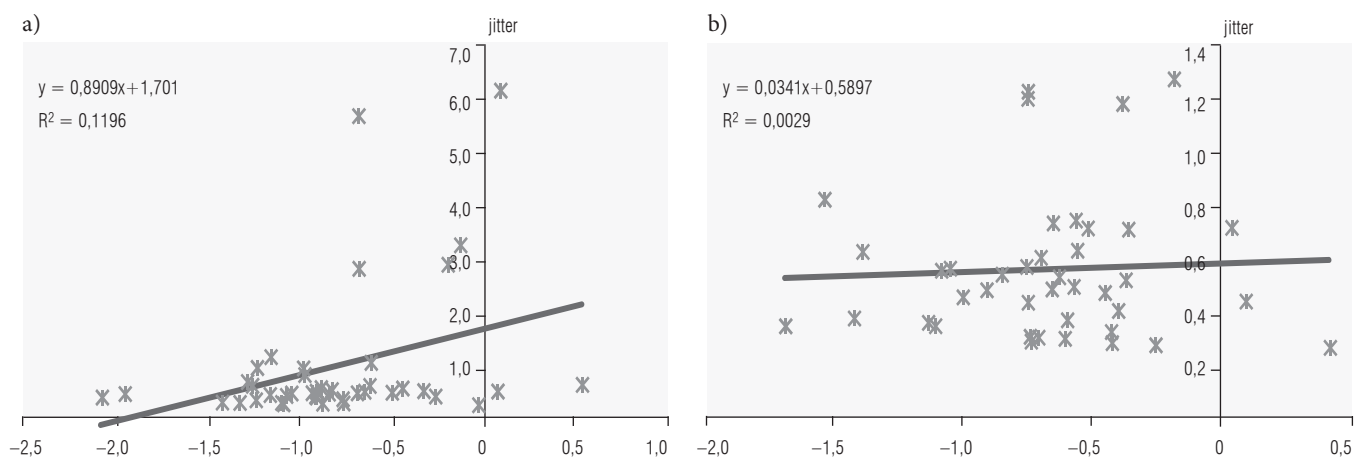
## MFCC4



## MFCC8



## MFCC10



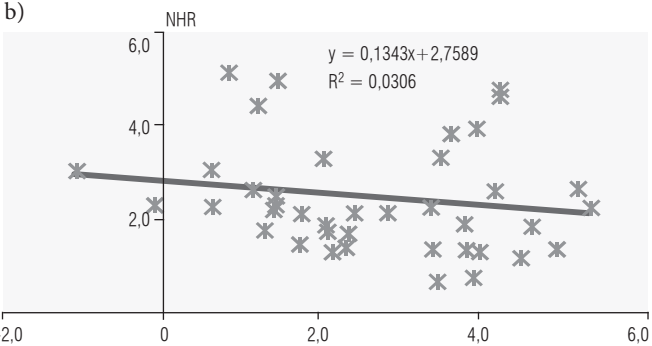
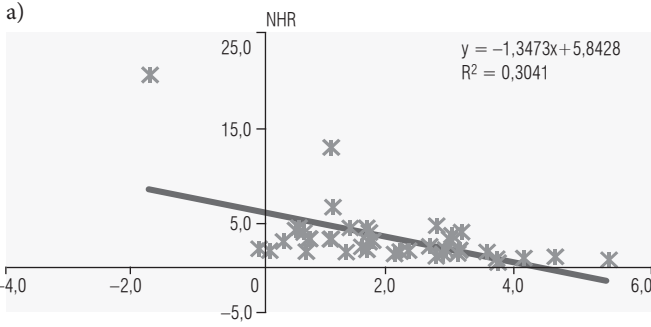
y – zmienna zależna / dependent variable: jitter.

Inne objaśnienia jak w rycinie 1 / Other abbreviations as in Figure 1.

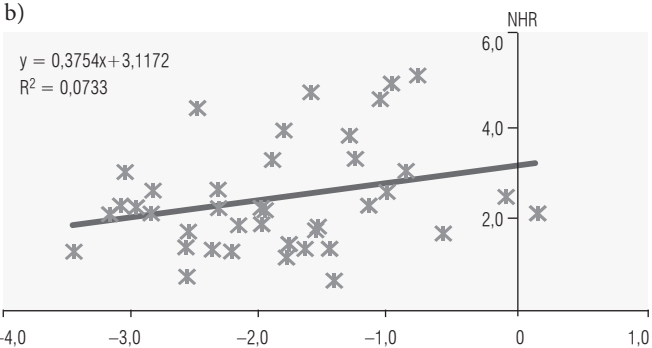
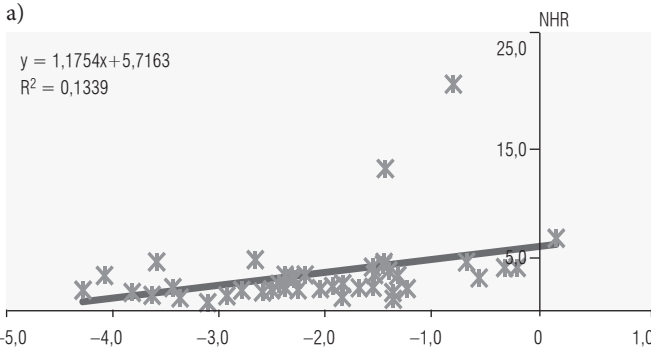
**Ryc. 2.** Ocena zależności między parametrem jitter a współczynnikami MFCC – istotne statystycznie wyniki analizy regresji liniowej  
**Fig. 2.** Relations between jitter parameter and MFCC coefficients – statistically significant results of linear regression analysis



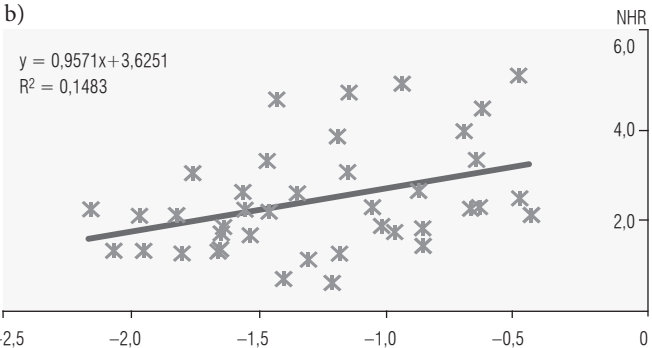
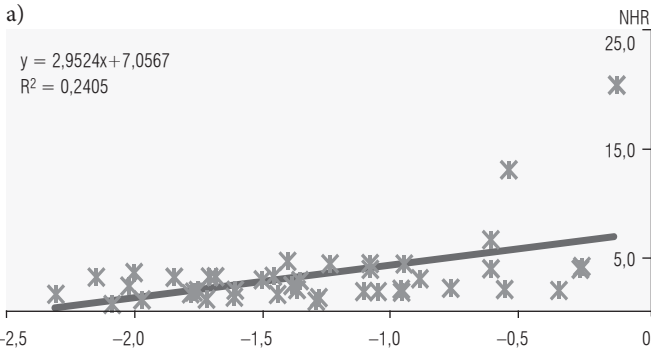
MFCC2



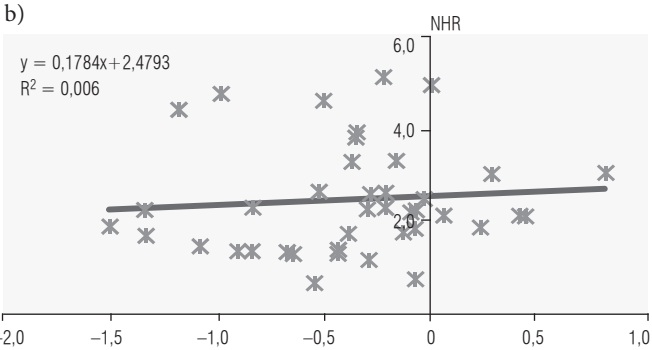
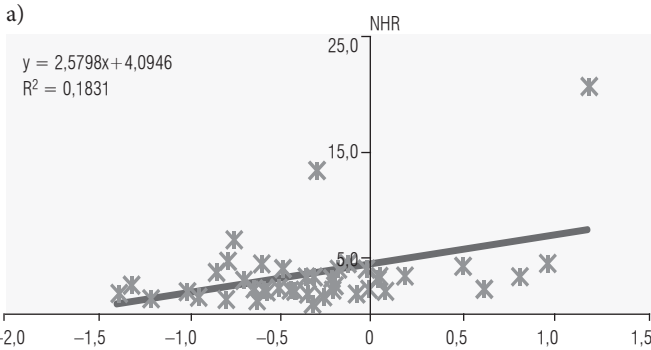
MFCC4



MFCC8



MFCC9



y - zmienna zależna / dependent variable: NHR.  
Inne objaśnienia jak w rycinie 1 / Other abbreviations as in Figure 1.

Ryc. 3. Ocena zależności między parametrem NHR a parametrami MFCC – istotne statystycznie wyniki analizy regresji liniowej  
Fig. 3. Relations between NHR parameter and MFCC coefficients – statistically significant results of linear regression analysis

**Tabela 5.** Współczynniki korelacji między VHI a MFCC w grupie badanej  
**Table 5.** Coefficients of correlation between VHI and MFCC in the study group

VHI	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC10	MFCC11	MFCC12
Funkcjonalne / Functional	-0,232	-0,186	0,091	0,175	-0,002	-0,067	0,138	0,227	0,263	0,271	<b>0,325 ↑</b>	0,188
Emocjonalne / Emotional	-0,146	-0,077	0,237	0,174	-0,202	-0,112	0,110	0,178	0,120	0,282	0,291	0,220
Fizyczne / Physical	0,077	0,211	0,159	-0,030	-0,141	0,049	-0,065	0,017	-0,138	0,106	-0,056	0,067
Ogółem / Total	-0,122	-0,031	0,187	0,128	-0,132	-0,055	0,076	0,166	0,102	0,255	0,224	0,185

VHI – skala niepełnosprawności głosowej / Voice Handicap Index.  
 Inne objaśnienia jak w tabeli 3 / Other abbreviations as in Table 3.

**Tabela 6.** Współczynniki korelacji między skalą GRBAS a MFCC w grupie badanej  
**Table 6.** Coefficients of correlation between the GRBAS scale and MFCC in the study group

Skala Scale	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC10	MFCC11	MFCC12	MFCC1
G	-0,095	-0,164	0,071	-0,055	-0,050	<b>-0,335 ↓</b>	-0,050	0,157	0,212	0,132	<b>0,369 ↑</b>	0,213	0,123
R	-0,172	0,177	0,212	-0,152	0,030	-0,113	0,030	0,104	-0,031	-0,157	-0,077	-0,209	-0,156
B	0,114	<b>-0,304 ↓</b>	-0,165	-0,022	-0,167	-0,007	-0,167	0,047	0,142	0,275	0,127	0,294	<b>0,402 ↑</b>
A	0,127	-0,182	-0,158	0,079	-0,170	-0,082	-0,170	<b>-0,323 ↓</b>	0,156	0,249	0,105	0,240	0,300
S	-0,121	-0,057	-0,038	-0,022	0,262	-0,246	0,262	0,068	0,103	-0,066	0,072	0,066	0,083

GRBAS: G – stopień chryplki / grade of hoarseness, R – szorstkość głosu / roughness, B – głos chuchający / breathiness, A – głos słaby / asthenic, S – głos napięty / strained.  
 Inne objaśnienia jak w tabeli 3 / Other abbreviations as in Table 3.

## OMÓWIENIE

W pracy przedstawiono ocenę przydatności współczynników MFCC w diagnozowaniu dysfonii zawodowych w porównaniu z parametrami subiektywnymi i obiektywnymi kompleksowej oceny, tj. wskaźnikiem niepełnosprawności głosowej – VHI, skalą percepcyjną GRBAS, parametrem aerodynamicznym MPT, parametrami akustycznymi z grupy jittera i shimmera oraz NHR i współczynnikiem chrypki Yg. Opisywana metoda analizy cepstralnej została początkowo opracowana do rozpoznawania głosów mówców, ale ostatnio wzrasta zainteresowanie nią w zakresie wykorzystania jej do celów klinicznych (12). Wynika to także z tego, że stosowana powszechnie analiza akustyczna oparta na wyznaczaniu parametrów oceniających względną zmianę częstotliwości (jitter) i amplitudy (shimmer) nie zawsze spełnia oczekiwania klinicystów, ponieważ jej wyniki często nie korelują z cechami oceny percepcyjnej głosu (7).

Ważną cechą analizy cepstralnej jest możliwość wydzielenia z sygnału głosu składowych, które są związane z działaniem generatora dźwięku, czyli krtani, a dokładniej jej środkowego piętra – głośni (8,13). Ma to szczególne znaczenie w ocenie zaburzeń głosu etiologicznie związanych z patologią głośni, np. w przypadku schorzeń narządu głosu o podłożu zawodowym (2,14).

W prezentowanych badaniach za pomocą analizy cepstralnej oceniano głosy kobiet z patologią głośni o podłożu zawodowym z grupy badanej oraz prawidłowe głosy kobiet z grupy porównawczej. Porównując współczynniki cepstralne w obu grupach, wykazano istotne różnice wartości, zwłaszcza dla współczynników MFCC1–MFCC12. Podobne wyniki obserwacji dotyczące wartości diagnostycznych MFCC otrzymali badacze hiszpańscy, analizując 117 głosów patologicznych i 53 prawidłowych (12,15).

W dalszych badaniach opisywanych w niniejszej pracy wykazano istotne korelacje niektórych współczynników MFCC1–MFCC12 z parametrami konwencjonalnej analizy akustycznej, tj. jittera, RAP, PPQ, shimmera, APQ i NHR. Te znamienne korelacje zostały potwierdzone także w analizie regresji liniowej. Stwierdzono też istotne zależności kilku MFCC z parametrem aerodynamicznym MPT i współczynnikiem chrypki Yanagihary.

Z kolei współczynnik MCFF8, który wykazywał istotne korelacje ze wszystkimi ww. obiektywnymi parametrami oceny głosu, znamienne korelował z parametrem A skali subiektywnej GRBAS, określającym głos słaby, asteniczny. Jest to charakterystyczna cecha głosu, często występująca w dysfonii zawodowej (14). Łatwa

męczliwość głosu i zaniki głosu po dłuższym mówieniu są objawem hypotonii głośni, rozwijającej się najczęściej na skutek przeciążania głosowego. Ta choroba krtani występuje często u osób pracujących głosem, a jej istotą jest osłabienie mięśni przywodzących i napinających fałdy głosowe (16). W grupie badanej ww. percepcyjna cecha głosu występowała w natężeniu znacznym lub ciężkim u ok. 60% nauczycielek. Należy zwrócić uwagę, że korelacja współczynnika MFCC8 z parametrem A skali GRBAS pozwala przypuszczać, że współczynnik ten może być predyktorem głosu astenicznego.

Inni autorzy potwierdzają, że zaletą analizy cepstralnej jest większa w porównaniu z konwencjonalną analizą akustyczną możliwość odniesienia wyników do percepcyjnej oceny głosu i cech dystynktywnych (17). Autorzy amerykańscy za pomocą analizy cepstralnej z wykorzystaniem wyznaczonej wartości maksimum cepstrum (cepstral peak prominence – CPP) i jego odchylenia standardowego badali głosy patologiczne, których cechą dystynktywną była szorstkość (roughness – R) lub głos chuchający (breathiness – B) (18). Porównując 14 głosów „szorskich” i 14 głosów „chuchających” z 14 głosami prawidłowymi, ww. badacze wykazali dobrą korelację zastosowanego wskaźnika cepstralnego z badanymi cechami percepcyjnymi głosu (18).

Z kolei Peterson i wsp. łącząc wyniki analizy cepstralnej i spektralnej, utworzyli wskaźnik CSID (cepstral spectral index of dysphonia – cepstralno-spektralny wskaźnik dysfonii) i wykorzystali go do oceny głosów grupy osób o zróżnicowanej patologii głośni (19). Stwierdzili, że CSID dobrze korelował ze stopniem dysfonii ocenianym odsłuchowo przez doświadczonych patologów głosu. Ponadto ww. autorzy podkreślają, że diagnostyka akustyczna zaburzeń głosu za pomocą analizy cepstralnej jest bardziej praktyczna, ponieważ umożliwia badanie dłuższych próbek głosu, np. zdań czy dłuższych wypowiedzi, a nie tylko izolowanej głoski (tak jak w stosowanej powszechnie krótkookresowej analizie tonu podstawowego). Badania poprzednio prezentowane (10) przez autorów niniejszej publikacji również wykazały przydatność tej metody w analizowaniu zdań, potwierdzając doniesienia opisywane przez Peterson i wsp. (19).

## WNIOSKI

Analiza cepstralna głosu oparta na wyznaczaniu współczynników MFCC jest dobrze rokującym obiektywnym narzędziem do diagnostyki zaburzeń głosu o podłożu zawodowym. Bardziej bowiem niż konwencjonalne

metody analizy akustycznej odzwierciedla cechy percepcyjne głosu. Współczynnik MFCC8, który u kobiet z zawodowymi zaburzeniami głosu korelował istotnie ze wszystkimi badanymi parametrami obiektywnymi, wykazał też istotną zależność od percepcyjnej cechy dysfunkcyjnej A (asthenity) subiektywnej skali GRBAS, charakterystyczną dla głosu w dysfoniach zawodowych.

## PIŚMIENNICTWO

- Rubin J.S., Satalof R.T., Korovinn G.S. [red.]: *Diagnosis and treatment of voice disorders*. Thomson Delmar Learning, New York 2003
- Dejonckere P.H.: *Occupational voice-care and cure*. Kugler Publications, Hague 2001
- Schutte H.K.: *Proceedings of 7th Pan European Voice Conference Aug 29 – Sept 1 2007, Groningen, The Netherlands*. Voice Research Lab, Groningen 2007
- Wiskirska-Woźnica B.: *Kompleksowa ocena głosu w schorzeniach organicznych i czynnościowych krtani [rozprawa habilitacyjna]*. AM, Poznań 2002
- Fulljames N., Harris S.: Voice outcome measures: correlations with patients' assessment of their condition and the effectiveness of voice therapy. *Log. Phon. Vocol.* 2006;31: 23–35, <http://dx.doi.org/10.1080/14015430500295749>
- De Jong F.: An introduction to the teacher's voice in a biopsychosocial perspective. *Folia Phoniatr. Logop.* 2010;62:5–8, <http://dx.doi.org/10.1159/000239058>
- Maryn Y., Roy N., de Bodt M., van Cauwenberge P., Corthals P.: Acoustic measurement of overall voice quality: A meta-analysis. *J. Acoust. Soc. Am.* 2009;126(5): 2619–2634, <http://dx.doi.org/10.1121/1.3224706>
- Wszolek W.: *Analiza dźwięków mowy do celów medycznych*. W: Tadeusiewicz R., Augustyniak P. [red.]: *Podstawy inżynierii biomedycznej*. Tom I. Wydawnictwo AGH, Kraków 2009, ss. 229–252
- Choi S.H., Lee J., Sprecher A.J., Jiang J.J.: The effect of segment selection on acoustic analysis. *J. Voice* 2012;26(1): 1–7, <http://dx.doi.org/10.1016/j.jvoice.2010.10.009>
- Niebudek-Bogusz E., Grygiel J., Strumiłło P., Śliwińska-Kowalska M.: Nieliniowa analiza akustyczna w ocenie zawodowych zaburzeń głosu. *Med. Pr.* 2013;64(1):29–35, <http://dx.doi.org/10.13075/mp.5893/2013/0004>
- DiagNova Technologies: *Analiza głosu* [cytowany 3 października 2013]. Adres: <http://www.diagnova.pl/pl/index.php/oferta/produkty/analiza-akustyczna-glosu>
- Godino-Llorente J.I., Gomez-Vilda P., Cruz-Roldan F., Blanco-Velasco M., Fraile R.: Pathological likelihood index as a measurement of the degree of voice normality and perceived hoarseness. *J. Voice* 2009;24(6):667–677, <http://dx.doi.org/10.1016/j.jvoice.2009.04.003>
- Samlan R.A., Story B.H.: Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling. *J. Speech Lang. Hear. Res.* 2011;54:1267–1283, [http://dx.doi.org/10.1044/1092-4388\(2011/10-0195\)](http://dx.doi.org/10.1044/1092-4388(2011/10-0195))
- Śliwińska-Kowalska M. [red.]: *Zasady orzekania o przedyspozycjach zawodowych do pracy w narażeniu na hałas i nadmierny wysiłek głosowy*. Warszawa 2011
- Arias-Londoño J.D., Godino-Llorente J.I., Markaki M., Stylianou Y.: On combining information from modulation spectra and mel-frequency cepstral coefficients for automatic detection of pathological voices. *Log. Phon. Vocol.* 2011;36(2):60–69, <http://dx.doi.org/10.3109/14015439.2010.528788>
- Obrębowski A.: *Narząd głosu i jego znaczenie w komunikacji społecznej*. Wydawnictwo Naukowe Uniwersytetu Medycznego im. Karola Marcinkowskiego w Poznaniu, Poznań 2008
- Sáenz-Lechón N., Fraile R., Godino-Llorente J.I., Fernández-Baillo R., Osmá-Ruiz V., Gutiérrez-Arriola J.M. i wsp.: Towards objective evaluation of perceived roughness and breathiness: An approach based on mel-frequency cepstral analysis. *Log. Phon. Vocol.* 2011;36(2):52–59, <http://dx.doi.org/10.3109/14015439.2010.517551>
- Lowell S.Y., Colton R.H., Kelley R.T., Mizia S.A.: Predictive value and discriminant capacity of cepstral and spectral-based measures during continuous speech. *J. Voice* 2013;27(4):393–400, <http://dx.doi.org/10.1016/j.jvoice.2013.02.005>
- Peterson E.A., Roy N., Awan S.N., Merrill R.M., Banks R., Tanner K.: Toward validation of the cepstral spectral index of dysphonia (CSID) as an objective treatment outcomes measure. *J. Voice* 2013;27(4):401–410, <http://dx.doi.org/10.1016/j.jvoice.2013.04.002>