

Wybrane problemy optymalizacji pracy głosem w zawodach artystycznych i pedagogicznych

Voice work optimization in artistic and pedagogical professions – selected issues

Andrzej W. Mitas^{1AC-E}, **Agata Szkielkowska**^{2A-CEF}, **Mirostaw Czak**^{1AEF},
Julia M. Nycz^{3BE}

¹ Politechnika Śląska, Katedra Informatyki Medycznej i Sztucznej Inteligencji, Wydział Inżynierii Biomedycznej, Zabrze

² Instytut Fizjologii i Patologii Słuchu, Światowe Centrum Słuchu, Klinika Audiologii i Foniatrii, Warszawa/Kajetany

³ Institut für Biomedizinische Technik und Informatik, Technische Universität Ilmenau, Niemcy

Wkład autorów:
A Projekt badania
B Gromadzenie danych
C Analiza danych
D Interpretacja danych
E Przygotowanie pracy
F Przegląd literatury
G Gromadzenie funduszy

Streszczenie

W artykule przedstawiono wybrane zagadnienia pracy głosem w aspekcie monitorowania jego stanu. W kontekście dotychczasowych (statycznych) metod analizy statusu aparatu mowy i śpiewu przedstawiono znane metody pomiaru biomarkerów fizjologicznych służących do diagnozowania potencjalnych patologii. Zaprezentowano także metodę badania i konstrukcji systemu pomiarowego, który może być wykorzystany do bieżącej oceny stanu narządów mowy i śpiewu, tak by w czasie rzeczywistym możliwe było podjęcie adekwatnych działań, polegających zasadniczo na czasowym zaniechaniu wydawania dźwięków lub ograniczenia ich intensywności. W materiale wyeksponowano przewagę systemu pomiarowo-kontrolnego, stanowiącego techniczne wsparcie dla beneficjenta, nad systemem autokontroli – zawodnym przeważnie z przyczyn socjologicznych lub z powodu niedostatecznie czytelnych sygnałów ostrzegawczych, wysyłanych przez nasz organizm.

Słowa kluczowe: systemy analizy głosu • systemy wspomaganie diagnostyki • inżynieria biomedyczna • telemedycyna

Abstract

The article presents selected issues of working with voice in the aspect of monitoring its health. In the context of the present (static) methods of analyzing the status of the speech and singing apparatus, some known methods of measuring physiological biomarkers, typically used to diagnose potential pathologies, have been presented. On this background, a proposal for a test method and a measurement system design is presented, which can be used for the current assessment of the condition of the speech and singing organs, so that in real time it would be possible to take adequate activities, essentially consisting in temporarily ceasing to emit sounds or limiting their intensity. The material emphasizes the advantage of the measurement and control system, which provides technical support for the beneficiary, over the self-control system, which is unreliable mostly for sociological reasons or because of insufficiently clear warning signals sent by our body.

Key words: systems for voice analysis • diagnostic support system • biomedical engineering • telemedicine

Adres autora: Andrzej W. Mitas, Wydział Inżynierii Biomedycznej, Katedra Informatyki Medycznej i Sztucznej Inteligencji, Politechnika Śląska, Roosevelta 40, 41-800, Zabrze; email: andrzej.mitas@polsl.pl

Wprowadzenie

Niekończące się pasmo sukcesów, jakie każdy z nas odnotowuje na co dzień w swojej zawodowej praktyce jest w ujęciu biocybernetycznym sygnałem dodatniego sprzężenia zwrotnego prowadzącego nieuchronnie do przesterowania obiektu, czyli ciała i umysłu. Wobec nowych możliwości trudno jest zaakceptować opcję zredukowania własnych wymagań, nawet wówczas, gdy obciążenie stresem zawodowym przynosi negatywne skutki zdrowotne. Wykonywanie każdej pracy wiąże się z wykorzystaniem elementów ciała, w tym mózgu. Funkcjonujące w organizmie naturalne narzędzia, tj. mięśnie i mózg, są w systemie wczesnego ostrzegania ignorowane, ponieważ zwykle nie dostrzegamy związku ekonomicznego nadmiernej eksploatacji i wysokich kosztów odnowy, nie liczymy także kosztów leczenia uszkodzonego aparatu głosu i mowy. Praktyka nauczycielska nie potwierdza zjawiska zaplanowanej relaksacji, a komasacja zajęć wykładowych na uczelni nie należy do rzadkości. Wiedza o gospodarowaniu zasobami naturalnymi, nawet w ogólnym wymiarze, jest nadal słabo spopularyzowana. Jej pozyskanie jest przeważnie trudne, a respektowanie zasad jest problematyczne, ponieważ wymaga rezygnacji z doraźnych korzyści.

Aspekty głosu

Dźwięk jest prawdopodobnie najważniejszym antropologicznie sygnałem służącym do przetrwania w kontekście stanów alarmowych. Słyszymy dźwięki nawet we śnie i reagujemy na informację o zagrożeniu. Stereofonię i zjawisko Dopplera wykorzystujemy w codzienności, np. uczestnicząc w niebezpiecznym z natury ruchu drogowym, wyróżniając przestrzennie tor i zbliżanie lub oddalanie się pojazdu. W domowej ciszy rzask zamka w drzwiach wejściowych brzmi niekiedy jak wystrzał i wywołuje często nieadekwatne emocje.

Głos człowieka wpisuje się w świat pochłanianych przez środowisko dźwięków, a jego brzmienie stanowi rzeczywistość równoległą do semantyki. Znaczenie leksykalne słów nie musi ostatecznie korespondować z brzmieniem, choć instynkt nam podpowiada, jaki jest ton lekkiej przygany i czym się odróżnia od brzmienia słów pocieszenia czy nagrody. Trudno jest przy tym jasno zdefiniować, jakie spełnić trzeba kryteria w dziedzinie częstotliwościowej i czasowej, by fala akustyczna była matematycznie interpretowalna w sensie wyrażania treści.

Głos i mowa, w tym ich artystyczne przedłużenie, czyli śpiew, do komunikacji są wykorzystywane zawsze, gdy chcemy wyrazić nasze wrażenia, oczekiwania czy potrzeby. Są również podstawowymi elementami estetyzacji rzeczywistości, zapewne jako antropologiczna pochodna prawdopodobieństwa sukcesu reprodukcyjnego. W ujęciu holistycznym można przyjąć, że jakość komunikacji akustycznej jest zależna od brzmienia głosu. Prosta egzemplifikacją tego tematu jest przewaga konkurencyjna jednych piosenkarzy czy lektorów nad innymi.

Anatomia aparatu głosu i mowy jest gatunkowo zależna, a możliwość uogólnienia opisu świadczy o istotnych podobieństwach strukturalnych. Krtań przyjmuje swoją postać w 32 tygodniu życia płodowego i jest gotowa do

pracy w chwili urodzin. W okresie noworodkowym i niemowlęcym krtań jest nie w pełni uwapniona, delikatna i niewielka. Skala głosu dziecka w pierwszych okresach życia jest również wąska i niezależna od płci. W okresie dzieciństwa zmienia się częstotliwość podstawowa głosu (do 10 roku życia obniża się do 265 Hz u dziewczynek i do 235 Hz u chłopców). W okresie dojrzewania następuje przemiana dziecięcej krtań w dojrzały organ. Głos obniża się u mężczyzn o 1 oktawę, a u kobiet o 2–3 tony. U osoby dorosłej krtań znajduje się na wysokości 5 lub 6 kręgu szyjnego i jest w pełni ukształtowana. Po przebytych mutacji głos charakteryzuje się indywidualną barwą i siłą, a jego skala się poszerza. Brzmienie głosu różni się w miarę powiększania się organów służących do generowania mowy i silnie zależy od spersonalizowanej odmienności anatomicznej rezonatorów.

Dorosłość określa zazwyczaj okres stałych parametrów wykorzystywania głosu przez człowieka. Ważne jest, aby podobnie jak w teorii niezawodności okres normalnego użytkowania trwał możliwie długo, a intensywność występowania uszkodzeń była niewielka. Ostatnim fragmentem tej krzywej jest okres starzenia, gdy nieunikniona destrukcja poszczególnych frakcji aparatu mowy prowadzi do pogorszenia jakości wydawanych dźwięków, zarówno w mowie, jak i w śpiewie. Zadaniem współczesnej nauki jest optymalizacja normalnej eksploatacji, by koszty etapu starzenia były minimalne.

Wybrane biomarkery w systemie ochrony aparatu mowy

Akwizycja danych wydaje się współcześnie zadaniem „łatwym i przyjemnym”. Gdyby tak w istocie było, to nagrywanie piosenki nie wymagałoby studia, lecz można byłoby to zrobić w warunkach domowych z wykorzystaniem sprzętu powszechnego użytku, który (jak to niekiedy wynika ze „specyfikacji”, czyli danych przedstawionych bez odpowiednich referencji) umożliwi rejestrację nawet z gigantyczną szybkością 96 kHz. Tymczasem w praktyce nagranie zwykłej, prostej kompozycji z wykorzystaniem takiego urządzenia doraźnie destabilizuje poczucie estetyki. Z tego powodu sprzęt profesjonalny musi być, niestety, złożony i kosztowny, a jego przykłady przedstawiono w dalszej części. Propozycję systemu, wykorzystywanego do własnych badań, także z udziałem probantów z obszaru zawodowego muzyki profesjonalnej, również scharakteryzowano pokrótce w dalszej części materiału celem opisu sposobu monitorowania w warunkach rzeczywistych dostępności sprzętu i oprogramowania, które daje szansę samoobserwacji i wczesnej profilaktyki.

Celem ogólnego przedstawienia złożoności modelu matematycznego poniżej opisano kilka spośród wręcz setek biomarkerów (zwanymi często w medycynie parametrami), które można, a niekiedy nawet należy uwzględnić w analizie ilościowo-jakościowej sprawności aparatu mowy i śpiewu. Ta lista najpowszechniej używanych w różnych analizach wybranych biomarkerów zawiera wielkość zarówno w dziedzinie czasowej, jak i częstotliwościowej.

1. Wskaźnik modulacji amplitudy (ang. *shimmer*) – powszechnie stosowany deskryptor głosu fizjologicznego lub patologicznego; wahania amplitudy dźwięku związane są z ciśnieniem podgłośniowym, zmieniającym się

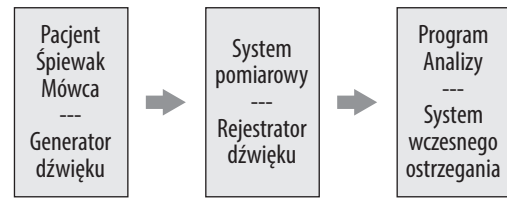
- pod wpływem zmian amplitudy drgań i napięcia fałdów głosowych.
- Wskaźnik zmienności częstotliwości podstawowej (ang. *jitter*) – równie powszechny wskaźnik opisu jakości głosu; brak kontroli drgań strun głosowych implikuje jego zmienność; w przypadku śpiewu „w pobliżu” ustalonej częstotliwości wyznacza się *jitter* uśredniony.
 - Wskaźnik zmiany amplitudy – względne odchylenie standardowe amplitud próbek wskazujących na początki okresów; używa się tego markera do oceny wahań krótkotrwałych lub długotrwałych oraz regularnych lub incydentalnych.
 - Wskaźnik zmiany częstotliwości podstawowej – względne odchylenie standardowe ze zbioru F0; eksponuje poziom zmian częstotliwości podstawowej badanego sygnału.
 - HNR (ang. *harmonic to noise ratio*) – służy do analizy udziału składowych harmonicznym sygnału w stosunku do szumu; wyznaczany z funkcji autokorelacji w całkowitym przebiegu sygnału.
 - NHR (ang. *noise to harmonic ratio*) – opisuje relację szumu i składnika harmonicznego sygnału.
 - Środek ciężkości widma gęstości mocy – dobry deskryptor ostrości głosu, wskazujący czy przeważają wysokie, czy niskie formanty.
 - Rozrzut widma gęstości mocy (ang. *spread spectrum*) – deskryptor do oceny rozkładu energii sygnału w stosunku do jego środka ciężkości, co pozwala wyróżnić sygnały szumowe i dźwięczne.

Złożoność sposobu opisu dźwięku stanowiącego mowę lub śpiew jest oczywista, a powyższe definicje kilku z zestawu biomarkerów pełnią jedynie rolę ilustracji. Zdecydowanie bardziej skomplikowane jest ich matematyczne modelowanie umożliwiające porównywanie i analizę o charakterze diagnostycznym.

Aspekty biologiczne i techniczne monitorowania głosu

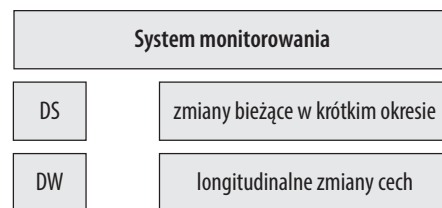
W artykule przedstawiono system optymalizacji w aspekcie ergonomii pracy, przy czym ograniczono zbiór badanych wielkości do tych, które eksponują szerokość pasma i wydajność amplitudową w wyznaczonych zakresach częstotliwości. **Rycina 1** ilustruje uproszczoną zasadę działania takiego systemu, w którym główną rolę odgrywa system pomiarowy. Współcześnie może to być rejestrator oparty na wbudowanym w smartfon przetworniku analogowo-cyfrowym, którego rozdzielczość jest czasami nadmiernym powodem do dumy, a szybkość próbkowania wielokrotnie większa od realnych potrzeb użytkownika. Dyskusja o parametrach tego rodzaju przetwarzania jest potencjalnym przedmiotem odrębnego artykułu metrologicznego. Na działanie prezentowanego tu systemu nie mają realnego wpływu takie parametry jak np. nieliniowość całkowita czy stabilność temperaturowa.

Na podstawie profesjonalnych autorskich badań, realizowanych w Instytucie Fizjologii i Patologii Słuchu, oraz danych literaturowych ocenia się, że pasmo użytkowe mówcy ogranicza się do kilku kiloherców (kHz). Do śpiewu potrzebne są te, jak też wyższe częstotliwości, ale ograniczenie audiogramu do 8 kHz w badaniu podstawowym, 10 kHz w praktyce radiokomunikacji rozsiewczej



Rycina 1. System monitorowania stanu aparatu głosu i mowy – opracowanie własne

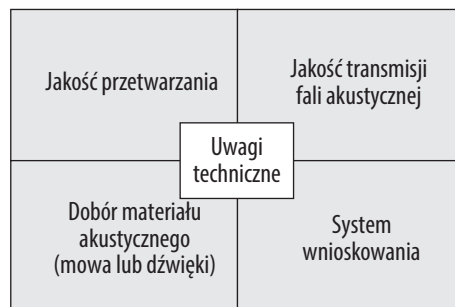
Figure 1. Voice and speech status monitoring system – own elaboration



Opis: DW – dane wolnozmiennie; DS – dane szybkozmiennie.

Rycina 2. Rodzaje zmian głosu – opracowanie własne

Figure 2. Types of voice changes – own elaboration



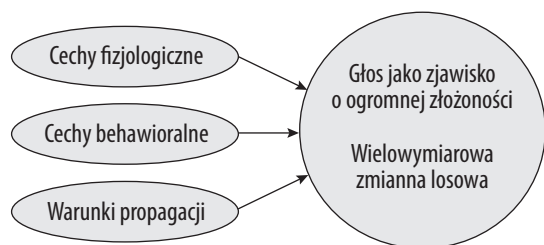
Rycina 3. Czynniki warunkujące monitorowanie głosu

– opracowanie własne

Figure 3. Factors determining voice monitoring – own elaboration

czy 12–14 kHz dla wieku dalece przedemerytalnego powoduje, że częstotliwość próbkowania na poziomie 28 kHz jest wystarczająca do oceny stanu aparatu mowy i głosu. Powszechnie 44,1 kHz wystarcza więc z nawiązką.

Głos poddawany analizie jest wytwarzany przez człowieka w trakcie procesu fizjologicznego. Na przebieg tego procesu mają wpływ czynniki występujące doraźnie, jak i te longitudinalne, których oddziaływanie jest postrzegane jako skutek procesów inwolucyjnych. Odpowiednio należy więc monitorować stan wyczerpania, objawiający się obniżoną jakością emisji lub pogorszeniem funkcjonowania aparatu mowy i głosu, nawet wówczas, gdy doraźnie zmiany te nie są eksponowane lub ich uwidocznienie jest maskowane przez odpowiedni trening (**rycina 2**).



Rycina 4. Głos jako wielowymiarowa zmienna losowa – opracowanie własne

Figure 4. Voice as a multidimensional stochastic variable – own elaboration

Przetwarzanie i analiza głosu to zagadnienia wieloaspektowe. Problem w dużym uproszczeniu zilustrowano na rycinie 3. Wynik w postaci cyfrowej jest oczywiście łatwy do przechowywania, ale w pierw musi być przetworzony z postaci analogowej, co potencjalnie jest źródłem niedokładności i musi być wykonane z należytą starannością. Dane muszą zostać przekazane do przetwarzania, a wynik

końcowy zależy również od jakości transmisji fali głosowej (mechanicznej) do rejestratora. Nie bez znaczenia jest dobór materiału dźwiękowego. Dopiero wówczas, gdy jakość składowych jest dobra, można myśleć o poprawności systemu wnioskowania.

Celem analizy danych następującej po pomiarze jest poszukiwanie określonych informacji. Obecnie wykonuje się to zarówno sprawdzonymi metodami algorytmicznymi, dopuszczalnymi w medycynie heurystykami, ale również poprzez stochastyczne penetrowanie zasobów występującej pod nazwą sztucznej inteligencji. Na rycinie 4 zilustrowano problem analizy głosu ludzkiego, który może być postrzegany w trzech aspektach, przy czym rozróżnienie wpływu cech fizjologicznych i behawioralnych nierzadko jest zadaniem samym w sobie. Sumarycznie warto głos potraktować jako wielowymiarową zmienną losową, z należytym szacunkiem podchodząc do entropii informacji. W istocie dopiero zastosowanie poważnego sprzętu obliczeniowego umożliwi realną analizę cech, przy czym za punkt wyjścia należy przyjąć konieczność opracowania biologiczno-medycznego opisu, zwanego systemem objaśniającym, uprawniającemu do podejmowania decyzji terapeutycznych.

Tabela 1. Narzędzia stosowane do monitorowania głosu

Table 1. Tools applicable for voice monitoring

Nazwa systemu	Producent	Dedykowany HW	Cechy funkcjonalne	Klasa systemu
KAY CSL 4500b	Pentax Medical	tak	Rozbudowana wersja systemu (hardware i software) Computerized Speech Lab (CSL). Sprzęt dedykowany do podłączenia wysokiej klasy mikrofonów oraz wejść liniowych (4 linie), pozwala uzyskać dynamikę większą o 20–30 dB w stosunku do najlepszych kart dźwiękowych w komputerach. Łącznie 4 wyjścia dźwięku: na słuchawki oraz głośniki. Dedykowana karta interfejsu w PC do połączenia z modułem audio – praca tylko ze stacjonarnymi komputerami PC. Pasma przenoszenia sygnału od 0 do 88 kHz; 11 aplikacji do różnorodnych analiz głosu i dźwięku, w tym MDVP pozwalający uzyskać 34 biomarkery głosowe.	profesjonalny, do diagnostyki klinicznej
Visi-Pitch	Pentax Medical		Tańsza wersja hardware łączona z komputerami za pomocą USB. Pasma przenoszenia od 5 Hz do 24 kHz. Pozwala na pracę z przenośnymi komputerami typu notebook. Dwa wejścia mikrofonowe lub liniowe. Software jak w CSL 4500b.	profesjonalny, do diagnostyki klinicznej
Praat	University of Amsterdam	nie	Darmowy software stosowany przede wszystkim w analizach do celów naukowych. Umożliwia analizę spektralną, formantową, wyznaczenie biomarkerów oraz wiele form przekształceń dźwięku, a także syntezę mowy.	do analiz akademickich
Oxford Wave Research	Vocalize	nie	Software do biometrii głosu używany m.in. innymi w laboratoriach śledczych policji. Przeznaczony przede wszystkim do analizy treści mowy oraz rozpoznawania mówcy. W analizie używana jest jedna z dwóch metod: analiza formantowa oraz analiza oparta na MFCC. Działa w połączeniu z oprogramowaniem BioMetric.	profesjonalny do analiz kryminalistycznych
Oxford Wave Research	Spectrum View	nie	Wizualizacja w czasie rzeczywistym spektrogramu rejestrowanego przez mikrofon smartfonu lub tabletu pracującego pod kontrolą systemu operacyjnego IOS (iPhone, iPad itp.). Pozwala analizować częstotliwości słyszalne lub niesłyszalne. Dźwięki można nagrywać i przekazywać za pomocą e-mail do dalszej analizy.	półprofesjonalna, personalna
Centre for Digital Music, Queen Mary University of London	Sonic Visualizer	nie	Aplikacja typu open source na Windows, Linux i Mac; umożliwia wizualizację nagranych dźwięków w postaci spektrogramów oraz wyświetlenia spektrum dźwięku	personalna

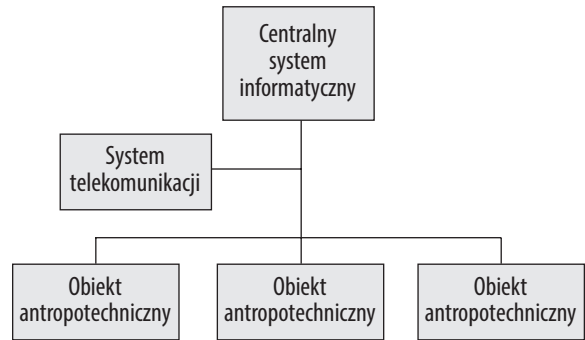
Opis/Note: HW – HardWare; CSL – Computerized Speech Lab; MDVP – Multidimensional Voice Program; MFCC – Mel-Frequency Cepstral Coefficient.

Narzędzia medyczne i laboratoryjne do monitorowania głosu

Od wielu lat w laryngologii i foniatrii dostępne są systemy pomiarowe, zwłaszcza w formie oprogramowania, nierzadko ze wsparciem specjalistycznego sprzętu metrologicznego. Przegląd i skróconą charakterystykę takich rozwiązań przedstawiono w tabeli 1. Podstawowe cechy funkcjonalne umożliwiają pobieżną kwalifikację do stosowności w celu przeprowadzenia diagnozy zaburzeń funkcjonowania aparatu głosu i mowy.

AVM i ergonomia pracy głosem

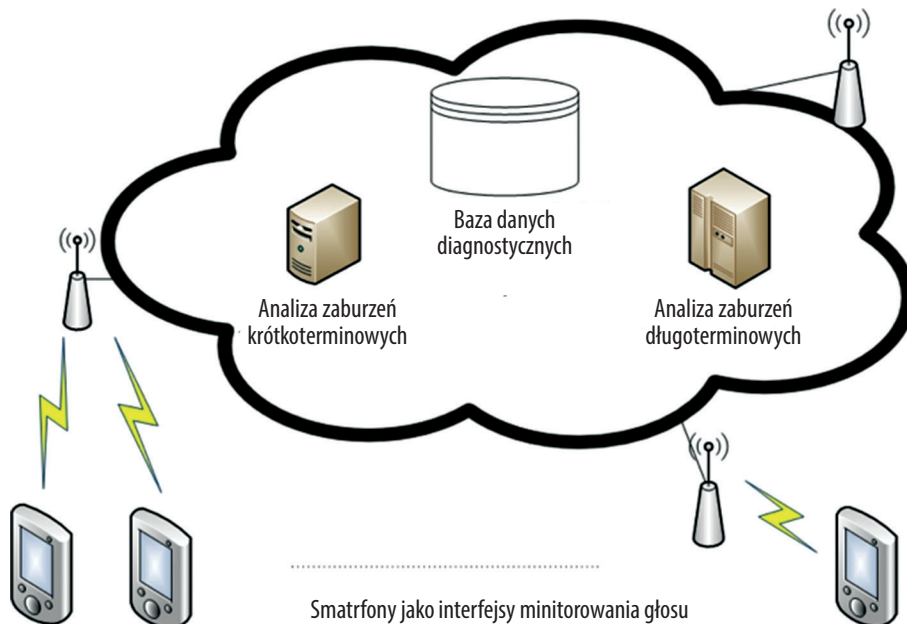
Wykorzystanie głosu przez człowieka, mimo znacznego postępu ergonomii pracy, jest nadal zjawiskiem niekontrolowanym. Sygnałem zwrotnym świadczącym o potencjalnych zaburzeniach jest dolegliwość bólowa, której może towarzyszyć zaburzenie pierwotnego brzmienia głosu. Niestety często występuje też zjawisko maskowania funkcjonalnego, polegającego na korekcie sygnału akustycznego (wielu śpiewaków wykorzystuje metody „usprawnienia” głosu w przypadku niedyspozycji zdrowotnych). Łatwo domyślić się, że takie zachowanie, powielane bez umiaru, prowadzić może do poważnych dysfunkcji aparatu głosu i mowy. W obliczu często występującego przymusu ekonomicznego i społecznego zalecenia wczesnej profilaktyki nie są realizowane. Pragnienie awansu zachęca do maskowania niedostatku, choć ogólnie wiadomo, jak trudno jest przywrócić sprawność organu, którego prawidłowe funkcjonowanie zostało poważnie zaburzone. Powagę tego problemu dostrzega środowisko lekarzy laryngologów i foniatorów. Koszty sanacji poważnych zaniedbań są bowiem z reguły wielokrotnie wyższe od obciążeń związanych z działaniami prewencyjnymi.



Rycina 5. Struktura systemu AVM – opracowanie własne
Figure 5. Structure of the AVM system – own elaboration

Inicjatywa prac badawczych doprowadziła do podjęcia wspólnych prac przez środowisko lekarskie i inżynierów z Politechniki Śląskiej ukierunkowanych na system wczesnego ostrzegania przed przeciążeniem pracy głosem.

Proponowany system antropotechnicznego systemu monitorowania głosu (ang. *anthropotechnical voice monitoring system*, AVM) opiera się na analizie wybranych biomarkerów oraz ocenie na podstawie złożonego modelu matematycznego kondycji emisyjnej pacjenta w czasie pracy. Drugą funkcjonalnością jest wyznaczanie linii trendów zmian niezauważalnych doraźnie, a stanowiących długofalowo poważne zagrożenie dla pacjenta. Podstawowym założeniem funkcjonalnym systemu pomiarowego jest poręczność i odporność na uchybienia eksploatacyjne. Współczesne systemy metrologiczne mają postać obiektów zwanych antropotechnicznymi, ponieważ jako urządzenia techniczne służą człowiekowi do



Rycina 6. Przykładowe rozwiązanie techniczne systemu pomiarowego wraz z terminalami metrologicznymi użytkownika – opracowanie własne

Figure 6. An example of a technical solution of a measuring system with user's metrological terminals – own elaboration

realizacji określonych zadań. Obiekty takie rozmieszczone są w rozległej sieci teleinformatycznej, zapewniającej łączność (dziś przeważnie bezprzewodową, dla wygody użytkownika) z systemem informatycznym ulokowanym w tzw. chmurze, czyli z wirtualnym środowiskiem archiwizacji i przetwarzania danych. Możliwość akwizycji gigantycznej ilości informacji pozwala na użycie stochastycznych metod poszukiwania wiedzy, czyli sztucznej inteligencji. Na **rycynie 5** przedstawiono schemat takiej struktury systemu AVM, której główną cechą jest przeniesienie złożonych obliczeń do wydajnego przetwarzania chmurowego, dzięki czemu w rękach użytkownika pozostaje czytelny terminal wyposażony w adekwatny rejestrator dźwięku (smartfon).

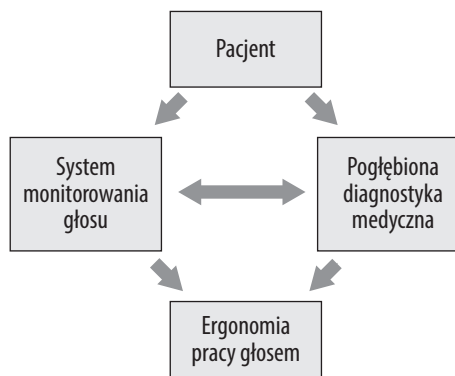
Można przyjąć, że system pomiarowy AVM, służący do spersonalizowanego monitorowania zaburzeń głosu, zwłaszcza w aspekcie obciążenia zawodowego, będzie miał postać przedstawioną na **rycynie 6**.

Użycie sprzętu powszechnego użytku (typu smartfon) jest oczywiście uwarunkowane jakością wbudowanego przetwornika analogowo-cyfrowego, za pomocą którego rejestrowany jest dźwięk. Możliwość zaprogramowania parametrów przetwarzania przez dyktafon jest oczywiście również niezbędna, by osiągnąć zadaną szybkość i rozdzielczość konwersji. W kontekście obecnych trendów nie przewiduje się w tym względzie trudności technicznych.

Algorytm optymalizacji pracy głosem

Oczekiwanym wynikiem działania projektowanego systemu pomiarowego jest informacja o bieżącym wyczerpaniu aparatu głosu i mowy. Pacjent pracujący głosem ma zawsze możliwość pozyskania szczegółowej wiedzy o jakości wydawanych dźwięków. Przeważnie diagnozy dokonuje on z wykorzystaniem własnego odczucia bólu (informacji płynących z receptorów bólowych), ale na tej podstawie nie zawsze podejmuje odpowiednią decyzję. Można przyjąć, że osoba eksploatująca głos znacznie poważniej podejździe do swojej niedyspozycji głosowej, kiedy otrzyma wyraźny sygnał – adekwatny do aktualnego stanu narządu głosu – informujący o progu przekroczenia granicy wydolności głosowej. Dlatego w odniesieniu do głosu tak ważna jest powszechna wiedza o profilaktyce i ergonomii pracy głosem.

Rejestracja przekroczenia granicy wydolności głosowej działa na wyobraźnię użytkownika znacznie silniej niż świadomość permanentnego, lecz niedestruktywnego przekraczania dopuszczalnych wartości (oczywiście chodzi tu o typową reakcję, jest bowiem wielu pacjentów, którzy do rozmaitych badań okresowych i bieżącego automonitorowania podchodzą z należytą starannością). Zakłada się więc, że na podstawie zagregowanej informacji użytkownik podejmie decyzję o doraźnym odpoczynku. Przywrócenie właściwej gotowości do pracy pozwoli na uniknięcie przeciążenia i przekroczenia granicy wydolności głosowej. Opierając się na wynikach pracy systemu pomiarowego użytkownik może podjąć decyzję o wykonaniu pogłębionej diagnostyki medycznej, o ile zachodzi uzasadnione podejrzenie dolegliwości przewlekłej. Schematycznie ilustruje to **rycina 7**.



Rycina 7. Schemat działań prewencyjnych – opracowanie własne
Figure 7. Scheme of preventive actions – own elaboration

Podsumowanie i wnioski

Artysta muzyk, pracownik dydaktyczny, nauczyciel czy po prostu człowiek pracujący głosem powinien dysponować odpowiednim oprzyrządowaniem do obserwacji stanu swojego najcenniejszego narzędzia, czyli aparatu głosu i mowy. Umiejętne wykonywanie zawodu wymaga wiedzy z zakresu ergonomii, by unikać zużycia narzędzi i stanu wyczerpania. Z reguły człowiek dostrzeże niekorzystne zjawiska dopiero wówczas, gdy trzeba dokonywać naprawy, która zwykle jest kosztowna. Wcześniejsze starania o wymiarze prewencyjnym skutkują przeważnie znacznie dłuższym okresem poprawnej eksploatacji.

W przypadku emisji głosu nie jest to ani łatwe, ani oczywiste, ponieważ przyczyny dysfunkcji mogą mieć charakter zarówno fizjologiczny, jak i behawioralny. Wielowymiarowość takiej zmiennej losowej uniemożliwia precyzyjne autonomiczne sterowanie. Dlatego też konieczne jest realne wspomaganie oceny poprawności funkcjonowania aparatu głosu i mowy. Wymaga to opracowania czytelnego sygnału, w szczególności składającego do doraźnego lub długookresowego odpoczynku. Wynikiem takiej świadomej relaksacji ma być przywrócenie funkcjonalności i możliwość kontynuacji pracy. Warto wykorzystać w tym przypadku informację o sobie agregowaną przez długi czas, jak również informację o podobnych pacjentach z kompatybilnej grupy wiekowej i zawodowej, ponieważ ostatecznie taka wielka liczba danych umożliwia zastosowanie stochastycznych metod poszukiwania wiedzy. Użycie sztucznej inteligencji jest uzasadnione realną trudnością dokonania analizy przyczynowo-skutkowej w przypadku słabo określonych wielowymiarowych zmiennych losowych. Proponowany w tekście AVM jest przeznaczony do monitorowania zjawiska biologicznego i umożliwi wsparcie decyzji o pogłębionej diagnostyce medycznej w przypadku stwierdzenia nieprawidłowości. Uwalnia to od konieczności definiowania systemu objaśniającego, potencjalnie uprawniającego do diagnozowania lub terapii. System AVM jest więc realnie jedynie „inteligentnym” wzmacniaczem w pętli biocybernetycznego sprzężenia zwrotnego, w której pacjent ma szansę na zauważenie informacji o spadku własnej bieżącej wydolności oraz o zaburzeniu trajektorii długookresowej charakterystyki głosu, sugerującym postępujące, acz drobne zmiany cech, jako skutek narastającej patologii.

Nie bez znaczenia jest rejestracja informacji o artefaktach zmęczeniuowo-patologicznych oraz trendach zmian cech. W przypadku spersonalizowanego rozproszonego systemu monitorowania z archiwizacją danych w chmurze obliczeniowej informacja, nawet stanowiąca dane osobowe (w tym dane o głosie pacjenta), jest wprawdzie bezpieczna, ale nieprzemijająca, o ile oczywiście użytkownik sam jej nie zneutralizuje.

Funkcjonalności systemu AVM wraz z postępowaniem badań będą rozszerzane i dostosowywane do różnych grup beneficjentów. Możliwość użycia przez dowolnego użytkownika będzie zapewne stymulatorem dalszych badań nad informacyjną zawartością głosu człowieka.

Piśmiennictwo

- Mitas AW, Bugdol MD. Biometric aspect of the voice timbre. W: *Biometry – special issues. Monograph*. Gliwice: Centrum Inżynierii Biomedycznej; 2010, 93–102.
- Mitas AW, Bugdol MD. An idea of the human voice reaction measurement system under the aspect of behavioural biometric. W: *Information Technologies in Biomedicine, vol. 2*, Piętka E, Kawa J (red.). Berlin: Springer; 2010, 219–28.
- Mitas AW, Bugdol MD, Konior W, Ryguła A. Phase angles of sound as a biometric feature. *Information Technologies in Biomedicine*. Berlin: Springer; 2012, 256–65.
- Bugdol MD, Mitas AW. Multimodal biometric system combining ECG and sound signals. *Pattern Recognit Lett*, 2014; 38: 107–22.
- Mitas AW. Dźwiękowy system antropotechniczny w aspekcie edukacji. *Konteksty Pedagogiczne*, 2015; 2(5): 33–44.
- Bugdol MD, Bugdol M, Lipowicz AM, Mitas AW, Bieńkowska M, Wijata A i wsp. Longitudinal voice study (LoVoiS) methodology and preliminary research results. *Proceedings of the 5th International Conference Information Technologies in Medicine*, Kamień Śląski, 20–22 czerwca 2016. Springer, 2016; 389–97.
- Mitas AW, Lipowicz AM. Numeryczna analiza emisji głosu w aspekcie jakości brzmieniowej. W: *Edukacyjne konteksty doświadczania świata sztuki*. Kochanowska E, Majzner R, Zawada E (red.). Kraków: StudioMoto; 2017, 155–72.
- Markova D, Richer L, Pangelinan M, Schwartz DH, Leonard G, Perron M i wsp. Age- and sex-related variations in vocal tract morphology and voice acoustics during adolescence. *Horm Behav*, 2016; 81: 84–96.
- Al-Nasheri A, Muhammad G, Alsulaiman M, Ali Z, Mesallam TA, Farahat M i wsp. An investigation of Multidimensional Voice Program parameters in three different databases for voice pathology detection and classification. *J Voice*, 2017; 31(1): 113.e9–113.e18.
- Arbeiter M, Petermann S, Hoppe U, Bohr Ch, Doellinger M, Ziethe A. Analysis of the auditory feedback and phonation in normal voices. *Ann Otol Rhinol Laryngol*, 2018; 127(2): 89–98.
- Lopes LW, Nascimento Alves J, da Silva Evangelista D, Pereira França F, Dias Vieira VJ, Bonfim de Lima-Silva MF i wsp. Accuracy of traditional and formant acoustic measurements in the evaluation of vocal quality. *Codas*, 2018; 30(5): e20170282.
- Alves Spazzapan E, Moraes Cardoso V, Gradim Fabron EM, Berti LC, Ghedini Brasolotto A, de Castro Marino VC. Acoustic characteristics of healthy voices of adults: from young to middle age. *Codas*, 2018; 30(5): e20170225.
- Lechien JR, Huet K, Finck C, Blecic S, Delvaux V, Piccaluga M i wsp. Are the acoustic measurements reliable in the assessment of voice quality? A methodological prospective study. *J Voice*, 2021; 35(2): 203–15.
- Park Y, Stepp CE. Test-retest reliability of relative fundamental frequency and conventional acoustic, aerodynamic, and perceptual measures in individuals with healthy voices. *J Speech Lang Hear Res*, 2019; 62(6): 1707–18.
- Procter T, Joshi A. Cultural competency in voice evaluation: considerations of normative standards for sociolinguistically diverse voices. *J Voice*, 2022; 36(6): 793–801.
- Saggio G, Costantini G. Worldwide healthy adult voice baseline parameters: a comprehensive review. *J Voice*, 2022; 36(5): 637–49.
- Li G, Hou Q, Zhang C, Jiang Z, Gong S. Acoustic parameters for the evaluation of voice quality in patients with voice disorders. *Ann Palliat Med*, 2021; 10(1): 130–6.