

# ZASTOSOWANIE TECHNIK MULTIMEDIALNYCH W BADANIACH I TERAPII ZMYŚLÓW KOMUNIKACJI

dr inż. Piotr Ody

Politechnika Gdańska, Wydział ETI  
Katedra Systemów Multimedialnych



## Systemy badań przesiewowych

- systemy do badania słuchu, wzroku i mowy – przez Internet, ale nie tylko
- liczbę przebadanych dzieci można szacować na ponad 500 tysięcy
- zdobyły liczne nagrody, także międzynarodowe, m.in. złoty medal i puchar Grand Prix podczas Światowego Salonu Wynalazczości „Eureka” w Brukseli



## Systemy badań przesiewowych

- system „Słyszę...” powstał pod koniec roku 1999, pozostałe na przełomie roku 2000 i 2001
- współpraca z Instytutem Fizjologii i Patologii Słuchu
- główne założenia
  - wykorzystanie Internetu do prowadzenia badań
    - HTML
  - testy przesiewowe
  - duża ilość materiałów informacyjnych
  - dodatkowe urządzenia (kalibratory) pozwalające na uzyskanie powtarzalnych i porównywalnych wyników
  - przesyłanie wyników do bazy danych

## Ogólne założenia

- małe wymagania sprzętowe
- prosty kod w HTML-u
- prosty interfejs
- wykorzystanie ciastek (cookies) do zapisu wyników badań
- dwie wersje: on-line i off-line
  - w ostatnim czasie rozwijane wyłącznie wersję off-line

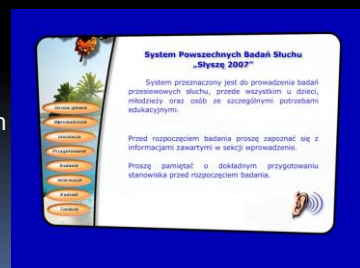
## „Słyszę...”

- miał działać na komputerach klasy 486
- problem z prędkością łączy
  - dźwięki były kompresowane do mp3, ściągane w tej postaci na dysk twardy użytkownika, specjalny program je wypakowywał, dekodował do WAV-a i dodawał szum
- problemy z kartami dźwiękowymi
- kalibrator wstępnie kalibrowany na sztucznym uchu



## „Słyszę...” – schemat badania

- ankieta osobowa z pytaniami
- test tonalny
  - 1kHz, 2kHz, 4kHz
  - dźwięki przeplatane ciszą
- testy mowy w szumie
  - szum CCITT
  - plansze z czterema wyrazami (test zamknięty)
  - test obrazkowy dla dzieci młodszych
    - wyrazy dwusylabowe
  - test słowny dla dzieci starszych i dorosłych
    - wyraz jednosylabowe
- diagnoza
  - progi zaliczeniowe dobrane na podstawie testów pilotażowych

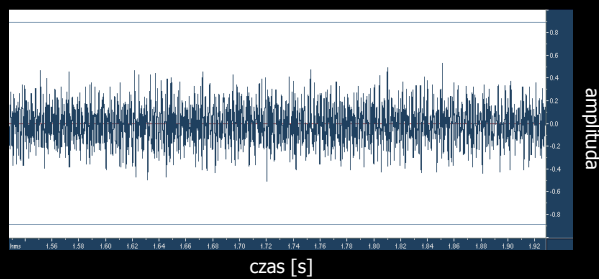


## Widmo szumu CCITT i szumu Fastla

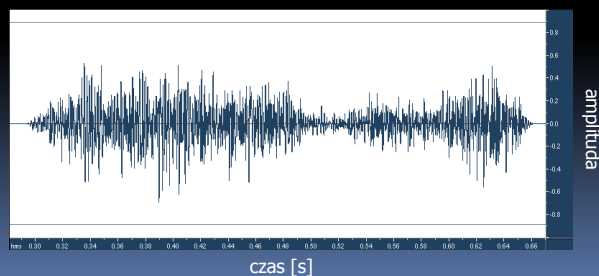


## Przebieg czasowy szumu CCITT i szumu Fastla

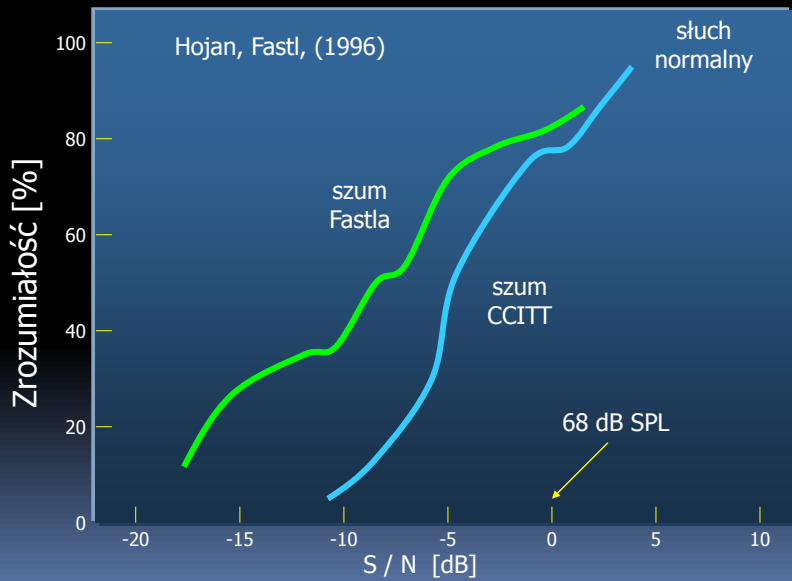
szum CCITT



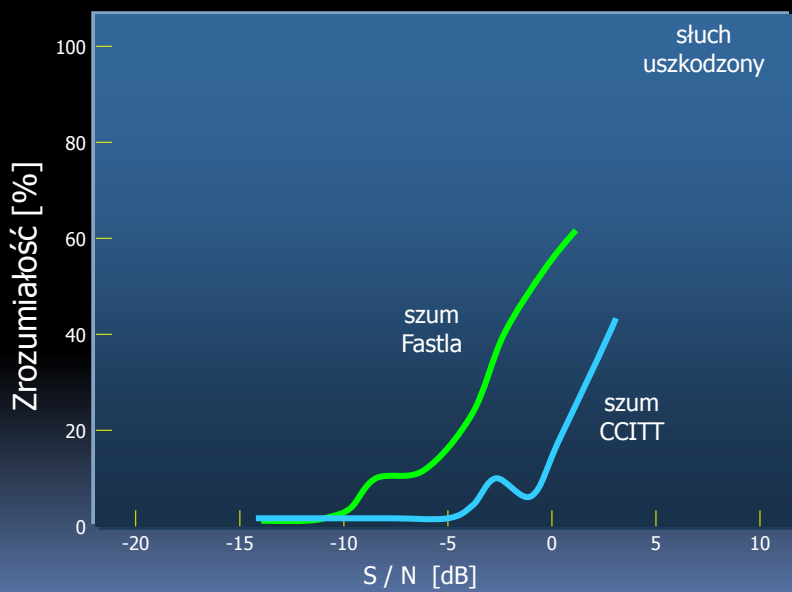
szum Fastla



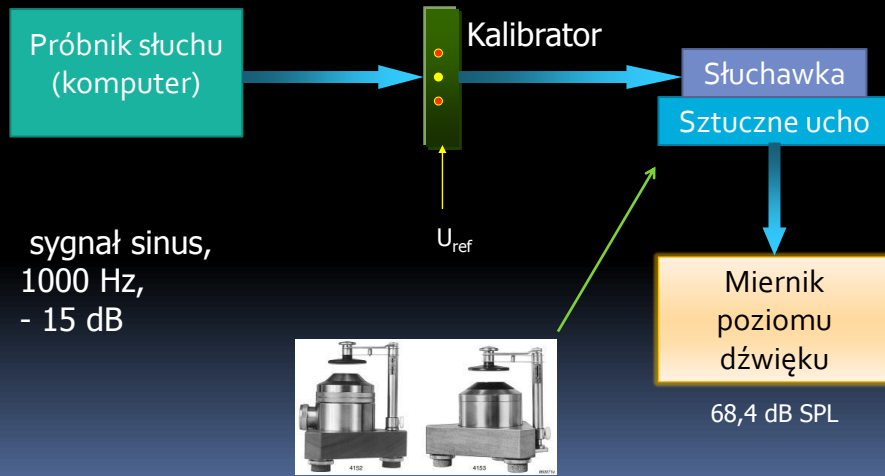
## Porównanie szumów



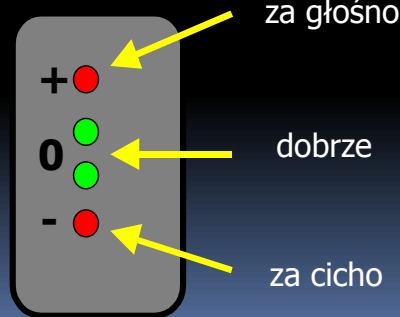
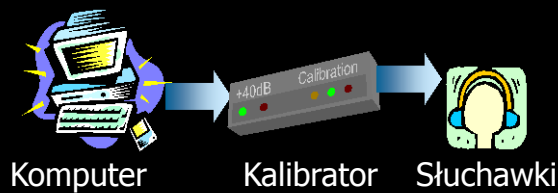
## Porównanie szumów



## Kalibracja (w procesie produkcji)



## Kalibracja (po stronie użytkownika)



## Audiometria słowna

- zaleta audiometrii słownej w szumie: względna skala audiogramu (ważny stosunek S/N)
- zmniejszenie wymogów kalibracyjnych
- łatwa procedura kalibracyjna
- zaleta audiometrii słownej: minimalizacja wpływu szumów otoczenia

## Widzę

- opracowany we współpracy z prof. Szaflikiem
- bardziej zaawansowany niż Słyszę
- konieczna kalibracja monitora
  - różnice między CRT, starymi LCD i nowymi LCD



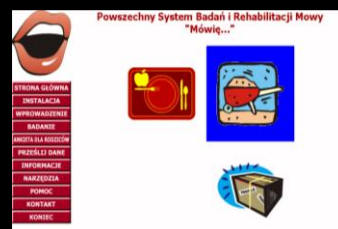
## Widzę – schemat badania

- ankieta osobowa z pytaniami
- test różnicowania kontrastu
  - pozwala wykryć poważniejsze problemy ze wzrokiem
  - polega na prezentowaniu kół wypełnionych liniami o zmiennym kontraście; dodatkowo linie te nachylone są pod różnymi kątami: 0, +45 i -45 stopni
- test widzenia barwnego
- test widzenia stereoskopowego
  - wykorzystuje anaglify



## Mówię

- największa liczba testów
  - ocena motoryki narządów mowy, słuchu fonemowego, słuchu fonetycznego, artykulacji, powtarzanie słów, powtarzanie ciągu wyrazów, słuchowa analiza głoskowa, słuchowa synteza głoskowa, ocena słownictwa, ocena gramatyki
  - polecenia czytane przez lektora
- najbardziej skomplikowany od strony graficznej
- konieczna osoba nadzorująca badania

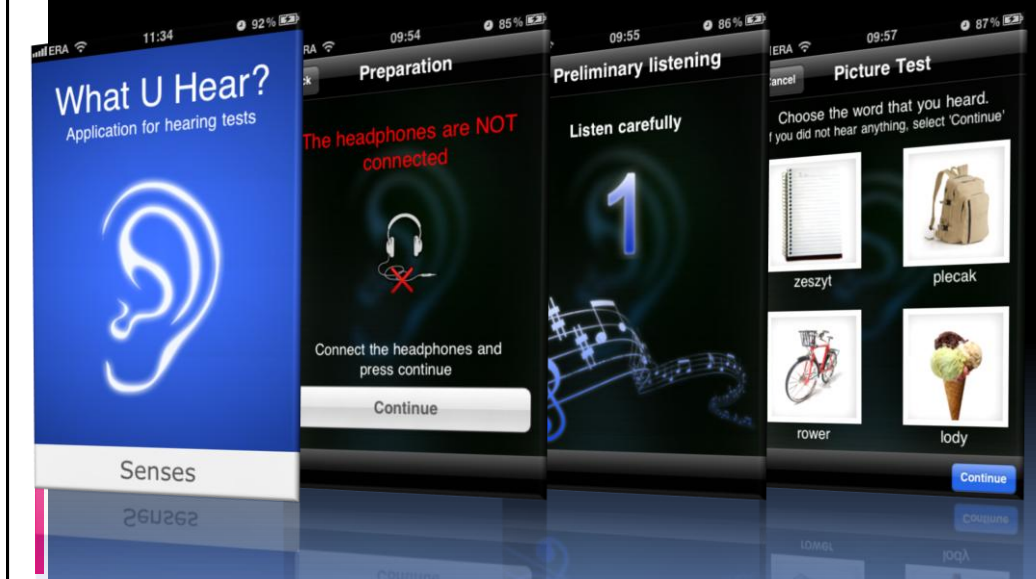




## Stan aktualny

- systemy w wersji Flash
- współpraca z YDP
- specjalne wersje dla dzieci przedszkolnych i ze specjalnymi potrzebami edukacyjnymi
  - dodanie elementów audiometrii behawioralnej: filtrowane dźwięki zwierząt zamiast testu tonalnego
- instalowane na twardym dysku użytkownika
- specjalny program do analizowania wyników
  - wyniki zapisywane w postaci plików tekstowych na dysku użytkownika
- obecnie także wersja na palmtopy
- wersje na iOS w opracowaniu

## Wersje na iOS



## Tinnitus

- miał służyć osobom cierpiącym na szumy uszne
- nie wykorzystywał testów
- pozwalał na odsłuchiwanie plików dźwiękowych
- praktycznie nie jest rozwijany



## Audiometria komputerowa

- od końca lat 90-tych powstało kilka wersji komputerowych audiometrów
- podstawowe założenia
  - dźwięki testowe nagrane na płycie CD w trybie Mixed-Mode
  - aplikacja sterujące odtwarzaniem dźwięku
  - zintegrowana baza danych
  - kalibrator

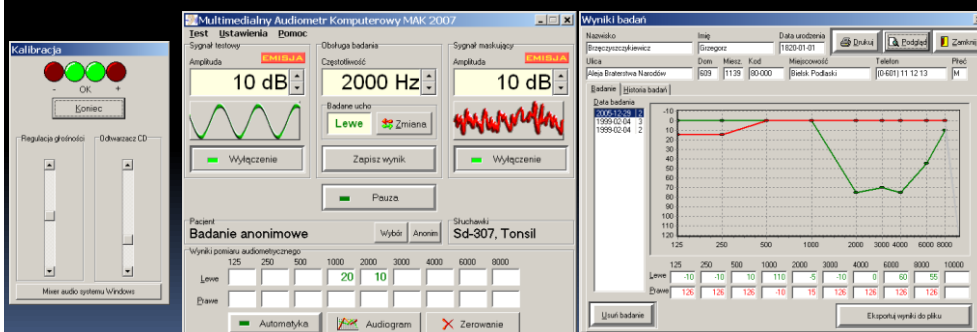


## Porównanie z tradycyjnym audiometrem

- testy wykazały, że MAK wskazuje na większe ubytki niż tradycyjny audiometr
- średnia różnica wskazań mieści się w zakresie  $\pm 5$  dB, za wyjątkiem najniższych częstotliwości
- różnice mogą być spowodowane
  - słuchawkami Tonsil
  - parametrami przetwornika C/A i wzmacniacza słuchawkowego

## MAK 2007

- zakres badania: 125 Hz do 8 kHz
- możliwość dodawania szumu



## Jąkanie

- Jąkanie jest jedną z częściej spotykanych wad mowy
  - ok. 0,5-1% populacji każdego kraju to osoby jękające się
  - jąkanie zakłóca proces komunikowania się
  - jąkanie jest zaburzeniem złożonym i długotrwałym
- Patogeneza jąkania nie jest do końca poznana
  - dziedziczne skłonności w formie niepełnowartościowego systemu wegatatywnego
  - przyuczanie dzieci leworęcznych do posługiwania się prawą ręką
  - ogólna niesprawność motoryczna
  - silne emocje
  - **defekty kontrolnych pętli sprzężenia zwrotnego**

## Kontrolne pętle sprzężenia zwrotnego

- nadzorowanie procesu artykułowania mowy
  - pętla dotykowa - prawidłowy ruch mięśni i stawów
  - pętla audytywna - kontrola jakości wytworzonych dźwięków
- dokładna analiza i kontrola wypowiedzianych słów oraz ruchów mięśni narządu mowy są niezbędne tylko na etapie nauki mówienia
  - płynne mówienie jest możliwe dzięki zautomatyzowaniu procesu tworzenia mowy
  - nadzorowanie przez mózg procesu mówienia po zakończeniu przyswajania mowy

## Audytywne sprzężenie zwrotne

- Modyfikacje wprowadzane w pętli audytywnego sprzężenia zwrotnego mogą poprawiać płynność mowy



- maskowanie sygnału mowy szumem (MAF – Masked Auditory Feedback)
  - opóźnianie sygnału mowy (DAF – Delayed Auditory Feedback)
  - transpozycja widmowa sygnału mowy (FAF – Frequency Altered Feedback)
- Rozwój technologii pozwala na wprowadzanie modyfikacji w pętli audytywnego sprzężenia zwrotnego

## Cyfrowy Korektor Mowy

- wynalazek prof. A. Czyżewskiego i dr M.R. Mieszkowskiego
- opracowany na początku lat 90.
- wykorzystuje algorytmy DAF i FAF



## Badania kliniczne skuteczności wybranych algorytmów korekcji mowy

- na przełomie 2002 i 2003 roku 100 korektorów trafia do poradni w całej Polsce
- opracowane zostają zasady prowadzenia terapii i oceny jej postępów
- wyniki nadsyłane z poradni są analizowane w celu oceny skuteczności algorytmów korekcji mowy
- próby wykonano z udziałem kilkuset osób, terapią objętych zostało 128 osób

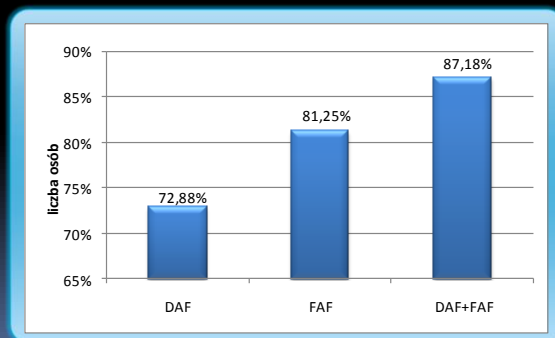


## Ocena skuteczności

- ponad 75% pacjentów dostrzega, że ich mowa w trakcie korzystania z korektora ulega poprawie
- ponad 70% osób twierdzi, że mówi bardziej płynnie także po wyłączeniu korektora
- w przypadku metody FAF pacjenci wybierali algorytmy o niewielkim przesunięciu na skali częstotliwości
- 66% osób noszących korektor czuje się pewniej
- 25% osób uznaje noszenie korektora za coś wstydliwego

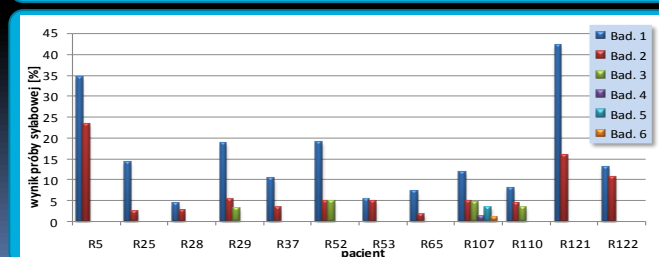
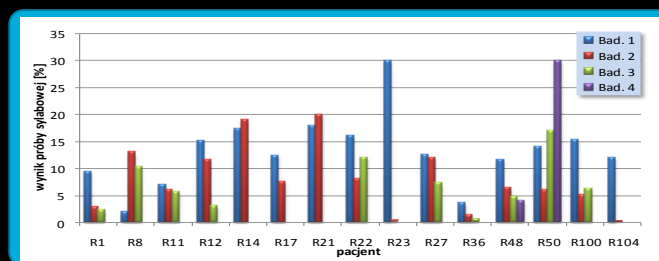
## Ocena skuteczności

- odsetek osób zauważających poprawę swojej mowy z podziałem na poszczególne algorytmy korekcji



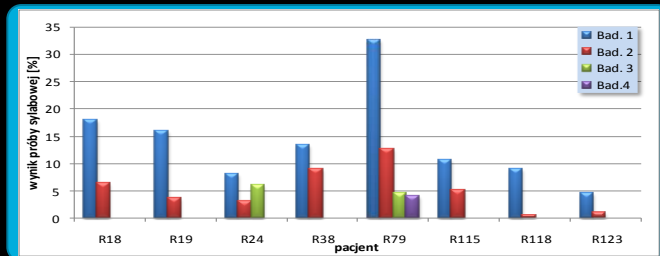
## Ocena skuteczności – porównanie metod

- wyniki prób sylabowych

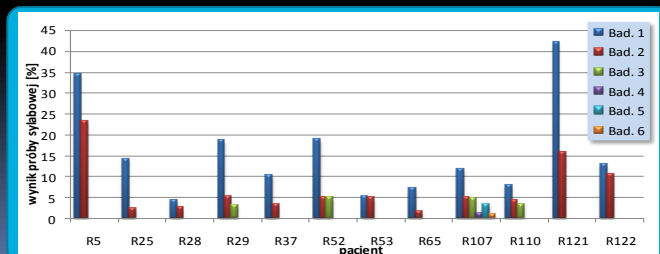


## Ocena skuteczności – porównanie metod

### wyniki prób sylabowych



metoda FAF



metoda DAF+FAF

## Korektor subminiaturowy - założenia

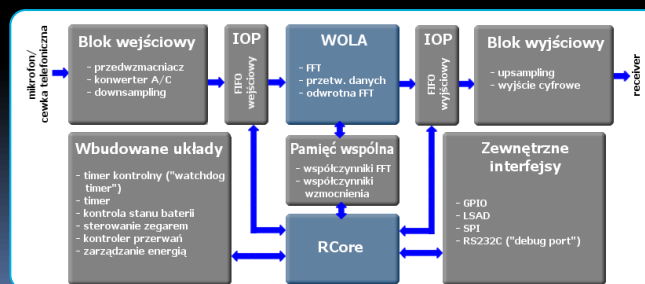
- wymiary identyczne z wymiarami wewnątrzusznych aparatów słuchowych
- nie jest konieczne korzystanie z dwóch aparatów (dwóch „słuchawek”)
- użycie algorytmów DAF, FAF, DAF+FAF
- dodatkowe możliwości obróbki sygnału, a co z tym związane nowe algorytmy
- subminiaturowy procesor sygnałowy





## Procesor subminiaturowy

- programowanie w języku assembler w celu uzyskania wysokiej wydajności
- parametry ustawiane z poziomu komputera PC
- częstotliwość zegara: 640kHz-3,84MHz
- częstotliwość próbkowania - od ok. 5kHz do 60kHz



## Subminiaturowy Korektor Mowy

- zaproponowano nowe algorytmy korekcji mowy
  - FAF-DAF
    - okresowa zmiana FAF na DAF
  - algorytm pogłosowy
    - rozwinięcie typowej metody DAF
  - modulacja opóźnienia
    - powoduje powstanie efektu chóralnego
- dodatkowe algorytmy
  - procesor dynamiki
  - korektor barwy
  - klucz głosowy
  - sygnalizacja stanu baterii

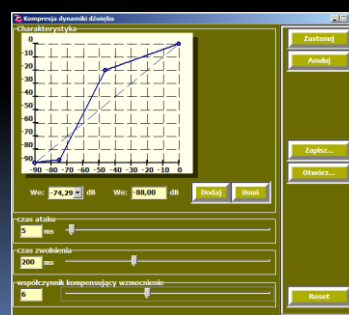
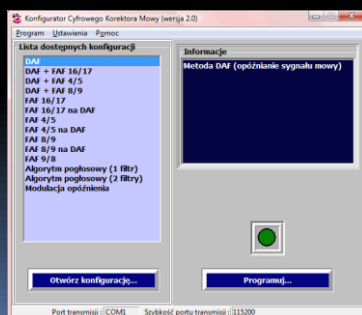
## Subminiaturowy Korektor Mowy

- parametry procesora:
  - częstotliwość zegara: 1,92MHz
  - częstotliwość próbkowania: 16kHz
  - liczba pasm: 16 (32-punktowa FFT)
- zaproponowano nowe algorytmy korekcji mowy
  - FAF-DAF
    - okresowa zmiana FAF na DAF
  - algorytm pogłosowy
    - rozwińnięcie typowej metody DAF
  - modulacja opóźnienia
    - powoduje powstanie efektu chóralnego
- dotatkowe algorytmy
  - procesor dynamiki
  - korektor barwy
  - klucz głosowy
  - sygnalizacja stanu baterii



## Subminiaturowy Korektor Mowy

- zmiana parametrów z poziomu komputera PC
  - przystawka podłączana do portu szeregowego komputera
  - dotatkowe oprogramowanie
    - umożliwia zmiany wszelkich parametrów korektora
    - wstępne ustawienia dobrane na podstawie eksperymentów

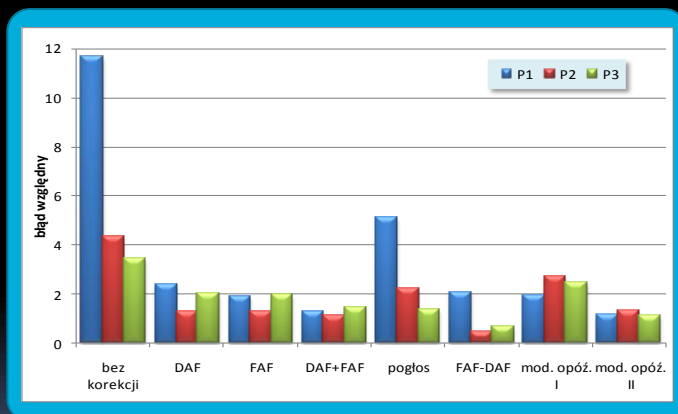


## Weryfikacja poprawności działania

- obiektywne sprawdzenie parametrów
  - pobór prądu (typowo ok. 300μA)
  - poprawność działania algorytmów
- sprawdzenie w warunkach klinicznych
  - trzy osoby, które wcześniej nie korzystały z elektronicznych metod korekcji mowy
  - dwie sesje testowe
  - teksty: ok. 200-250 sylab
  - każda nie płynność liczona jako jeden błąd
  - suma błędów odniesiona do liczby sylab w tekście wyznacznikiem stopnia nasilenia jąkania
  - mierzony także czas wypowiedzi

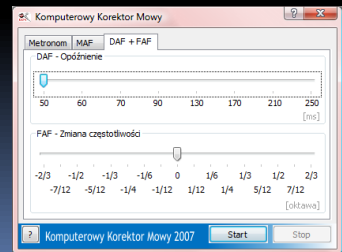
## Ocena skuteczności

- błąd względny



## Komputerowy korektor mowy

- pracuje na dowolnym komputerze klasy PC wyposażonym w kartę dźwiękową (oraz mikrofon i słuchawki)
- zaimplementowane algorytmy
  - metronom
  - maskowanie mowy szumem (MAF)
  - opóźnienie sygnału mowy (DAF)
  - przesunięcie sygnału mowy w dziedzinie częstotliwości (FAF)

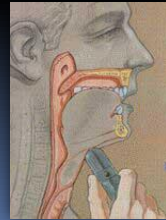


## Pomoce dla osób po laryngektomii

- mowa jest jedną z najbardziej podstawowych form komunikacji między ludźmi
- jednym z najpoważniejszych zabiegów ingerujących w funkcjonowanie aparatu mowy jest laryngektomia
- zaproponowane rozwiązania:
  - Cyfrowa Krtań Elektroniczna
  - Syntetyzer Komunikatów Głosowych

## Sztuczna krtąń

- zawiera wyłącznie elementy elektromechaniczne – generator impulsów oraz cewkę z nurnikiem uderzającym w membranę
- generowane drgania są wprowadzane do jamy ustnej i gardła poprzez przyłożenie urządzenia do skóry szyi
- na mowę nałożony jest szum (warkot) urządzenia, co negatywnie wpływa na jej zrozumiałość. Mowa brzmi przy tym bardzo sztucznie i monotonna. W efekcie zrozumiałość mowy szacuje się na ok. 60%.



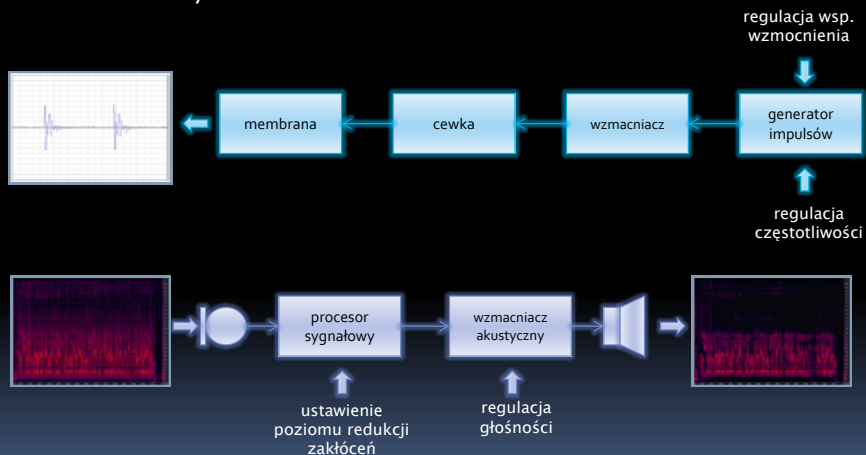
## Cyfrowa Krtąń Elektroniczna

- dzięki zastosowaniu cyfrowego przetwarzania sygnałów możliwe jest zredukowanie poziomu zakłóceń i poprawa jakości generowanej mowy



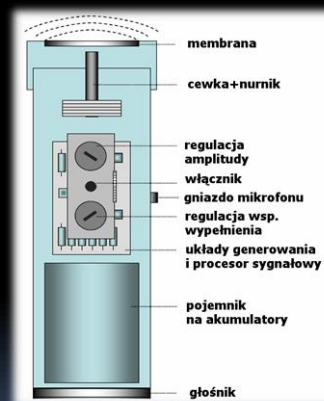
# Cyfrowa Krtań Elektroniczna

Schemat blokowy:



# Cyfrowa Krtań Elektroniczna

- Zastosowane algorytmy:
  - redukcja warkotu wibratora
    - odejmowanie widmowe
    - filtracja grzebieniowa
  - eliminacja sprzężeń zwrotnych
    - modulacja opóźnienia - wprowadza wolnozmienną modulację częstotliwości
    - transpozycja widmowa – przesuwamowę na skali częstotliwości



## Odejmovanie widmowe

- polega na systematycznym określaniu średniego widma sygnału i średniego widma szumu we fragmentach nagrania oraz dokonywaniu odejmowania obu reprezentacji widmowych
- sygnał zaszumiony może być opisany wzorem:

$$y[m] = x[m] + n[m]$$

gdzie  $x[m]$  to sygnał mowy a  $n[m]$  jest niepożądanym szumem/zakłóceniami (od mechanicznej części sztucznej krtani)

- w dziedzinie widma widmo sygnału ma postać:

$$X(j\omega) = Y(j\omega) - N(j\omega)$$

- ponieważ widmo zakłóceń  $N(j\omega)$  jest nieznane, używa się estymaty  $N_e(j\omega)$

$$|X(j\omega)|^2 = |Y(j\omega)|^2 - |N_e(j\omega)|^2$$

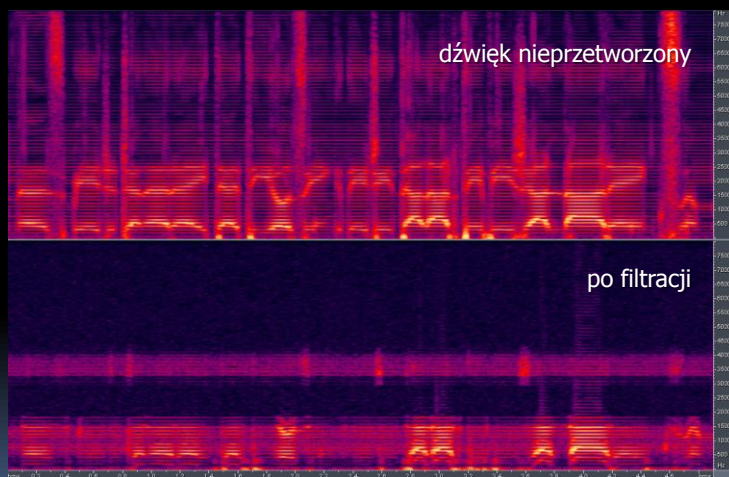
## Odejmovanie widmowe

- **problemy**
  - zarówno mowa, jak i zakłócenia mają to samo źródło i są ze sobą ściśle skorelowane dla dźwięcznych głosek
  - widmo zakłóceń musi być estymowane z sygnału rejestrowanego, gdy pacjent ma zamknięte usta
  - pacjenci muszą mieć zamknięte usta przez pierwszą sekundę po włączeniu sztucznej krtani

## Odejmovanie widmowe

- parametry procesora:
  - częstotliwość zegara: **1.28MHz**;
  - częstotliwość próbkowania: **10.7kHz**;
  - liczba podpasm: **64** (128 points FFT);
  - rozdzielczość częstotliwościowa wynosi ok. **83Hz**.

## Odejmovanie widmowe





## Filtracja grzebieniowa

- prostszy algorytm
- charakterystyka widmowa filtru umożliwia precyzyjne wstrojenie się w maksima sygnału zakłócającego
- główna zaleta: nie jest konieczne estymowanie widma zakłóceń, więc pacjent może mówić natychmiast po włączeniu urządzenia

## Filtracja grzebieniowa

- filtr grzebieniowy może być opisany wzorem:

$$y[n] = x[n] + \alpha x[n-k]$$

gdzie  $x[n]$  i  $y[n]$  oznaczają sygnał wejściowy i wyjściowy,  $\alpha$  jest współczynnikiem mnożenia,  $k$  oznacza opóźnienie w próbkach

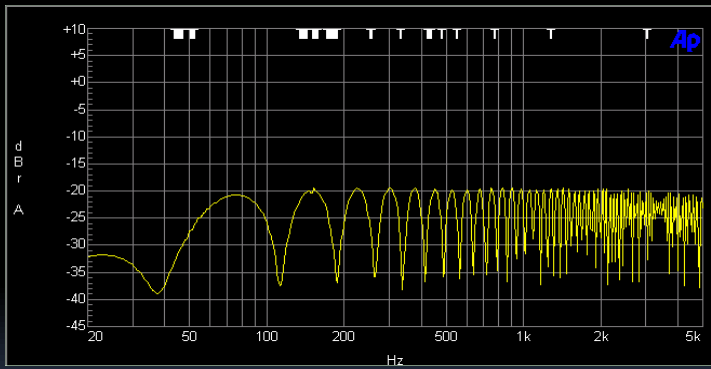
- widmo amplitudowe jest opisane wzorem:

$$|H(j\omega)| = [(1+\alpha^2)+2\alpha \cos(\omega k)]^{-1/2}$$

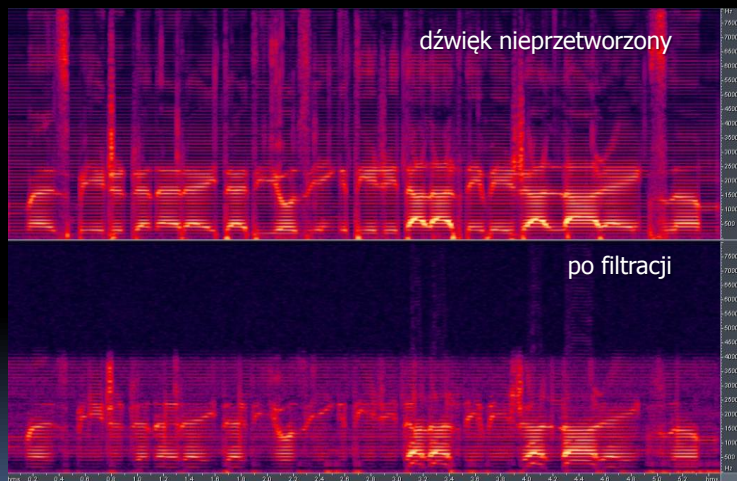
- **Implementacja**

- $\alpha = 1$
- $1 \leq k \leq 64$

# Filtracja grzebieniowa



# Filtracja grzebieniowa



## Eliminacja sprzężenia zwrotnego

- problem ze sprzężeniami zwrotnymi
  - niewielka odległość między mikrofonem i głośnikiem (ok. 15-20cm)
- testowane rozwiązania
  - filtr typu notch
  - modulacja opóźnienia
  - transpozycja widmowa

## Eliminacja sprzężenia zwrotnego

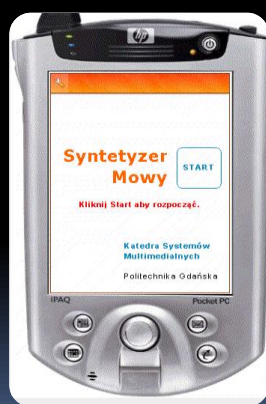
- filtr typu notch
  - brak możliwości wprowadzenia adaptacji - niewystarczające możliwości obliczeniowe procesora
- modulacja opóźnienia
  - wprowadza minimalne zmiany częstotliwości dźwięk
  - sygnał modulujący - sinus
  - okres sygnału modulującego: 1,5 s, średnie opóźnienie: 30-35 ms  
zakres zmian opóźnienia: +/-4 ms
- transpozycja widmowa
  - zbliżona do metody FAF
  - wielkość transpozycji: 6% w dół oktawy

## Wstępne testy

- pacjenci oceniali część mechaniczną porównywalnie lub nawet wyżej niż w przypadku tradycyjnych sztucznych krtani
- problemy z częścią DSP
  - pacjenci przyzwyczajani byli do mówienia natychmiast po włączeniu urządzenia, więc efekty działania odejmowania widmowego były słabe
  - pacjenci w zasadzie odrzucili modulację opóźnienia, wskazywali, że generowany dźwięk brzmi sztucznie
  - najlepsze wyniki uzyskano dla filtracji grzebieniowej z transpozycją widmową

## Syntetyzer Komunikatów Słownych

Urządzenie ma służyć osobom zmuszonym do tymczasowego zaprzestania komunikacji głosowej, z powodu rehabilitacji po laryngektomii







- elementy składowe:
  - przenośny komputer klasy PocketPC
  - oprogramowanie umożliwiające wybór, edycję i odtwarzanie komunikatów słownych
    - bardzo przyjazny interfejs użytkownika
    - duże, kolorowe ikony z opisem, odpowiadające poszczególnym kategoriom
    - możliwości wpisywania własnego tekstu
    - rozbudowane opcje sterujące wyświetlaniem tekstu na ekranie

## Syntetyzer Komunikatów Słownych

Przygotowany zestaw komunikatów głosowych dotyczy sytuacji występujących w codziennym życiu osób rehabilitowanych.

### Przykłady wygenerowanych zdań:

- zakupy
  - „Proszę chleb” 
- korzystanie ze środków transportu
  - „Poproszę bilet ulgowy do Gdańsk.” 
- wizyta u lekarza
  - „Bardzo boli mnie głowa” 
- nieskomplikowana rozmowa
  - „Cześć, co u Ciebie słychać?” 



DZIĘKUJĘ ZA UWAGĘ