

LITERATURA:

1. W. Jassem, *Podstawy fonetyki akustycznej*, PWN, Warszawa 1973.
2. F. Jaroszyk (red.), *Biofizyka – podręcznik dla studentów*, PZWL, Warszawa 2001.

LITERATURA UZUPEŁNIAJĄCA:

1. Cz. Basztura, *Źródła, sygnały i obrazy akustyczne*, WKiŁ, Warszawa 1988.
2. D. Halliday, R. Resnick, J. Walker, *Podstawy fizyki*, PWN, Warszawa 2003.

ZAGADNIENIA TEORETYCZNE:

1. Sygnały oraz ich analiza. Rodzaje sygnałów. Jednostki akustyczne. Spektrogram i widmo sygnału.
2. Wytwarzanie dźwięków mowy. Analiza oraz parametryzacja sygnału mowy. Pętla formantowa.

PROGRAM ĆWICZENIA:

1. Zarejestrowanie samogłosek /i/, /e/, /a/, /o/, /u/, a następnie wyznaczenie spektrogramu dla każdej z nich.
2. Obliczenie średnich wartości częstotliwości formantów F1, F2, F3 dla poszczególnych samogłosek.
3. Wyznaczenie pętli formantowej dla zarejestrowanych samogłosek.
5. Porównanie uzyskanych wyników z danymi literaturowymi.

WPROWADZENIE DO ĆWICZENIA:

Mówienie i słyszenie to jeden z najwydajniejszych sposobów pozwalających ludziom na komunikowanie się otoczeniem. Proces komunikacji z fizycznego punktu widzenia polega na generowaniu oraz obiorze bodźców akustycznych. Narząd mowy człowieka jest układem, który umożliwia generowanie szerokiej gamy dźwięków poprzez sterowanie strumieniem powietrza wypływającego z płuc. Daje nam to możliwość kodowania użytecznej informacji w postaci chwilowych zmian ciśnienia. Narząd słuchu umożliwiający odbiór tych informacji odpowiada za analizę częstotliwościową docierających do niego fal akustycznych i zmianę ich na ciągi impulsów neuronowych.

Z punktu widzenia analizy sygnałów, przebieg w czasie dowolnej wielkości fizycznej może być traktowany jako sygnał. Przykładem sygnału jest fala akustyczna powstająca na skutek drgań ciała w ośrodku sprężystym, ponieważ wprowadza ona chwilowe zmiany ciśnienia akustycznego. Mowa zatem jest sygnałem akustycznym, do którego można zastosować szereg metod analizy sygnałów.

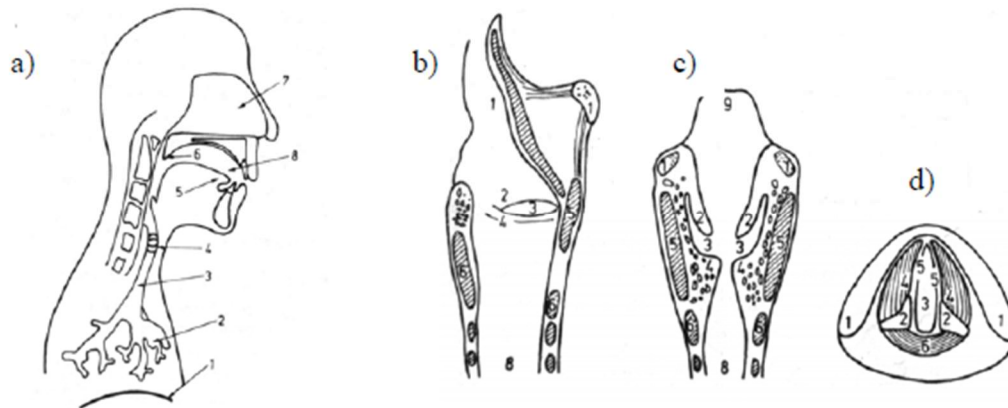
Sygnały można podzielić na zdeterminowane dające się opisać analitycznie oraz niezdeterminowane zwane losowymi, które nie dają się opisać zależnościami matematycznymi. W sygnałach okresowych przebieg powtarza się co jakiś czas T zwany okresem. Najprostszym sygnałem okresowym jest sygnał sinusoidalny (harmoniczny). Tego typu sygnał w pełni charakteryzują trzy parametry: amplituda, częstotliwość i faza początkowa. Funkcja opisująca sygnał sinusoidalnie zmienny przyjmuje postać:

$$x(t) = A \cos(\omega t + \varphi),$$

gdzie: A określa wartość amplitudy wyrażoną w jednostkach przedstawionej wielkości, φ to faza początkowa wyrażona w stopniach lub radianach, określająca przesunięcie sinusoidy względem początku układu współrzędnych, $\omega > 0$ to prędkość kątowna.

Znacznie częściej mamy do czynienia z sygnałami okresowymi złożonymi z wielu sygnałów sinusoidalnych, które nazywane są sygnałami poliharmonicznymi. Przykładem takich sygnałów są dźwięki muzyki, a także dźwięczne części mowy (samogłoski i spółgłoski dźwięczne) wytwarzane przy udziale drgań wiązań głosowych. Sygnały transjentowe, czyli takie, które można opisać za pomocą odpowiednich funkcji zmiennych, obejmują szeroką grupę sygnałów nieokresowych. Charakteryzują się one zmienną w czasie amplitudą o charakterze narastania lub zanikania. Przykładami takich sygnałów są dźwięki muzyki i mowy, które pozwalają na wydzielenie odcinków o nieustalonym przebiegu w czasie. Kolejną grupę sygnałów tworzą sygnały niezdeterminowane (losowe). Ich losowość w ogólności oznacza, że nie da się przewidzieć wartości ich parametrów w żadnej przyszłej chwili czasu. Jednym z najczęściej spotykanym i stosowanym sygnałem losowym (np. w badaniach słuchu) jest tak zwany szum biały. W sygnale tym występuje nieskończenie wiele składowych sinusoidalnych, których częstotliwości obejmują cały zakres słyszalny (do około 20 kHz). Amplitudy wszystkich składowych są identyczne, natomiast fazy początkowe są wartościami przypadkowymi. Sygnały wytworzone przez narząd mowy nie dają się jednoznacznie zakwalifikować do żadnej z wyżej wymienionych grup. Dzięki występowaniu niewielkich odcinków czasu w sygnałach mowy, które odpowiadają w przybliżeniu poszczególnym głoskom, można w wielu sytuacjach traktować je jako sygnały tego typu.

Za wytwarzanie mowy odpowiedzialny jest organ mowy, którego przekrój czynny został pokazany na Rysunku 1. Głównymi elementami tego organu są płuca, tchawica, krtani, gardło, nos, jama nosowa oraz usta. Leżąca powyżej krtani część drogi głosowej, nazywa się kanałem głosowym. Na skutek ruchów języka, warg i szczęk kształt przekroju poprzecznego kanału głosowego może się zmieniać, dzięki czemu możliwe jest wymawianie różnych głosek. Przestrzeń pomiędzy fałdami głosowymi, które otwierać się i zamykać nazywana jest głośnią. Dźwięk, który wytwarzany jest w trakcie wydostawania się powietrza z płuc przez fałdy głosowe, nazywa się tonem krtaniowym. Dźwięki takie nazywają się dźwięcznymi, są dźwiękami periodycznymi o dość małej częstotliwości podstawowej, zawierającymi harmoniczne ze znacznego zakresu częstotliwości.



Rysunek 1. Część a) przedstawia organ mowy: 1-przepona, 2-płuca, 3-tchawica, 4- więzadła głosowe, 5-język, 6-języczek, 7-jama nosowa, 8-kanał głosowy. W części b) przedstawiono przekrój profilowy krtani: 1-nagłośnia, 2-fałd kieszonki krtaniowej Morganiiego, 4-fałd głosowy, 5-chrząstka tarczowa, 6-chrząstka pierścieniowa, 7-kość gnykowa, 8-tchawica. W części c) przedstawiono przekrój czołowy krtani (oznaczenia jak w części b). Część d) przedstawia przekrój poziomy krtani: 1-chrząstka tarczowa, 2-chrząstki nalewkowe, 3-głośnia, 4-mięśnie głosowe, 5-więzadła głosowe, 6-mięśnie międzynałkowe.

Kanał głosowy możemy traktować jako układ filtrów o określonych częstotliwościach rezonansowych. Widmo tonu krtaniowego charakteryzuje się pewnymi maksimami i minimami lokalnymi. Te lokalne maksima nazywa się formantami i oznacza zwykle za pomocą litery F z liczbą określającą kolejny formant. Formant o najmniejszej częstotliwości nazywa się pierwszym F1, następny drugim F2, itd. Bezwzględne wartości częstotliwości poszczególnych formantów wykazują dość duże zróżnicowanie, zależne od płci oraz osobniczych cech budowy. Należy jednak pamiętać, że ich względne wartości nie zależą od indywidualnych cech głosowych. Poziomy poszczególnych formantów wykazują stałe w przybliżeniu wartości.

WYKONANIE ĆWICZENIA:

1. Uruchomić program PRAAT. Po uruchomieniu programu pojawiają się dwa okna, z których jedno „PRAAT objects” zawiera wszystkie opcje analizy, drugie okno „PRAAT picture” umożliwia tworzenie obiektów graficznych w celu dokumentowania wyników analiz.
Używając dostępnego mikrofonu zarejestrować kilkakrotnie (minimum 3 razy) samogłoski /i/, /e/, /a/, /o/, /u/. W tym celu w oknie *Praat Objects* klikamy *New->Record mono Sound*. (UWAGA – podczas rejestracji dźwięku w sali powinna panować cisza).
2. Po zarejestrowaniu wszystkich samogłosek należy zmierzyć średnie wartości częstotliwości formantów F1-F3 dla każdej z nich. W tym celu, z listy *Objects* wybieramy poszczególne nagrania i klikamy *View & Edit* w celu dokonania analizy. Otworzyć menu *Formant->Formant settings*. W przypadku analizy formantów dla głosu żeńskiego ustawić wartość pola „Maximum formant (Hz)” na 5500, a w przypadku głosu męskiego na 5000. Następnie obliczamy średnie wartości częstotliwości formantów. W tym celu należy zaznaczyć obszar widma częstotliwościowego i z menu *Formant* wybrać częstotliwości formantów F1-F3.
3. Sporządzić tabelę ze średnimi wartościami częstotliwości formantowych wszystkich zarejestrowanych samogłosek.
4. Wyznaczyć pętlę formantową dla zarejestrowanych samogłosek (tzw. czworobok samogłosek), określającą uśrednione położenie samogłosek na płaszczyźnie pierwszego i drugiego formantu (F1 i F2). Uzyskane wyniki porównać z danymi literaturowymi.

INFORMACJE DODATKOWE:

1. Częstotliwości formantów samogłosek podstawowych podane są w poniższej tabeli¹:

Samogłoska	F1 [Hz]	F2 [Hz]	F3 [Hz]
a	870	1750	2700
e	590	2280	2700
i	210	2750	3500
o	400	730	2300
u	270	615	2200
y	220	2550	3100

2. Zakresy formantów samogłosek podstawowych (dla 10 osób) podane są w poniższej tabeli¹:

Samogłoska	F1 [Hz]	F2 [Hz]	F3 [Hz]
a	683 – 1020	1130 – 1570	2390 – 2860
e	524 – 630	1580 – 2230	2470 – 3150
i	188 – 275	2090 – 2840	2670 – 3430
o	493 – 679	788 – 1100	2410 – 3030
u	243 – 330	557 – 789	2270 – 3190
y	262 – 391	1670 – 2360	

¹ Dane zgodnie z [1].