The left side of the slide features a decorative design with several vertical bars of varying heights and widths, some with a fine grid pattern. Below these bars are several overlapping circles of different sizes, all in a dark green color. The largest circle is positioned near the top left, with smaller circles scattered below and to its right.

# TEORIA WYTWARZANIA DŹWIĘKÓW MOWY, FORMANTY, MODELOWANIE WYTWARZANIA DŹWIĘKÓW MOWY.

mgr inż. Kuba Łopatka

1

# PLAN WYKŁADU

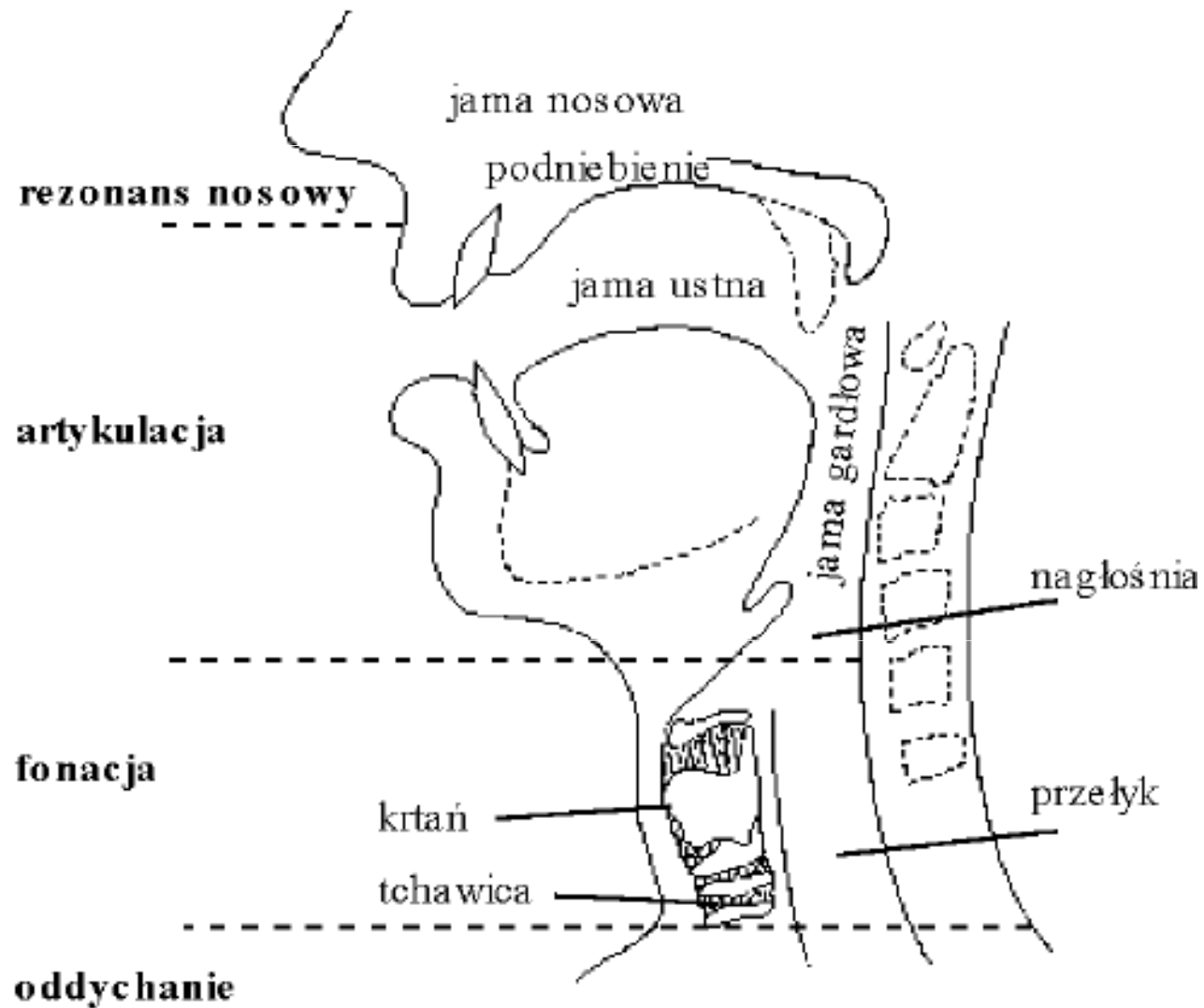
1. **Teoria wytwarzania dźwięków mowy**
  - Ogólna teoria wytwarzania dźwięków mowy
  - Ton krtaniowy
  - Pobudzenie szumowe
  - Transmitancja traktu głosowego
2. **Formanty**
  - Pojęcie formantu
  - Parametry formantowe
  - Formanty F1 i F2, trójkąt samogłosek
3. **Modelowanie wytwarzania dźwięków mowy**
  - Modelowanie fizyczne
  - Modelowanie elektryczne i komputerowe
  - Model źródło-filtr



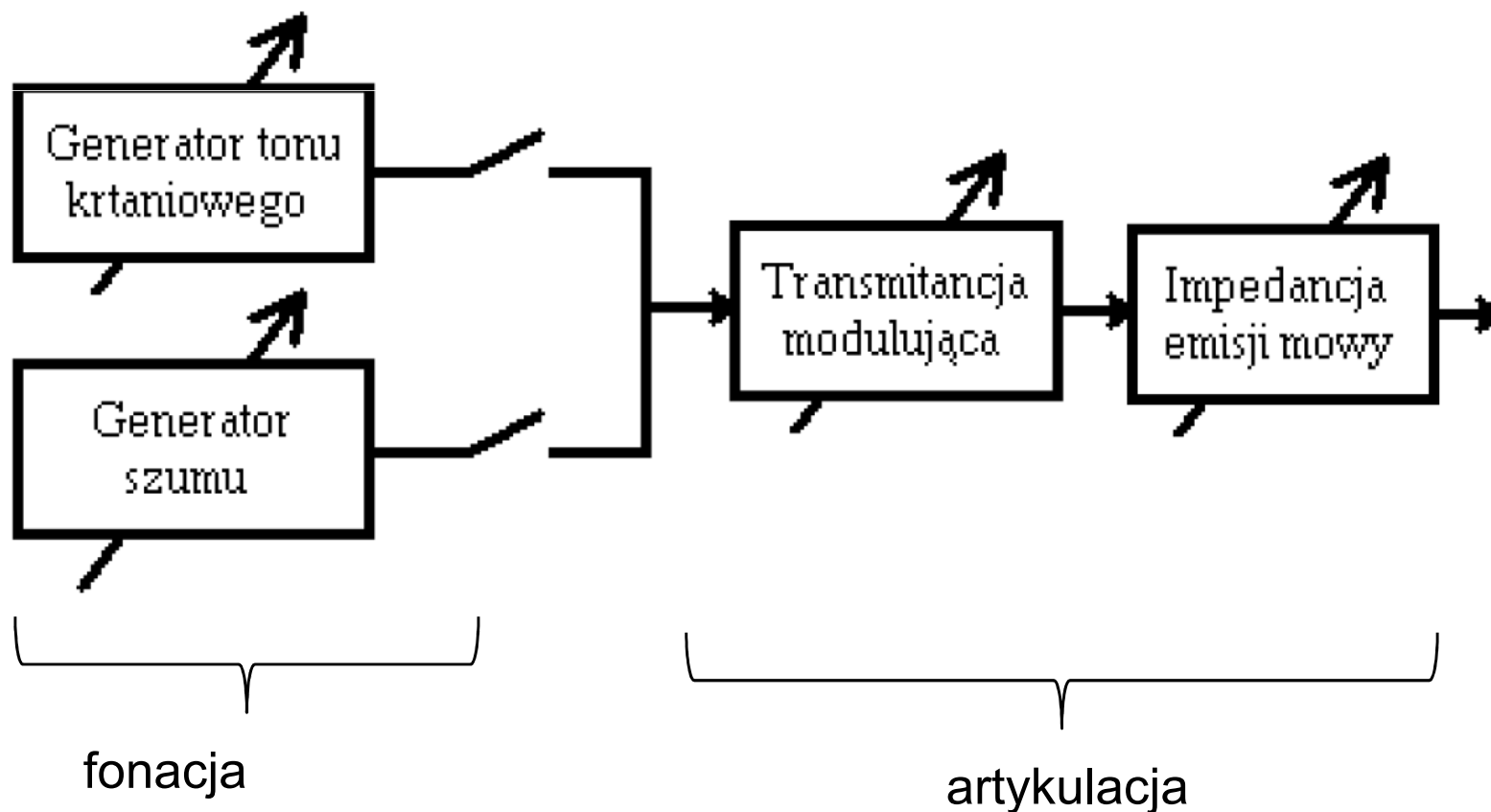
# TEORIA WYTWARZANIA DŹWIĘKÓW MOWY

3

# PRZYPOMNIENIE – ANATOMIA



# OGÓLNY SCHEMAT ZASTĘPCZY WYTWARZANIA MOWY



Podstawowym założeniem tej teorii jest **niezależność** charakterystyk tonu krtaniowego i artykulatorów.

# TON KRTANIOWY

- Jest to sygnał akustyczny wytworzony przez drgające struny głosowe.
- Stanowi pobudzenie dla głosek dźwięcznych i zwartych.
- Powstaje w procesie **fonacji**.
- Trudno go zbadać, bo zawsze jest w postaci zmodulowanej.
- Sam w sobie jest słaby i bezbarwny, ulega wzmocnieniu i nabiera brzmienia dopiero w wyższych partiach traktu głosowego.
- Podstawowy parametr – częstotliwość podstawowa –  $f_0$  (lub  $F_0$ )- zgodna z częstotliwością podstawową mowy.

# TON KRTANIOWY

Zakres częstotliwości podstawowej tonu krtaniowego:  
w mowie (typowe):

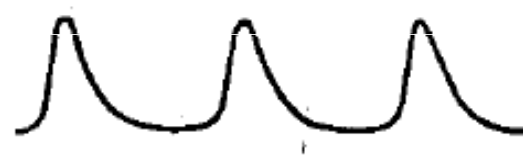
- mężczyźni: 85-180Hz
- kobiety: 165-255Hz
- dzieci: 250-300Hz

W śpiewie w zależności od rodzaju głosu:

- bas, baryton: 65Hz-349Hz
- tenor: 130-523Hz
- alt: 195-698Hz
- mezzosopran, sopran: 220-1280Hz

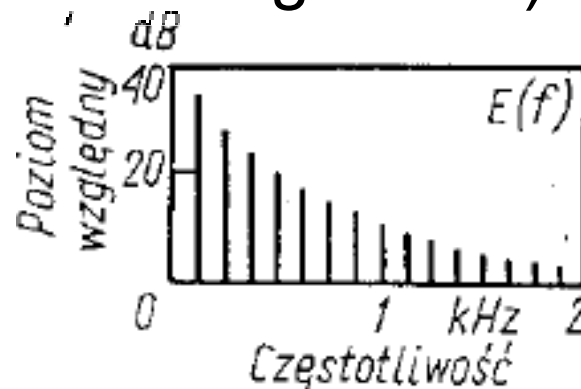
# TON KRTANIOWY

Postać czasowa:



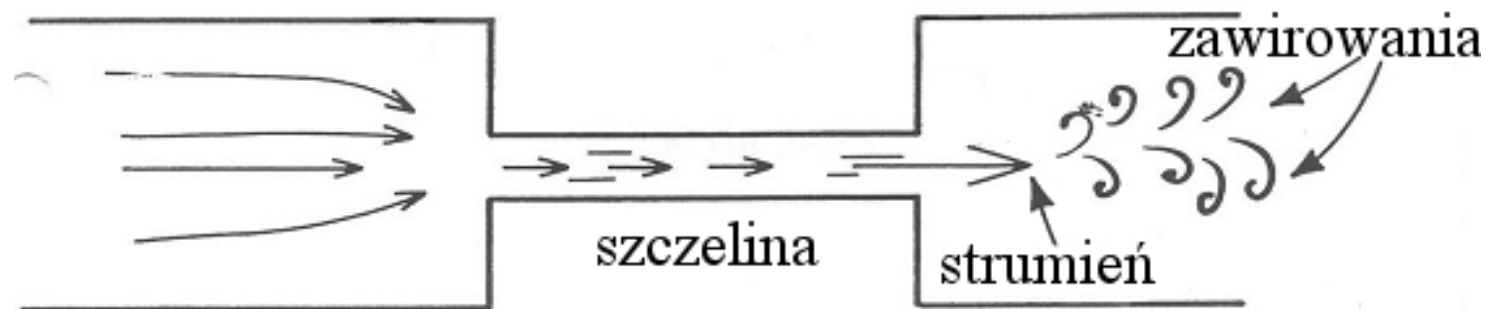
Widmo:

- Obecne są wszystkie harmoniczne częstotliwości podstawowej opadające ok. 6-12dB/okt. (w praktyce nie jest tak regularnie)



# POBUDZENIE SZUMOWE

Przy artykulacji głosek szumowych strumień wydychanego powietrza przechodzi z przepływu **laminarnego** do przepływu **turbulentnego** (zjawisko znane w hydrodynamice).

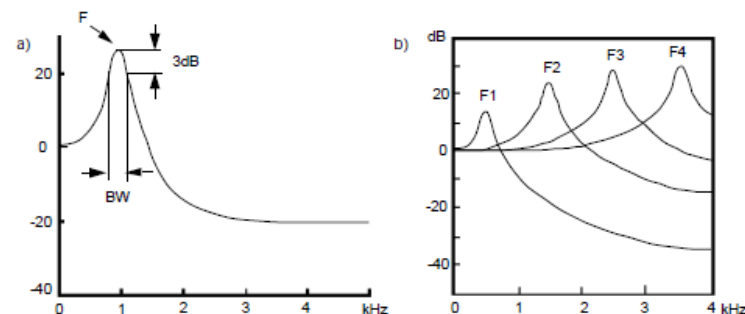


Przy przejściu powietrza przez szczelinę dokonuje się konwersja energii kinetycznej na akustyczną.

Przy nagłym otworzeniu drogi przepływu powietrza podczas artykulacji spółgłosek zwartych powietrze tworzy **falę uderową**, która jest nośnikiem energii akustycznej.

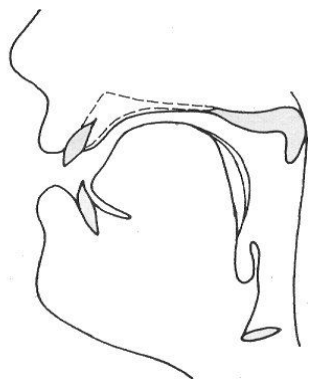
# TRANSMITANCJA TRAKTU GŁOSOWEGO

- Z punktu widzenia akustycznego trakt głosowy to układ rezonatorów.
- Kształt i połączenia między tymi rezonatorami zmieniają się w zależności od ułożenia narządów artykulacyjnych – żuchwy, języka, warg, podniebienia.
- Charakterystyka częstotliwościowa traktu głosowego jest połączeniem krzywych rezonansowych.

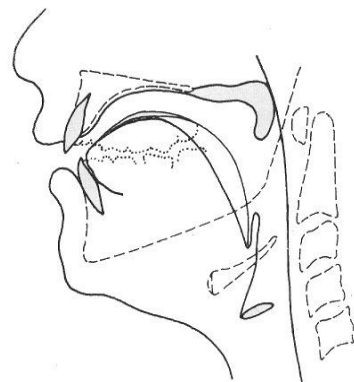


# TRANSMITANCJA TRAKTU GŁOSOWEGO

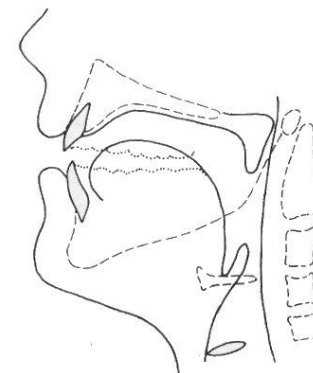
Różne konfiguracje narządów mowy:



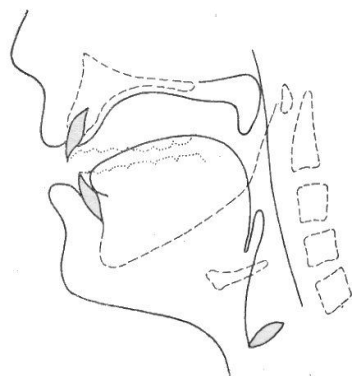
i



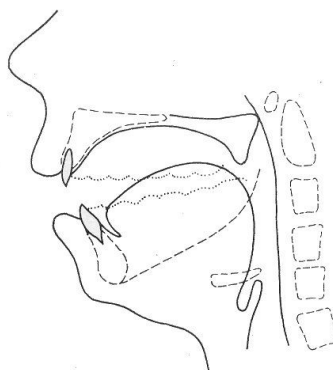
y



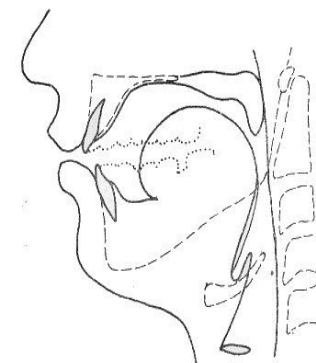
e



a



o



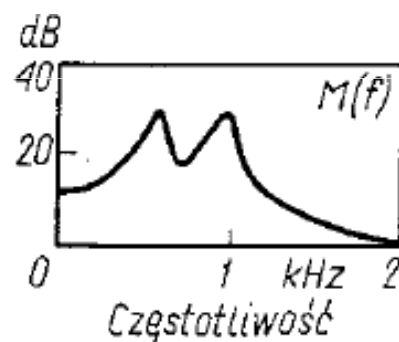
u

# TRANSMITANCJA TRAKTU GŁOSOWEGO

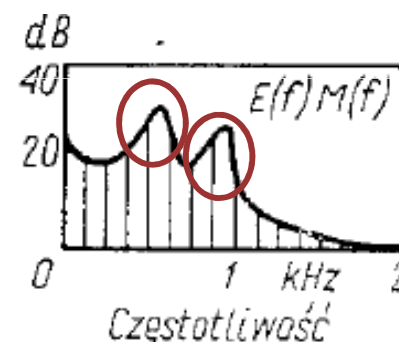
- Rezonans akustyczny narządów mowy powoduje podbicie pewnych częstotliwości (pewnych obszarów widma).
- To, które obszary widma zostaną podbite, zależy od ułożenia narządów mowy – przez zmianę ułożenia narządów mowy zmieniają się wymiary rezonatorów.
- W efekcie w widmie sygnału mowy obserwuje się charakterystyczne skupiska energii.



Widmo tonu krtaniowego



charakterystyka traktu głosowego



widmo sygnału mowy

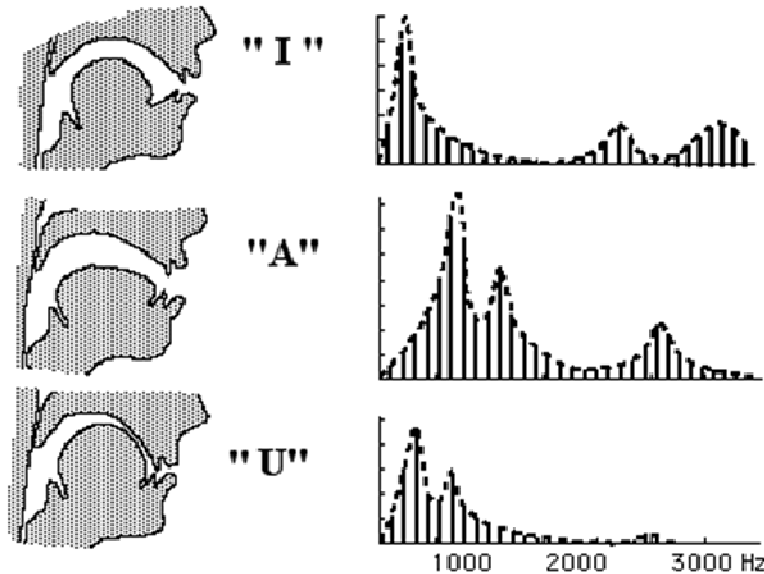


# FORMANTY

13

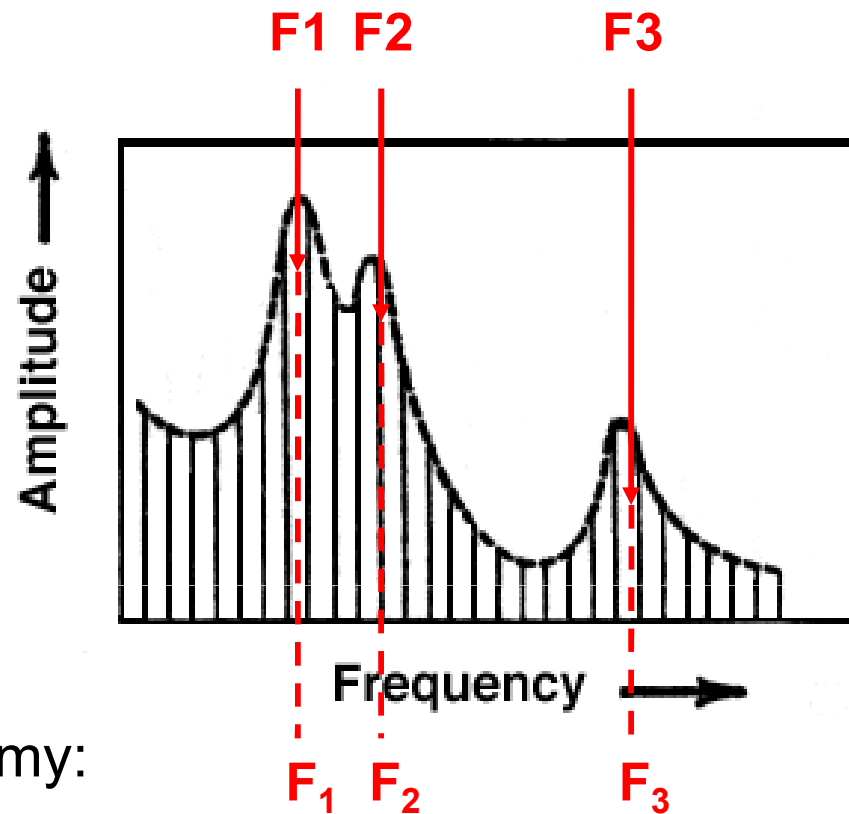
# FORMANTY

- Skupisko energii w widmie sygnału nazywamy **formantem**.
- Formant to jedno z najważniejszych pojęć w akustyce mowy.
- Rozmieszczenie formantów zależy od tego, jaka głoska jest artykułowana.



# FORMANTY

Formanty oznaczamy:



Częstotliwości  
formantowe oznaczamy:

Czasem formantem  $F_0$  nazywa się ton krtaniowy (o częstotliwości podstawowej  $F_0$ )

# FORMANTY

- Ucho ludzkie jest czułe na formanty.
- Teoria głosi, że formanty leżą u podstaw rozpoznawania mowy przez człowieka. W związku z tym wykorzystuje się je przy automatycznym rozpoznawaniu mowy.

# FORMANTY

Parametry formantowe to:

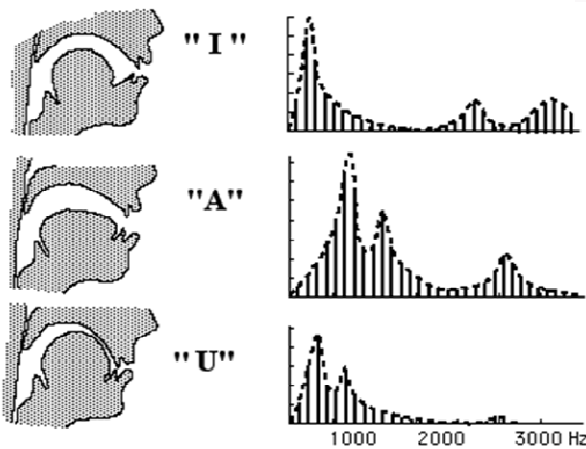
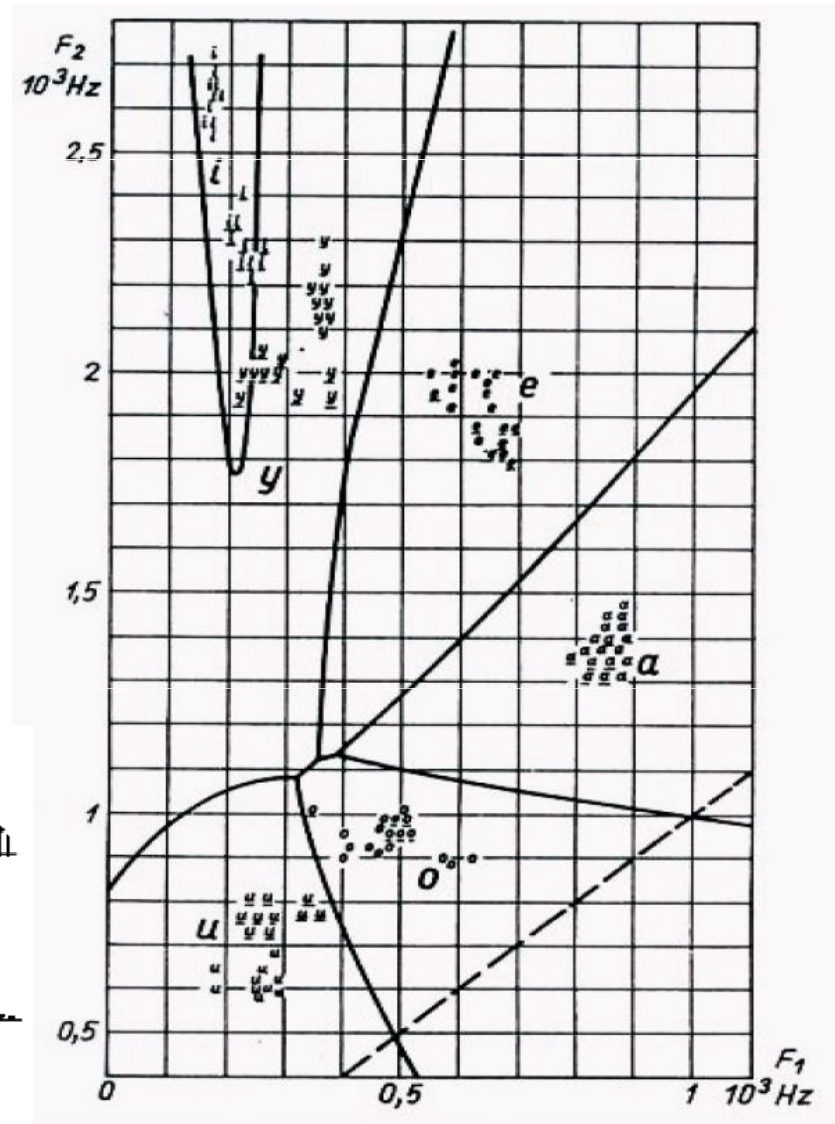
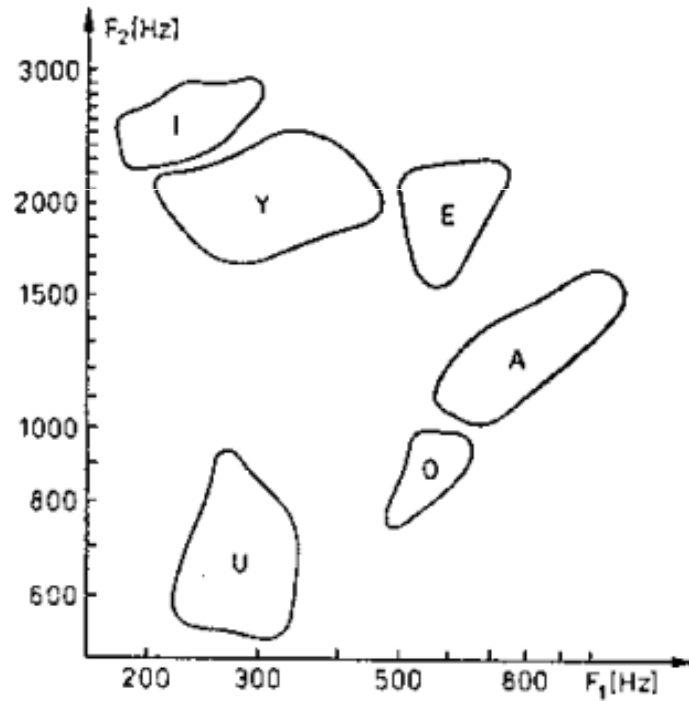
- Częstotliwość formantu  $F_N$
- Poziom formantu  $A_N$  (wyrażany w dB, względny unormowany do najsilniejszego formantu albo bezwzględny)

Fonem	częstotliwości [Hz]				poziomy względne [dB]			
i	210	2750	3500	4200	0	-15	-15	-27
e	380	2640	3000	3600	0	-12	-16	-20
a	780	1150	2700	3500	0	-7	-25	-25
y	240	1550	2400	3300	0	-12	-20	-30
o	400	730	2300	3200	0	-3	-30	-35
u	270	615	2200	3150	0	-13	-40	-50
w	600	1700	2900	4100	-9	0	-2	-10
sz	-	2300	2900	3600	-	-9	-8	0
h	500	1700	2500	4200	-12	0	-10	-17
z	-	1750	2950	4300	-	-6	-10	0

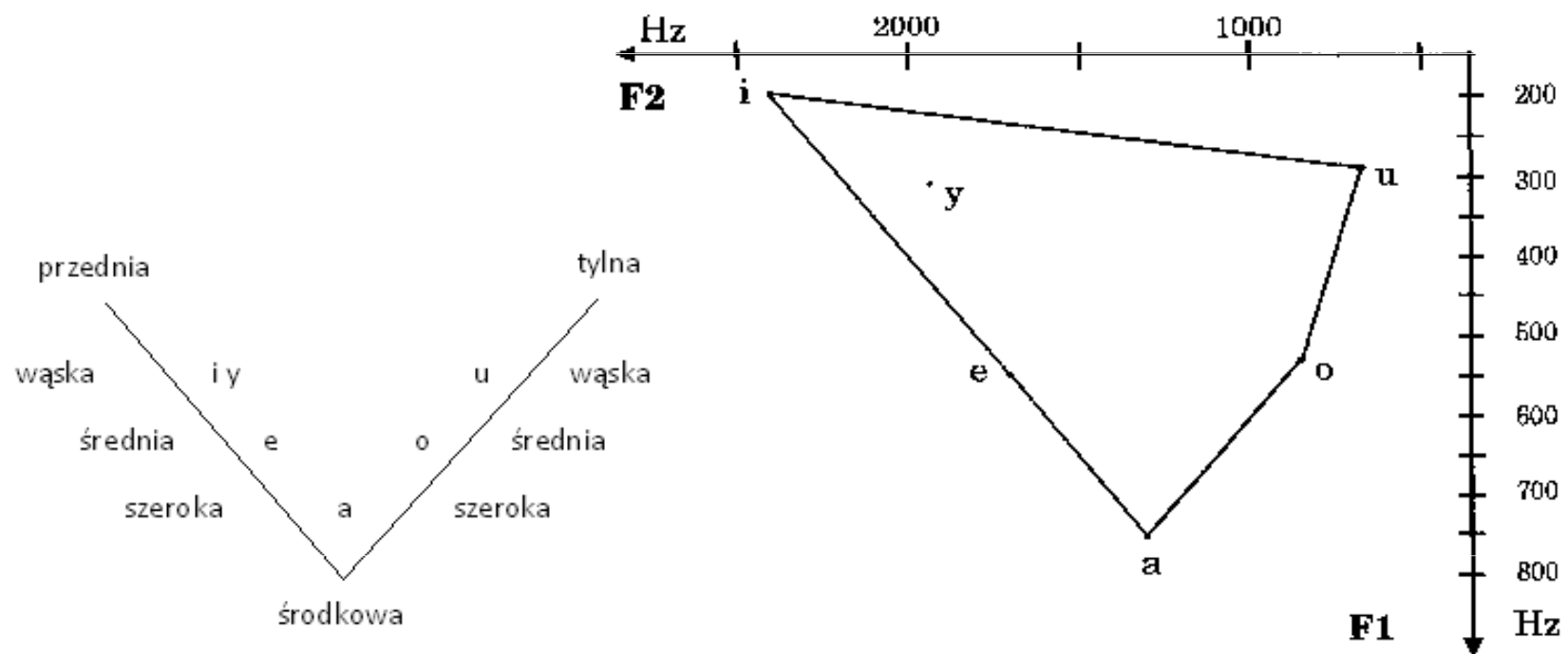
## DWA PIERWSZE FORMANTY

- Dwa pierwsze formanty (F1 i F2) zależą od ułożenia języka (dzielącego jamę ustną na dwie wnęki).
- Niektóre spośród głosek (zwłaszcza samogłoski) można odróżnić na podstawie tylko tych dwóch formantów.
- W akustyce mowy rozpowszechniony jest wykres na płaszczyźnie częstotliwości  $F_1$  i  $F_2$  – tzw. „trójkąt samogłosek”.

# TRÓJKAT SAMOGŁOSEK



# TRÓJKĄT SAMOGŁÓSEK





# MODELOWANIE MECHANIZMÓW WYTWARZANIA MOWY

21

# MODELOWANIE WYTWARZANIA MOWY

Modelowanie wytwarzania mowy wymaga 3 składników:

- Modelowanie pobudzenia
- Modelowanie traktu głosowego
- Modelowanie emisji mowy

W praktyce najważniejsze są dwa pierwsze elementy.

## MODELOWANIE POBUDZENIA

- o Głoski dźwięczne – modelowanie tonu krtaniowego

Ton krtaniowy modeluje się najczęściej przebiegiem piłokształtnym, który ma zbliżone właściwości widmowe.



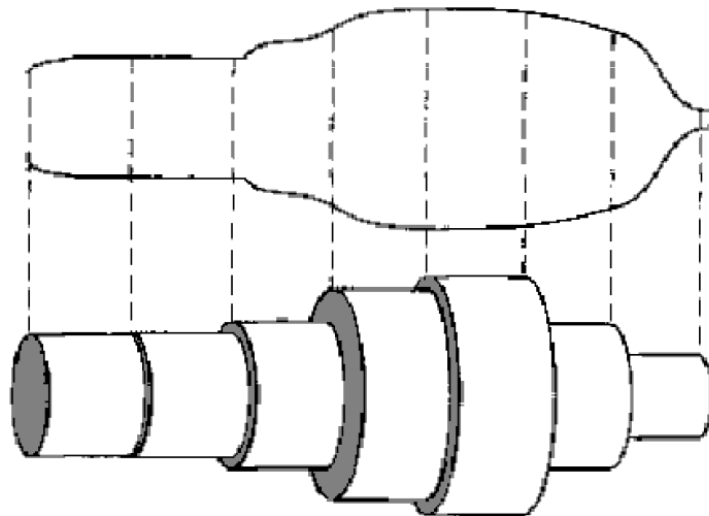
- o Głoski bezdźwięczne – pobudzenie szumowe.

Modeluje się szumem o widmie opadającym 6dB/okt.

# MODELOWANIE TRAKTU GŁOSOWEGO

- Model fizyczny

Fizycznie trakt głosowy można zamodelować jako połączenie ściętych stożków lub walców. W pierwszym przypadku powstaje model **tubowy**, zachowujący ciągłość przekroju, w drugim - model **cylindryczny**.



model tubowy

model cylindryczny

# MODELOWANIE TRAKTU GŁOSOWEGO

- rezonator Helmholtza (umożliwia modelowanie pojedynczego formantu)
- podwójny rezonator Helmholtza (umożliwia modelowanie dwóch formantów)
- modele złożone z kilku rur zakończonych płaską tarczą kołową (odgradą) imitującą
- charakterystykę promieniowania ust jako nadajnika dźwięku
- trójparametrowy model Fanta, uwzględniający rozkład biegunów i zer na płaszczyźnie zespolonej i podstawowe trzy parametry: miejsce artykulacji (miejsce największego przewężenia kanału), stopień tego przewężenia (powierzchnia przekroju) oraz kształt otworu wylotowego ust
- model Markela-Graya

# MODELOWANIE TRAKTU GŁOSOWEGO

- Model elektryczny

Odwzorowanie charakterystyki częstotliwościowej za pomocą układu filtrów elektrycznych o charakterystyce rezonansowej (LC).

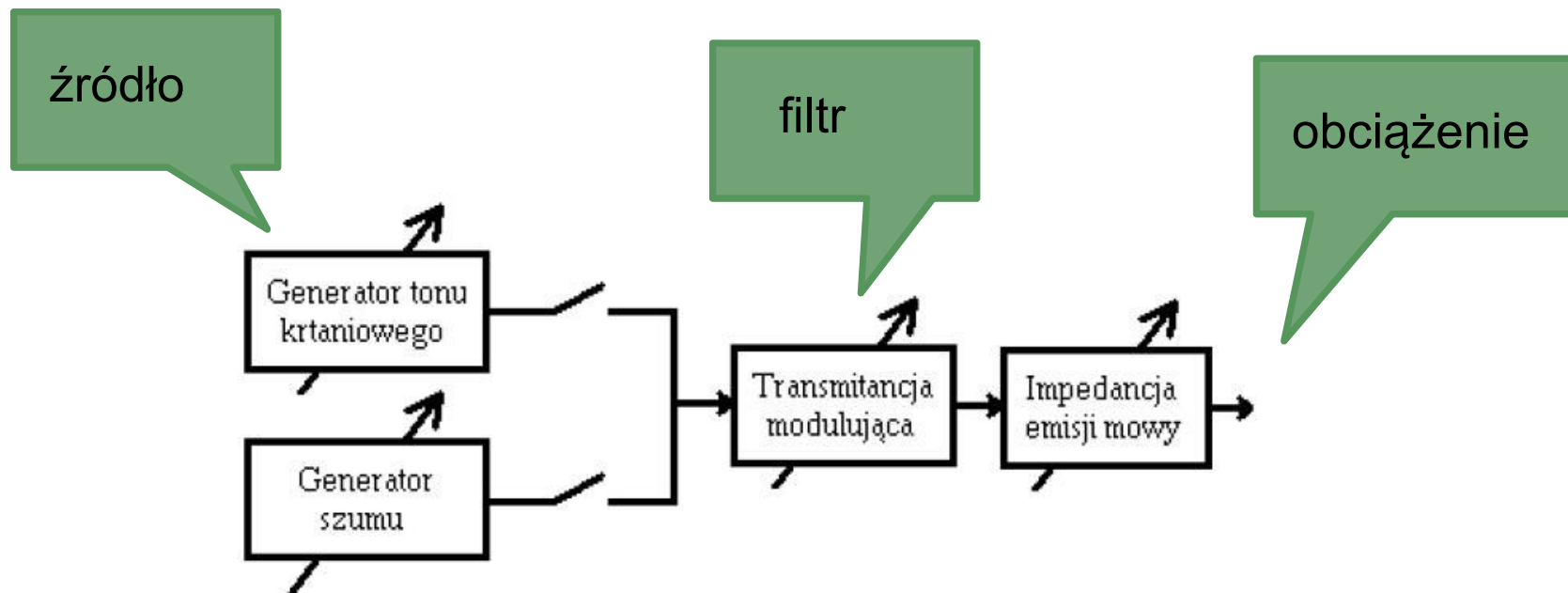
- Modele komputerowe:

- Model LPC – odwzorowanie charakterystyki traktu głosowego za pomocą filtra biegunowego.
- Model artykulacyjny – matematyczne odwzorowanie mechanizmów generowania mowy.

Więcej na wykładzie o **syntezie mowy**.

# MODEL ŹRÓDŁO-FILTR

W większości modeli wytwarzania dźwięków mowy powtarza się idea, by schemat zastępczy generowania mowy przedstawiać w układzie **źródło-filtr**.



# MODEL ŹRÓDŁO-FILTR

Jak w teorii obwodów, odpowiedź układu powstaje z wymnożenia charakterystyk pobudzenia i filtra.

