

Akustyka mowy

Podstawowe informacje o wytwarzaniu mowy i sygnale mowy

dr hab. inż. Józef Kotus

Historia mowy



- Postawa wyprostowana
- Obniżona krtań
- Odpowiednio rozwinięty mózg
- Zdolność abstrakcyjnego myślenia
- Początki mowy: 530 tyś lat temu

<https://naukawpolsce.pl/aktualnosci/news%2C142606%2Cpoczatki-mowy-mogly-byc-duzo-wczesniejsze-niz-sadzono.html>

<https://doi.org/10.1016/j.ygeno.2020.05.006>

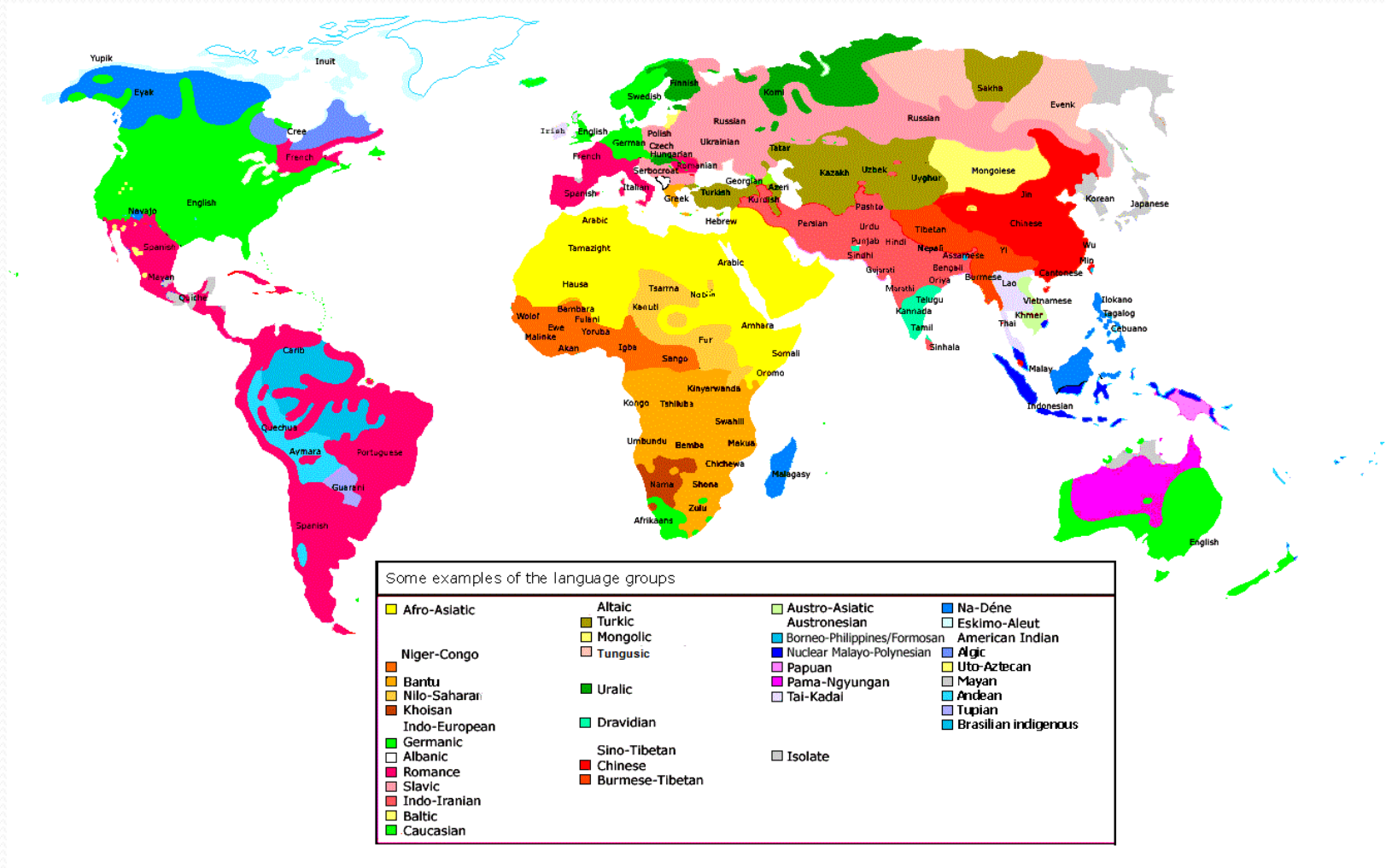
<https://histmag.org/Historia-komunikacji-od-mowy-do-Internetu-744>

<https://pl.wikipedia.org/wiki/Antropogeneza>

Znaczenie mowy

- Mowa – zbiór wokalnych symboli opartych na logicznej strukturze, umożliwiła przekazywanie informacji i wiedzy kolejnym pokoleniom w doskonalszy sposób.
- Doświadczenia przekazywane za pomocą języka stawały się coraz bogatsze, a człowiek mógł zmieniać się i przystosowywać się do otoczenia – bądź przystosowywać otoczenie do siebie – znacznie szybciej niż gatunki opierające się na genetycznej ewolucji.
- Mowa oznaczała możliwość koordynacji i współpracy, postępu technicznego, nauki, inżynierii, powstania pojęć abstrakcyjnych i kultury.
- Ulokowała człowieka na szczycie łańcucha pokarmowego, umożliwiając homo sapiens walkę i zwycięstwo nad zwierzęcymi drapieżnikami.

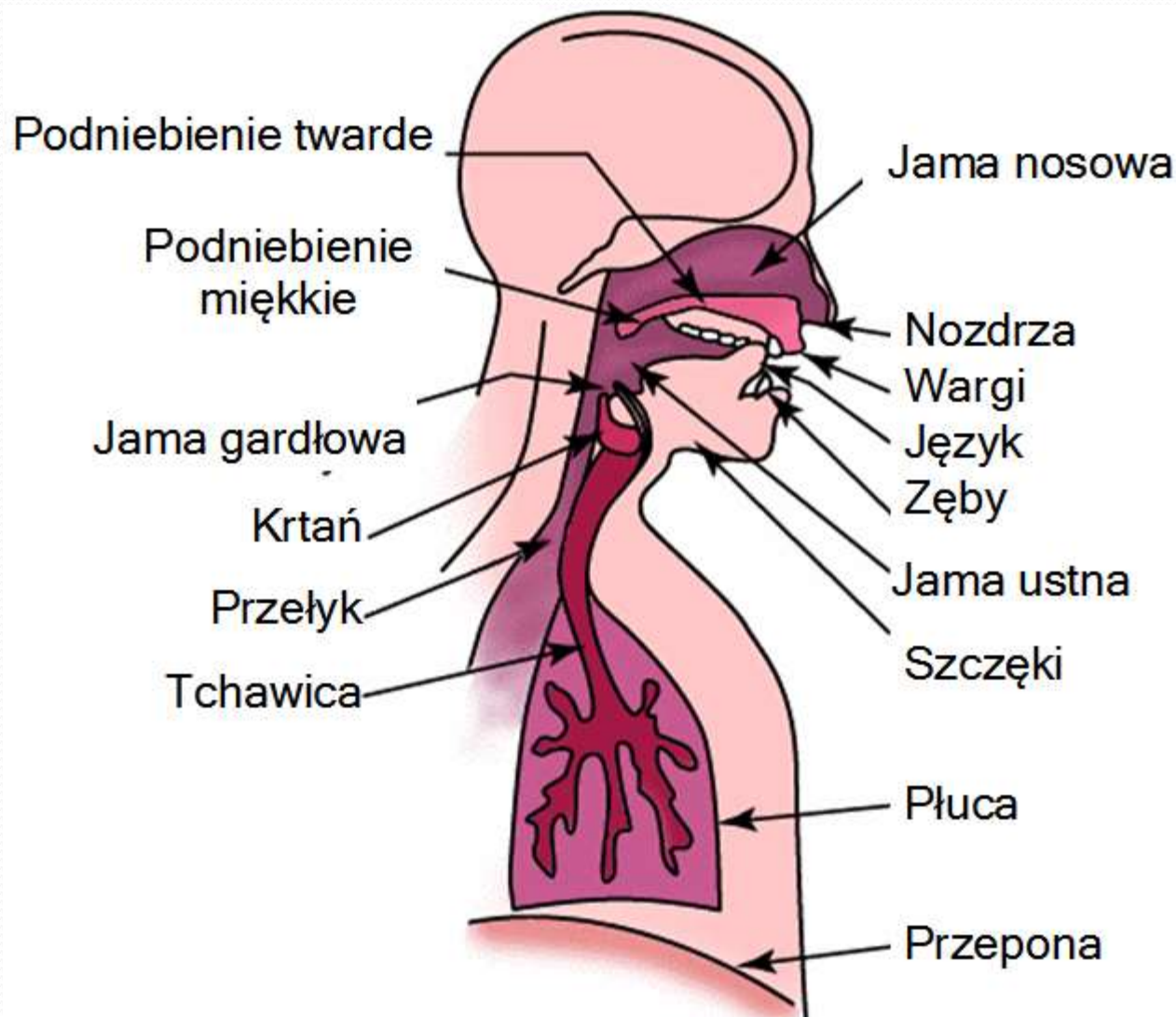
Języki na świecie



Procesy wytwarzania mowy

- Proces psychologiczny – przygotowanie informacji do przekazania (konceptualizacja),
- Proces neurologiczny – pobudzenie w ośrodkowym układzie nerwowym oraz na drodze eferentnej (ruchowej) mięśni narządu mowy,
- Proces fizjologiczny – artykulacja,
- Proces aerodynamiczny – przepływ powietrza i generowanie drgań o złożonej strukturze widmowo-czasowej.

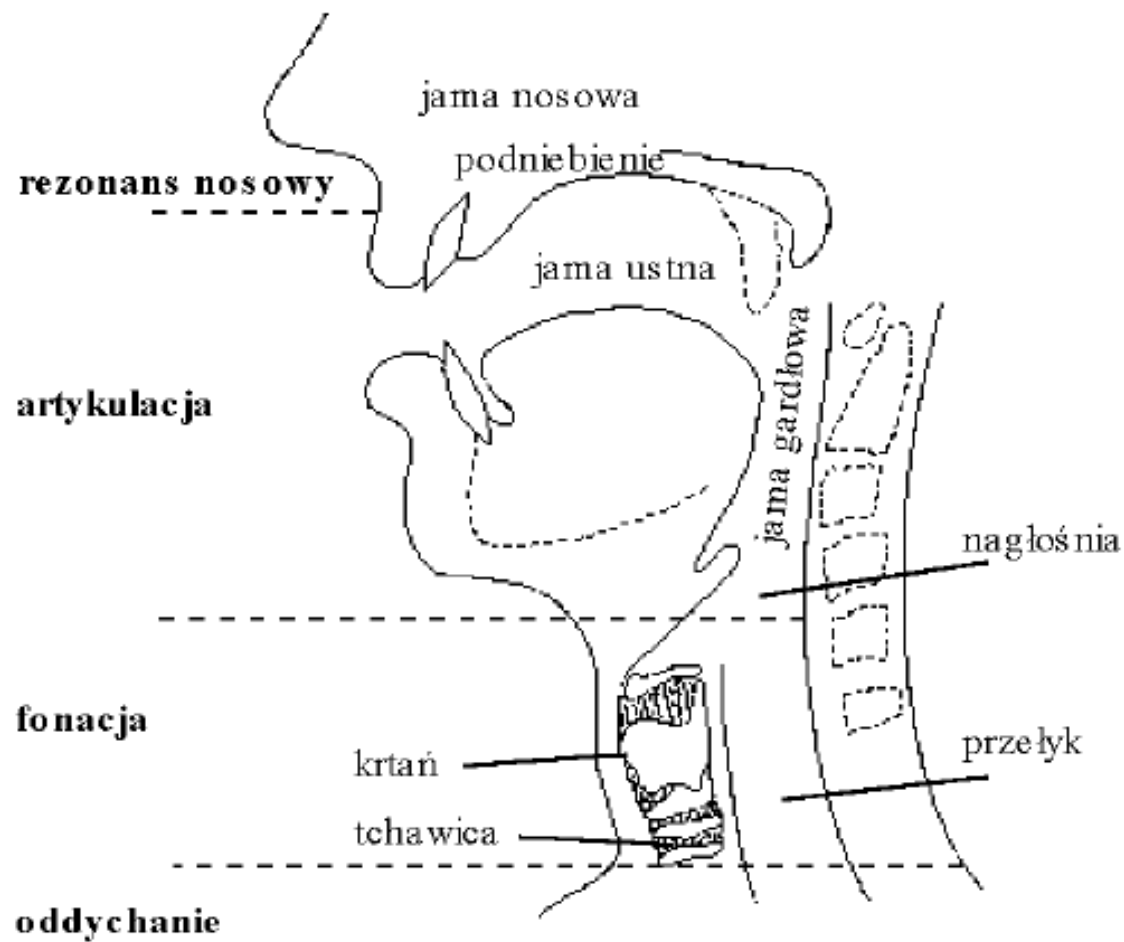
Narząd mowy – trakt głosowy



Wytwarzanie mowy

- Płuca, oskrzela i tchawica tworzą drogę doprowadzającą powietrze do krtani.
- W krtani znajdują się fałdy głosowe (pot. struny głosowe), które pod ciśnieniem powietrza zaczynają drgać i wytwarzają tzw. ton krtaniowy.
- Wygenerowany ton następnie jest „filtrowany” w dalszej części traktu głosowego. Na sposób filtracji wpływają takie narządy jak: język, języczek, podniebienie, zęby, usta, jama nosowa.

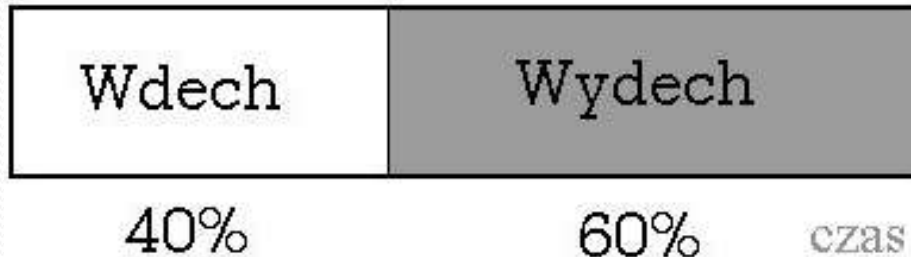
Wytwarzanie mowy



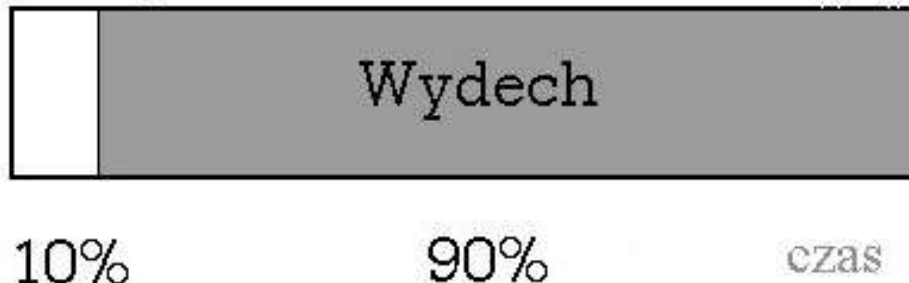
Cykle oddechowe

Ryszard Gubrynowicz

Cykl oddechowy w spoczynku



Cykl oddechowy podczas mówienia



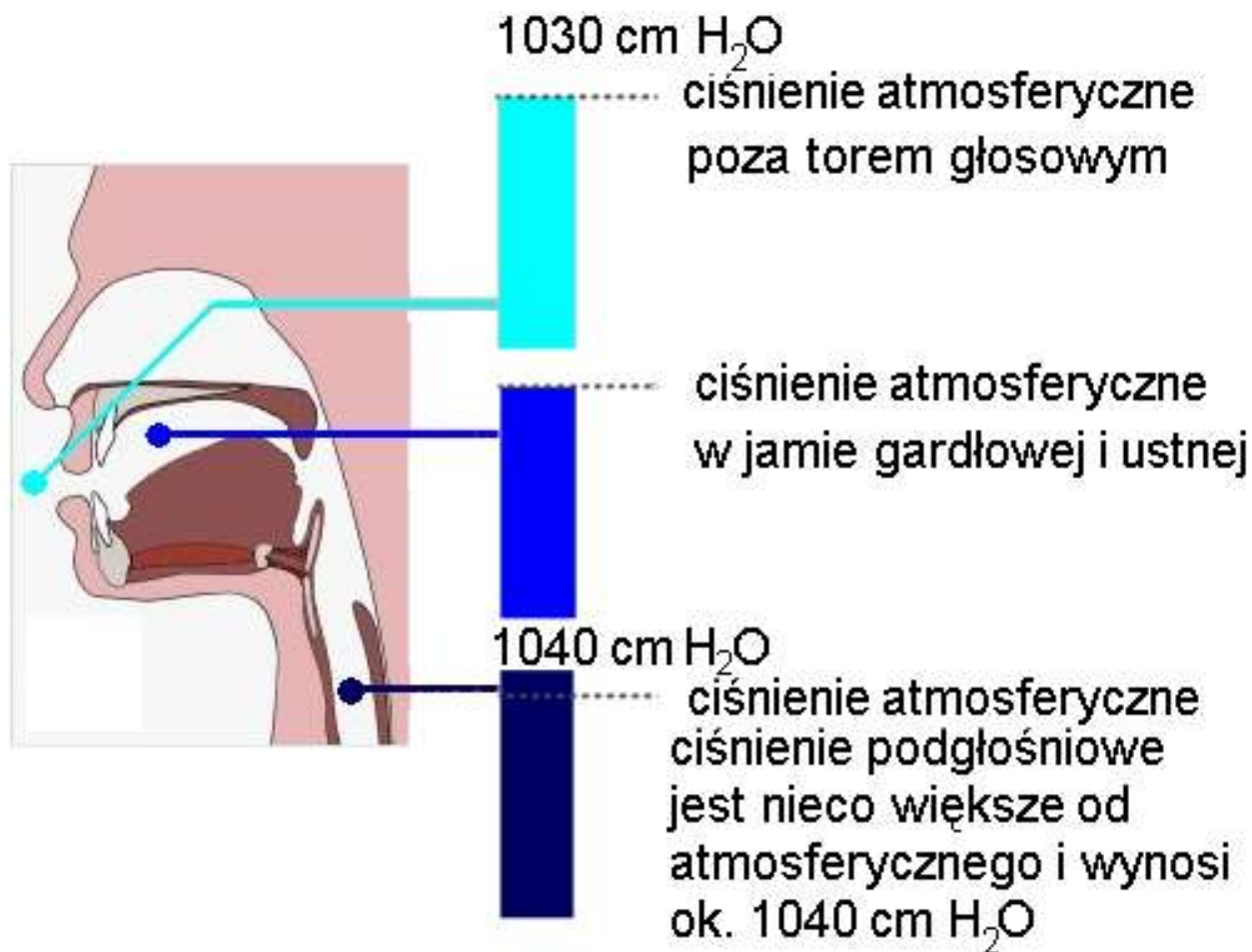
Max pojemność płuc – ok. 7 litrów

Pojemność minimalna – 2 litry stale w płucach.

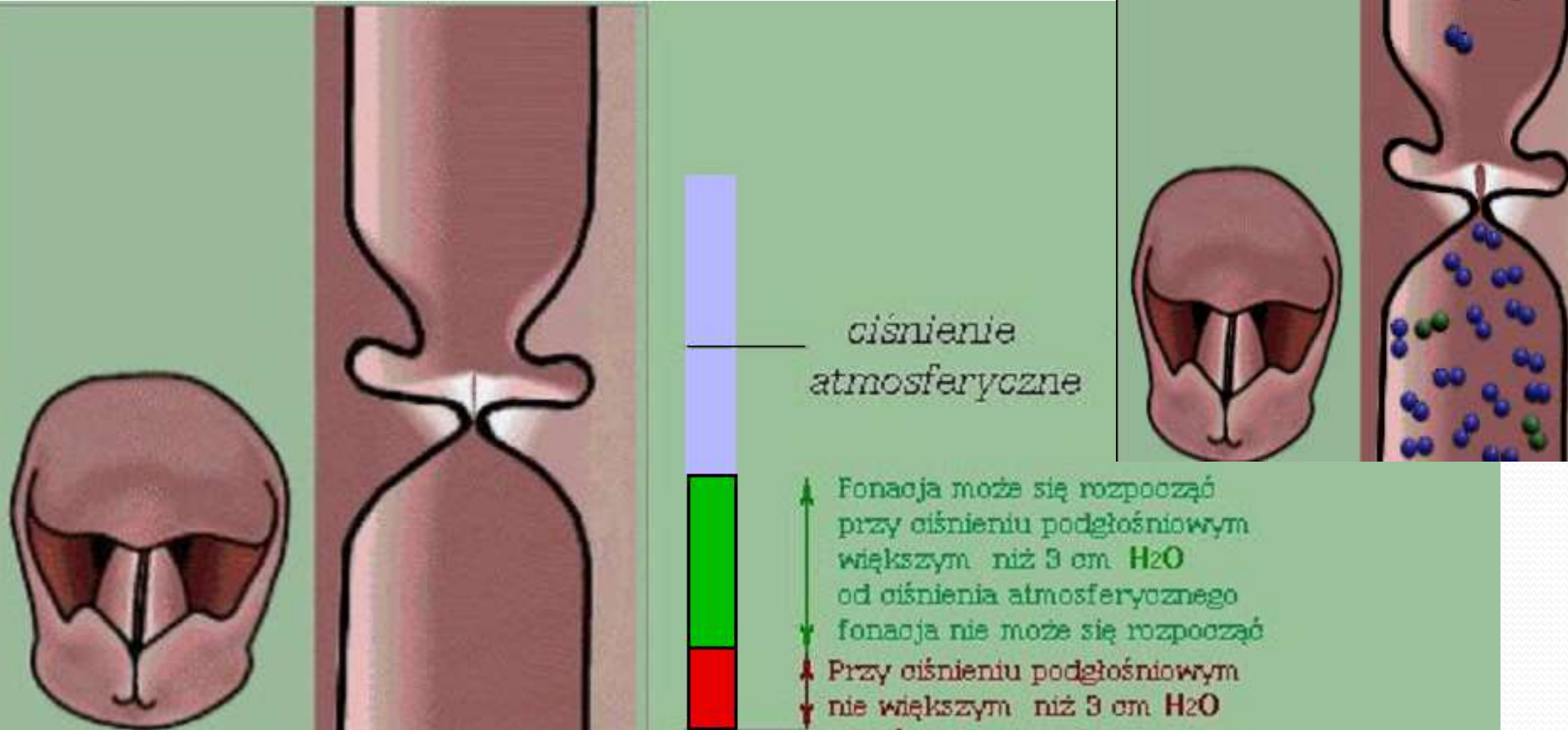
Objętość powietrza wymieniana podczas każdego cyklu oddechowego – 0.5 l

Częst. oddychania w stanie spoczynku – 12-20 cykli na minutę

Rozkład ciśnień powietrza w torze głosowym



efekt Bernoulliego



ciśnienie atmosferyczne

ciśnienie atmosferyczne

- ↑ Fonacja może się rozpocząć przy ciśnieniu podgłośniowym większym niż 3 cm H₂O od ciśnienia atmosferycznego
- ↓ fonacja nie może się rozpocząć
- ↑ Przy ciśnieniu podgłośniowym nie większym niż 3 cm H₂O od ciśnienia atmosferycznego
- ↓ fonacja nie może się rozpocząć

Cechy charakterystyczne

- Dwie warstwy wpływające na charakterystykę mowy:
 - Fizyczna – wynikająca z anatomicznych właściwości elementów traktu głosowego,
 - Psychiczna – indywidualny, wyuczony sposób fonacji i artykulacji.
- O charakterystycznej barwie głosu w dużej mierze decyduje ton krtaniowy, którego częstotliwość zależy od długości fałd głosowych w krtani.

Cechy mowy

- Mowę można analizować na kilku poziomach:
 - semantyczny – treść mowy,
 - osobniczy – cechy pozwalające zidentyfikować mówcę,
 - emocjonalny – poznanie stanu emocjonalnego, stanu zdrowia, pozycji społecznej itp.
 - prozodyczny – związany z akcentem, intonacją itp..

Sygnał mowy - struktura

- Sygnał mowy powstaje w wyniku **splotu** tonu krtaniowego (fonacja) i odpowiedzi traktu głosowego (artykulacja).
- Fonacja odpowiada za proces **wytwarzania energii** dźwięku, natomiast artykulacja kształtuje **charakterystykę częstotliwościową**.
- Niestety proces **rozplotu** nie jest operacją trywialną

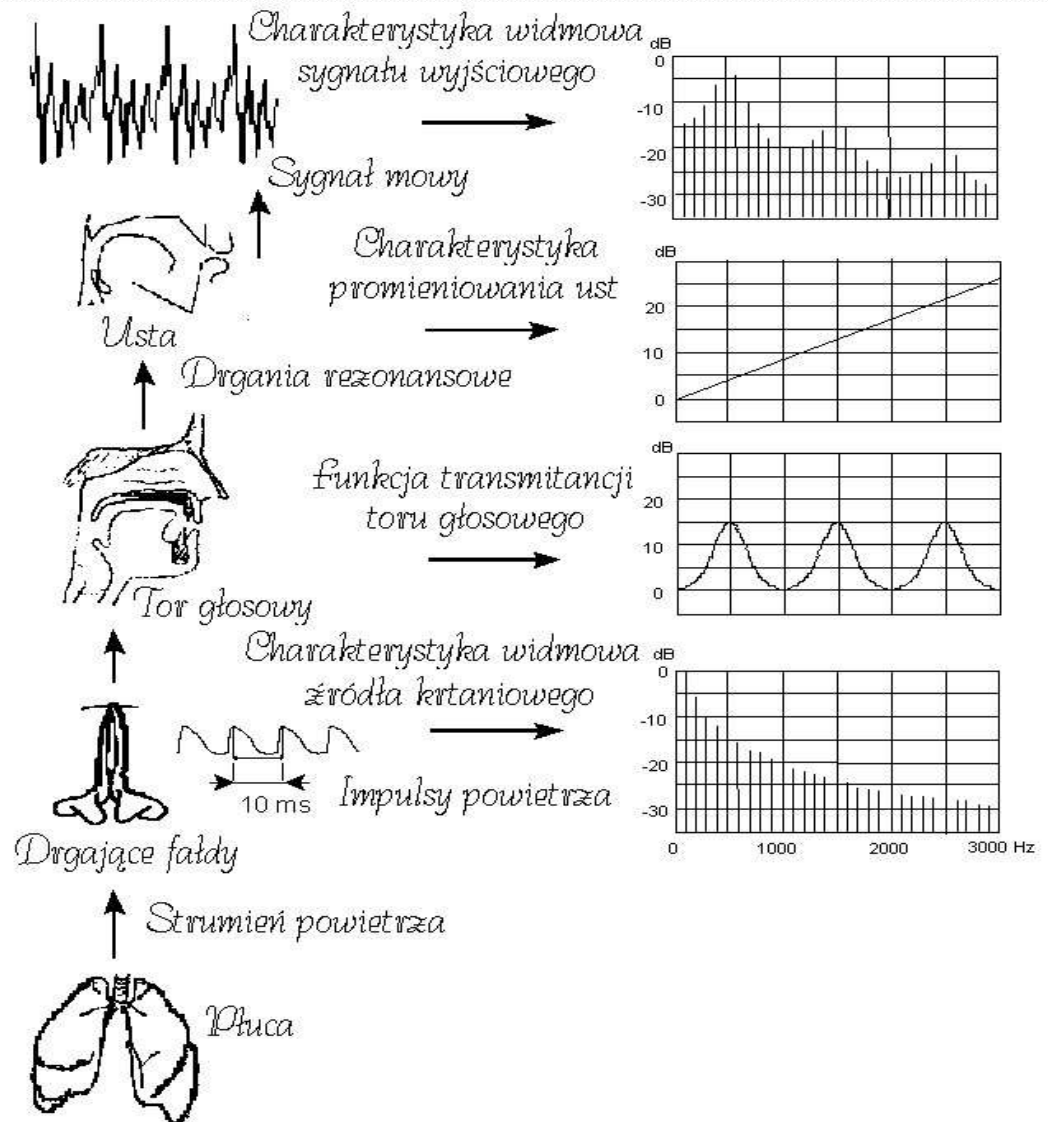
Sygnał mowy

- Zależność poziomu ciśnienia sygnału mowy od częstotliwości oznacza, że w określonych zakresach częstotliwości, składowe widma dźwięków mowy przybierają wartości znacznie wyższe niż w pozostałych zakresach częstotliwości.
- Te zakresy częstotliwości, w których składowe widma przyjmują maksymalne wartości nazywa się **formantami**, zaś odpowiadające im częstotliwości – **częstotliwościami formantowymi**.

Sygnał mowy

- Sygnał mowy charakteryzuje duża grupa parametrów akustyczno-fonetycznych, przy czym nie wszystkie parametry biorą bezpośredni udział w procesie percepcji.
- Redundancja danych pozwala percypować dźwięk nawet w trudnych warunkach akustycznych.
- Tony podstawowe mowy są w zakresie od 74 Hz do 1056 Hz. Istotne znaczenie dla zrozumiałości mowy mają częstotliwości nawet do 10 kHz.
- Poziom dźwięku mowy: 34 dB – 94 dB.

Proces formowania sygnału mowy



Poziomy percepcji mowy

- Proces percepcji mowy jest procesem złożonym i składa się z kilku faz:
 - Aerodynamiczna – drgania powietrza w przewodzie słuchowym,
 - Akustomechaniczna – przenoszenie drgań od błony bębenkowej do ślimaka,
 - Neurologiczna – przenoszenia i przetwarzanie impulsów w ośrodkowym układzie nerwowym,
 - Psychologiczna – rozpoznanie i zrozumienie przekazanej informacji.

Specyfika percepcji mowy

- Percepcja dźwięków mowy nie zależy wyłącznie od ich struktury czasowo-widmowej, ale również od treści lingwistycznej.
- Badania neurofizjologiczne pokazują, że u większości osób percepcja mowy odbywa się w lepszym stopniu za pomocą ucha prawego zaś w przypadku muzyki jest odwrotnie (dźwięki mowy są lepiej dekodowane przez lewą półkulę mózgu)

Język naturalny a sygnał mowy

- Aby dźwięki mowy miały określone znaczenie musi istnieć wzajemne przyporządkowanie między strukturą akustyczną sygnału mowy a przekazywaną informacją.
- Badaniem struktury dźwiękowej języka naturalnego zajmuje się **fonetyka**. Nauka ta zajmuje się selekcją dźwięków elementarnych, z których przez złożenie, powstają określone formy językowe.

Dźwięki elementarne mowy

- **Fonem** - minimalny segment dźwiękowy mowy, który może odróżniać znaczenie, lub inaczej klasa dźwięków mowy danego języka o różnicach wynikających wyłącznie z charakteru indywidualnej wymowy lub kontekstu.
- **Alofon** - wariant fonemu odróżniający się od innego alofonu cechami fonetycznymi a nie funkcją.
- **Difon** (diafon) –przejście (złączenie) dwóch fonemów.
- **Mikrofonem** – jednostka sygnału mowy o stałej długości (20-40 ms.)

Fonemy

- W języku polskim można wyróżnić 37 fonemów + 2 samogłoski nosowe
- Fonemy języka polskiego można sklasyfikować za pomocą binarnych cech dystynktywnych:
 - Spółgłoski – samogłoski,
 - Ponadkrtaniowe – krtaniowe,
 - Nosowe – ustne,
 - Łagodne-raptowne
 - Skupione-rozproszone
 - Jasne – ciemne,
 - Niskotonowe-wysokotonowe,
 - Długie-krótkie,
 - Dźwięczne - bezdźwięczne

Fonemy

- dźwięki o charakterze quasiperiodycznym:
 - 1) samogłoski sylabiczne (a, e, i, o, u, y)
 - 2) półsamogłoski niesylabiczne (j, ł)
 - 3) spółgłoski nosowe (m, n, ń, ą, ę)
 - 4) spółgłoski boczne (l)
- dźwięki o charakterze przebiegów nieperiodycznych - szumowych:
 - 1) spółgłoski bezdźwięczne trące (f, s, sz, ś, h)
 - 2) spółgłoski bezdźwięczne zwarto-trące (c, ć, cz)
- dźwięki o charakterze przebiegów nieperiodycznych - quasi-impulsowych:
 - 1) spółgłoski zwarte dźwięczne (b, d, g)
 - 2) spółgłoski zwarte bezdźwięczne (p, t, k)

Fonemy

- dźwięki o charakterze przebiegów będących superpozycją quasiperiodycznych i nieperiodycznych:
 - 1) spółgłoski trące dźwięczne (w, z, ź, ż)
 - 2) spółgłoski zwarto-trące dźwięczne (dz, dź, dż)



Dziękuję za uwagę