



Universidade do Minho
Escola de Ciências

II - Fala

António Mário Almeida

Departamento de Física

Universidade do Minho

Sumário

1. Definições
 1. Sistema respiratório
 2. Sistema fonador
 3. Tracto vocal
 4. Articulação
 5. Cordas vocais
2. A fala
 1. Fonemas
 2. Aspetos fonéticos
 1. Vocalização
 2. Local de articulação
 3. Modo de articulação
 3. Sinal acústico
3. Fonação / O som laríngeo
4. Articulação – o tracto vocal como filtro

Sistema respiratório

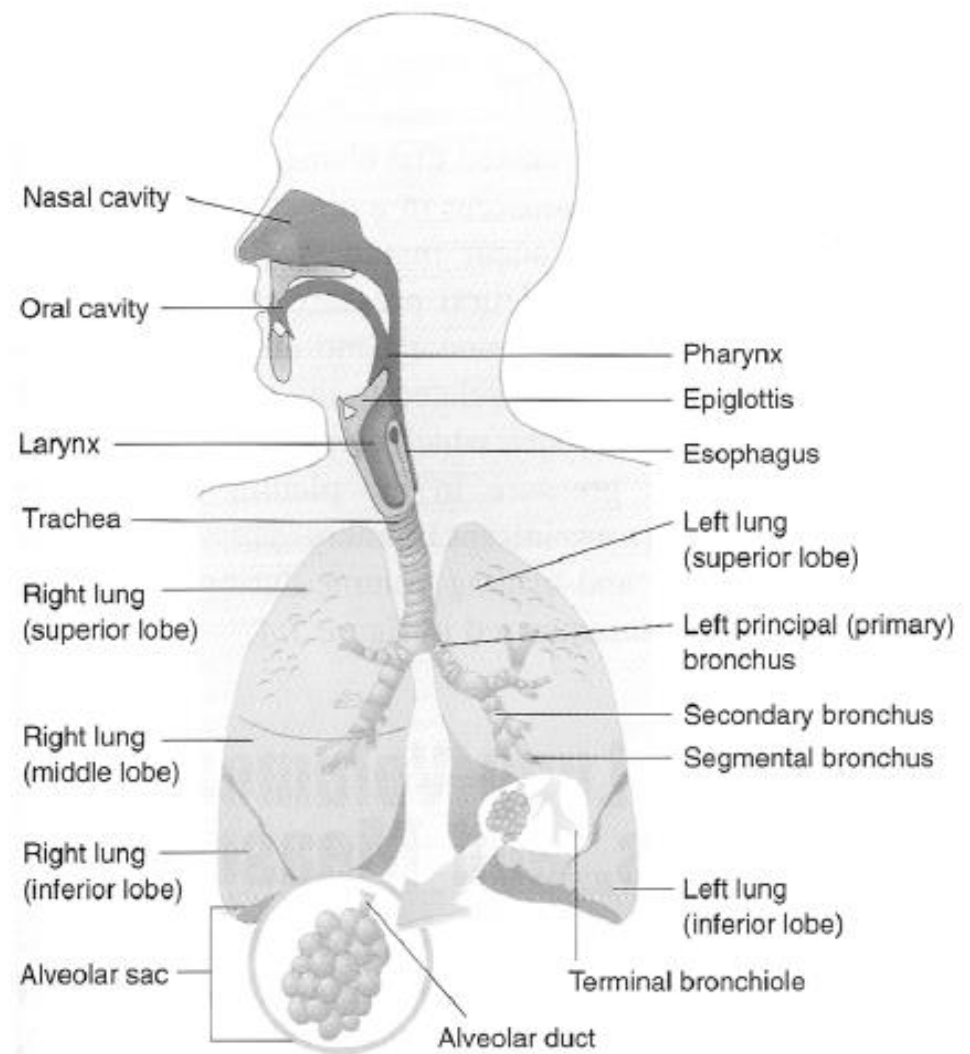


Fig. 9.1. Diagram of parts of the respiratory system. (These components are also important in voice production (Chap. 10). The vocal cords (or vocal folds) used in speaking are in the larynx.) (From [425]. Used with permission)

Sistema fonador

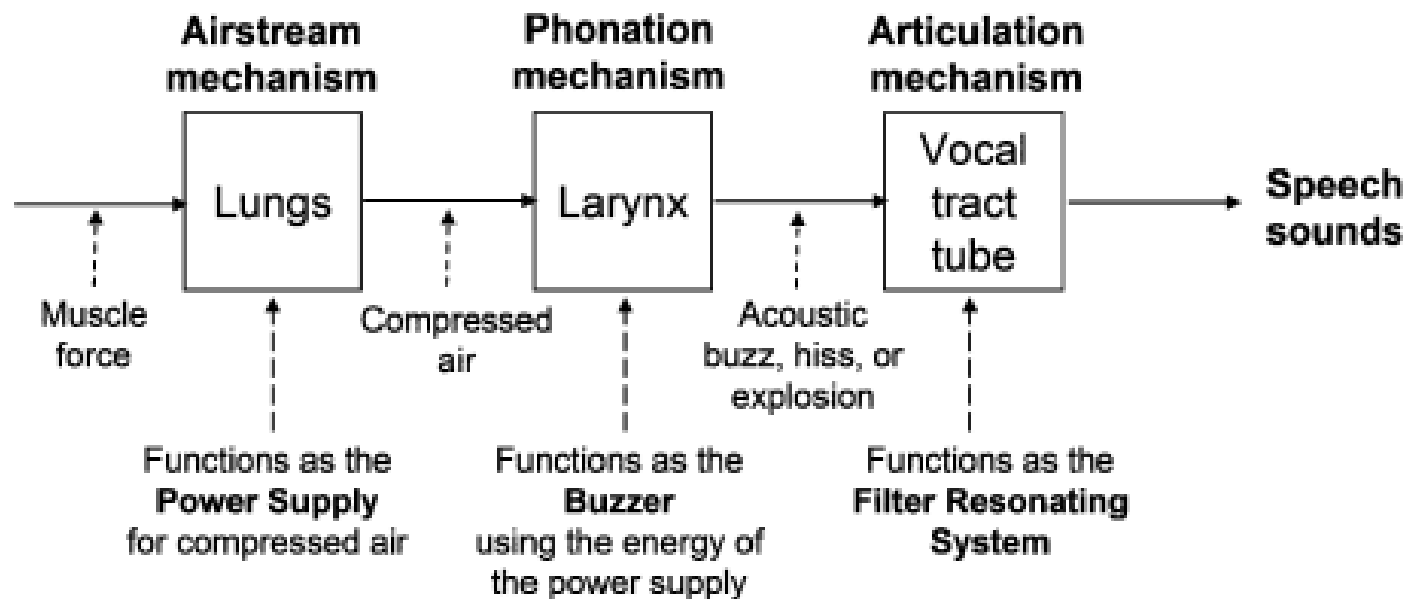
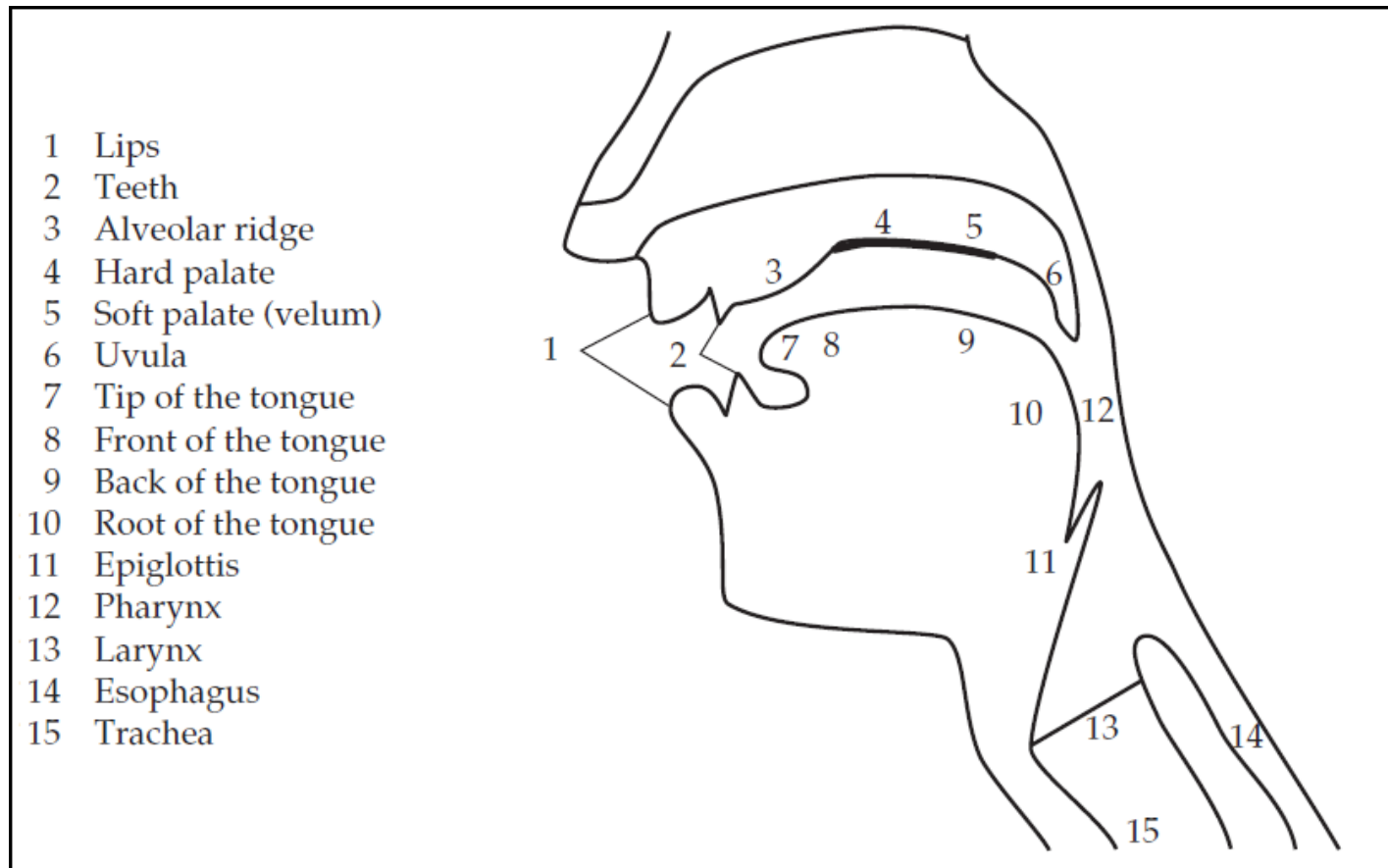


Fig. 10.10. The three neuromuscular systems in voice production. (Based on [460])

Mulheres: faringe 6,3 cm; cavidade oral 7,8 cm; tracto vocal de 14,1 cm.
Homens: faringe 8.9 cm; cavidade oral 8.1 cm; tracto vocal de 16,9 cm.

<http://talleringlesaq.blogspot.pt/2011/05/vocal-tract.html>

Tracto vocal: laringe (onde estão as cordas vocais), faringe, cavidade oral e cavidade nasal



A forma e o comprimento do tracto vocal podem ser modificados.

As alterações produzidas no tracto vocal no processo de produção de som constituem a

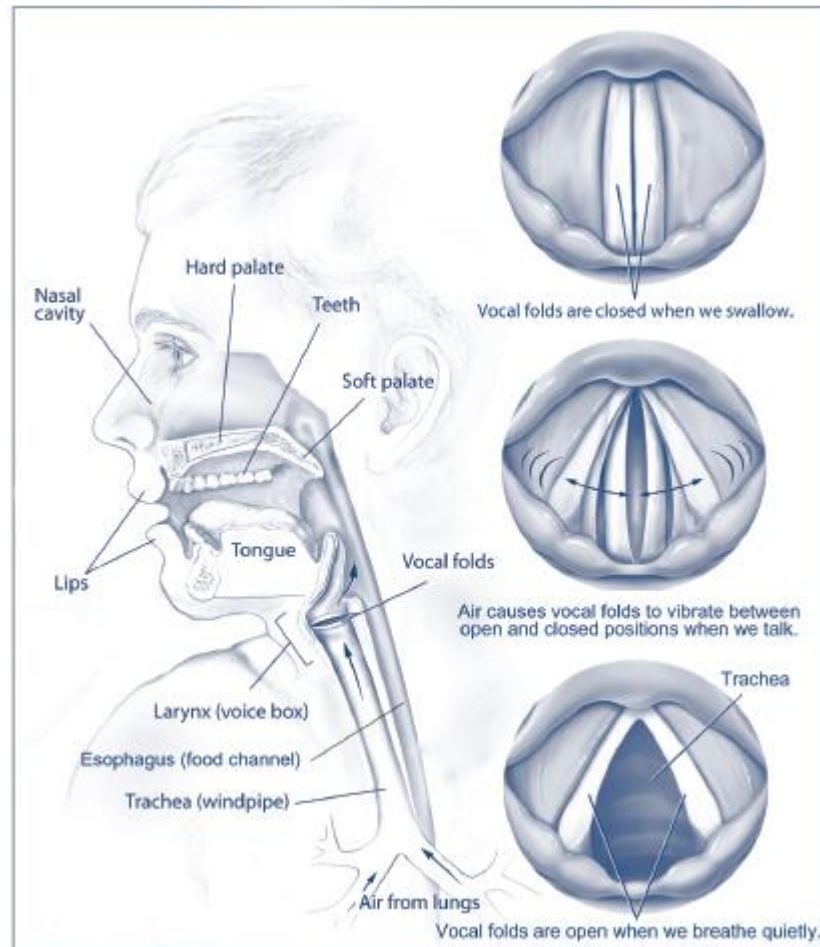
Articulação

Os movimento do palato móvel, da língua, dos lábios e do maxilar podem alterar a forma do tracto vocal em certos pontos - **articuladores**.

Mudar a posição dos lábios e subir ou baixar a laringe permite modificar o comprimento do tracto vocal.

As cordas vocais são duas pregas musculares e membranosas situadas na laringe, revestidas por uma fina camada muscosa e constituem o elemento vibratório no processo de produção de som.

Cordas vocais



2 - A fala

Os elementos básicos do discurso são classificados em:

- 1 – fonemas
- 2 – aspetos fonéticos (como são produzidos os sons)
- 3 – sinal acústico

A pronúncia do português europeu

http://cvc.instituto-camoes.pt/cpp/acessibilidade/capitulo2_1.html

1 - fonemas

São os segmentos mais curtos da fala.

Em inglês americano existem 14 vogais e 24 consoantes.

Já a descrição fonética do português de Portugal compreende

18 (?) vogais e http://www.fonologia.org/fonetica_vogais.php

19 (?) consoantes http://www.fonologia.org/fonetica_consoantes.php

As vogais são produzidas com o tracto vocal relativamente aberto, com diferentes formas.

As consoantes são produzidas por constrições/obstruções no tracto vocal.

Modo de articulação: configuração do tracto vocal durante a emissão vocal.

Ponto de articulação: zona do tracto vocal onde existe maior constrição.

2 – aspetos fonéticos

Há três aspetos a considerar na produção de fonemas:

- a) vocalização;
- b) local de articulação;
- c) modo de articulação.

a) Vocalização

- **Sons vozeados:** sons produzidos a partir de jatos de ar através das cordas vocais vibrantes.

Todas as vogais são sons vozeados.

Muitas consoantes também são: “d”, “m” e “v”.

- **Sons não vozeados:** sons produzidos pela passagem por constrições ou arestas formadas pela língua, dentes, lábios ou palato e sem a vibração das cordas vocais.

Por exemplo: “t”, “f”

b) Local de articulação

Classificação das obstruções que permitem produzir consoantes :

- **Bilabial** (ambos os lábios): “p”, “b”, “m”.
- **Labiodental** (lábio inferior e dentes frontais superiores): “f”, “v”.
- **Dental** (dentes): “th” (inglês)
- **Alveolar** (gengiva superior perto dos dentes): “d”, “t”, “s”, “z”, “n”, “r”, “l”.
- **Palato-alveolar** (palato): “ch”, “j”.
- **Velar** (palato móvel): “k”, “g”, “ng”.
- **Glotal** (glote): “h”.

Para uma leitura complementar, informação adicional e classificação completa:

http://www.fonologia.org/quadro_fonetico.php

c) Modo de articulação

Quanto ao modo de articulação, as consoantes de português de Portugal classificam-se como:

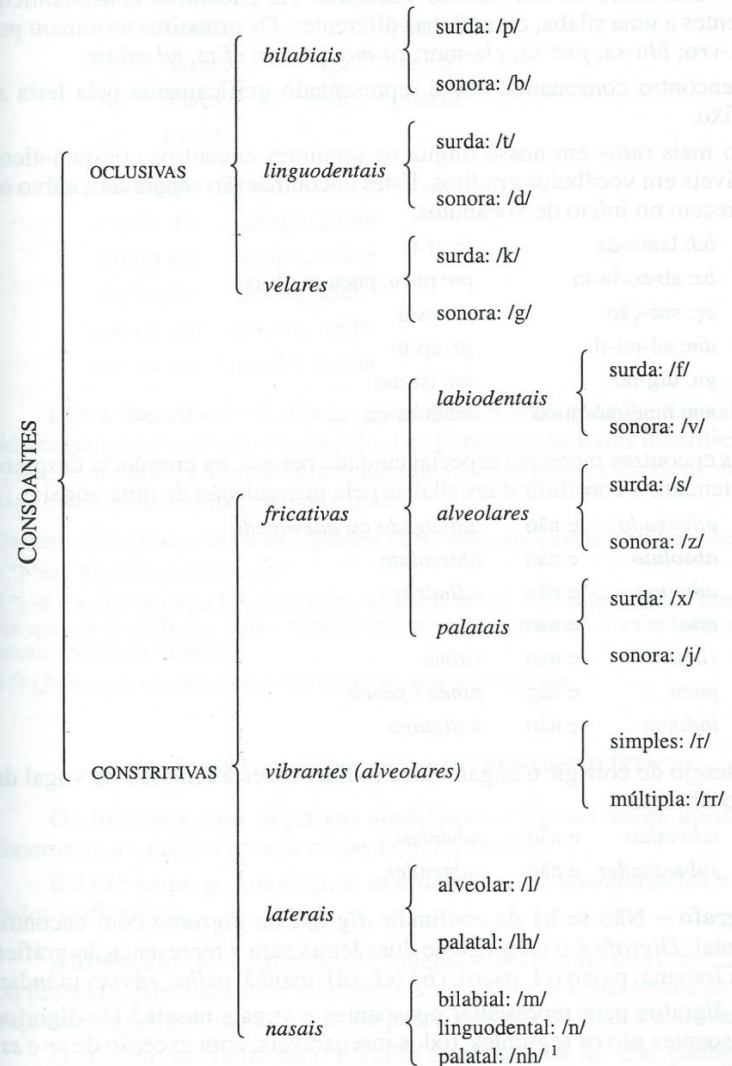
- **Oclusivas** (*plosive* ou *stop*) – passagem do ar totalmente obstruída nas cavidades supraglóticas; o som produz-se no momento em que os articuladores se afastam. Ex.: “p”, “t”, “k” (não vozeadas); “b”, “d”, “g” (vozeadas).
- **Fricativas** - passagem do ar parcialmente obstruída com produção simultânea de ruído. Ex.: “f”, “v”. Fricativos mais agudos como “s” ou “z” são **sibilantes**.
- **Laterais** – o fluxo de ar é obstruído pela língua contra os alvéolos ou parte central do palato ficando livres passagens pelas partes laterais da cavidade bucal. Ex.: “l”, “lh”.
- **Vibrantes** – o articulador móvel vibra contra o articulador fixo. Podem ser simples (“r”) ou múltiplas (“rr”).
- **Nasais** – o palato móvel desce e o ar flui repentinamente pela cavidade nasal. Ex.: “m”, “n”, “nh”.

c) Modo de articulação

Muito importantes no inglês e no português do Brasil são ainda:

- **Africativas** – início oclusivo seguido de fricativo. **Ex.:** “**ch**”, “**j**”.
- **Aproximantes** – produzidas ao mover um articulador em direção ao outro sem criar uma oclusão. **Ex.:** “**w**”, “**y**”, “**r**” e “**l**”.

QUADRO DA CLASSIFICAÇÃO DAS CONSOANTES



1 Para fugir a uma oposição errônea *surda/sonora/nasal*, preferimos, ainda com a aquiescência da NGB, colocar as *nasais* entre as constritivas. Há autores que fazem das *nasais* uma classe à parte, ou as põem entre as oclusivas, critérios também defensáveis.

3 – sinal acústico

O sinal acústico é o conjunto das frequências acústicas e das intensidades ao longo do tempo.

O espectrograma da figura (a) contém informação mais útil do que o sonograma (b) para a mesma expressão.

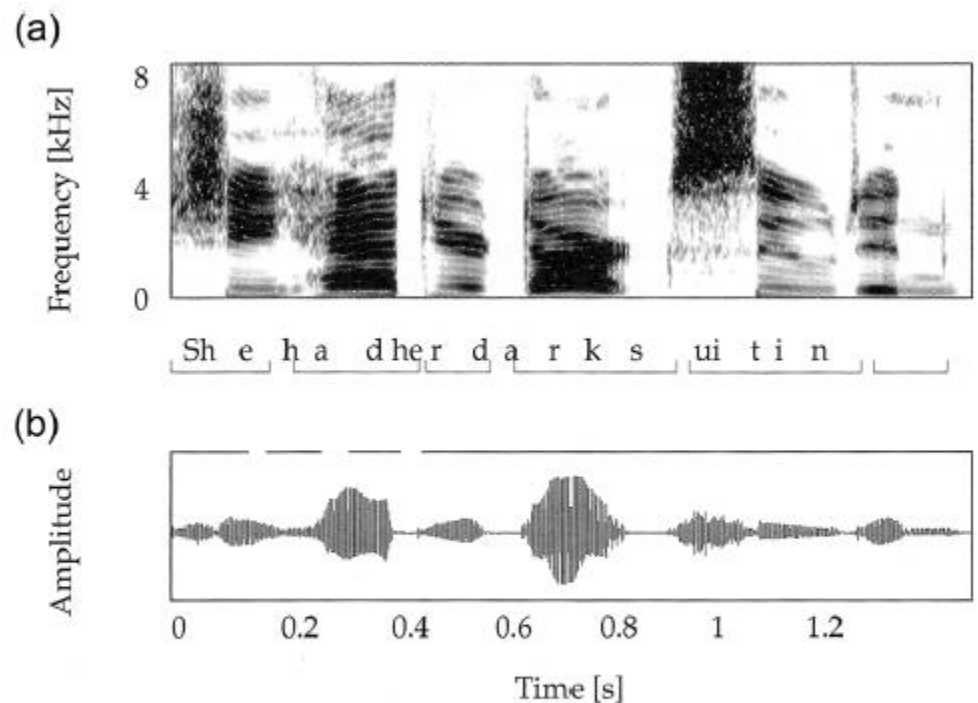
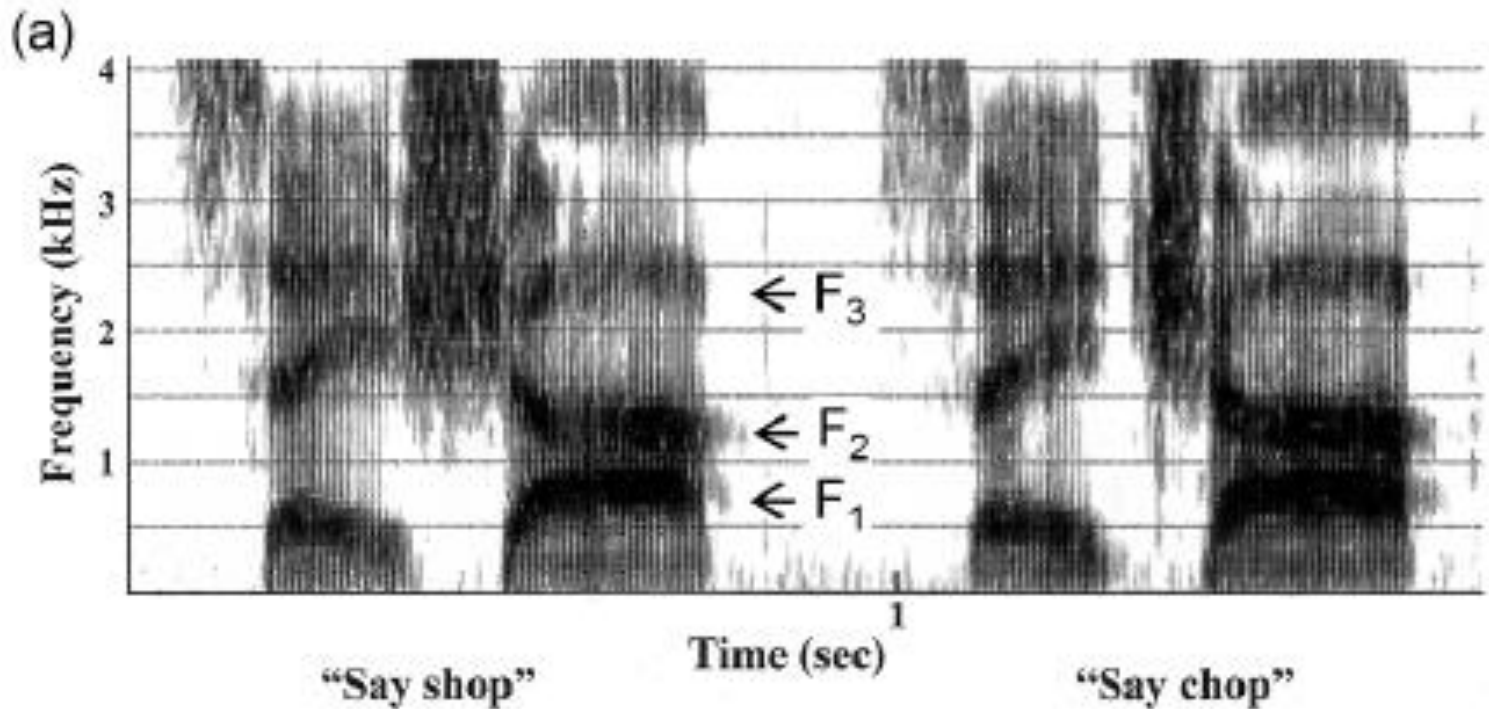
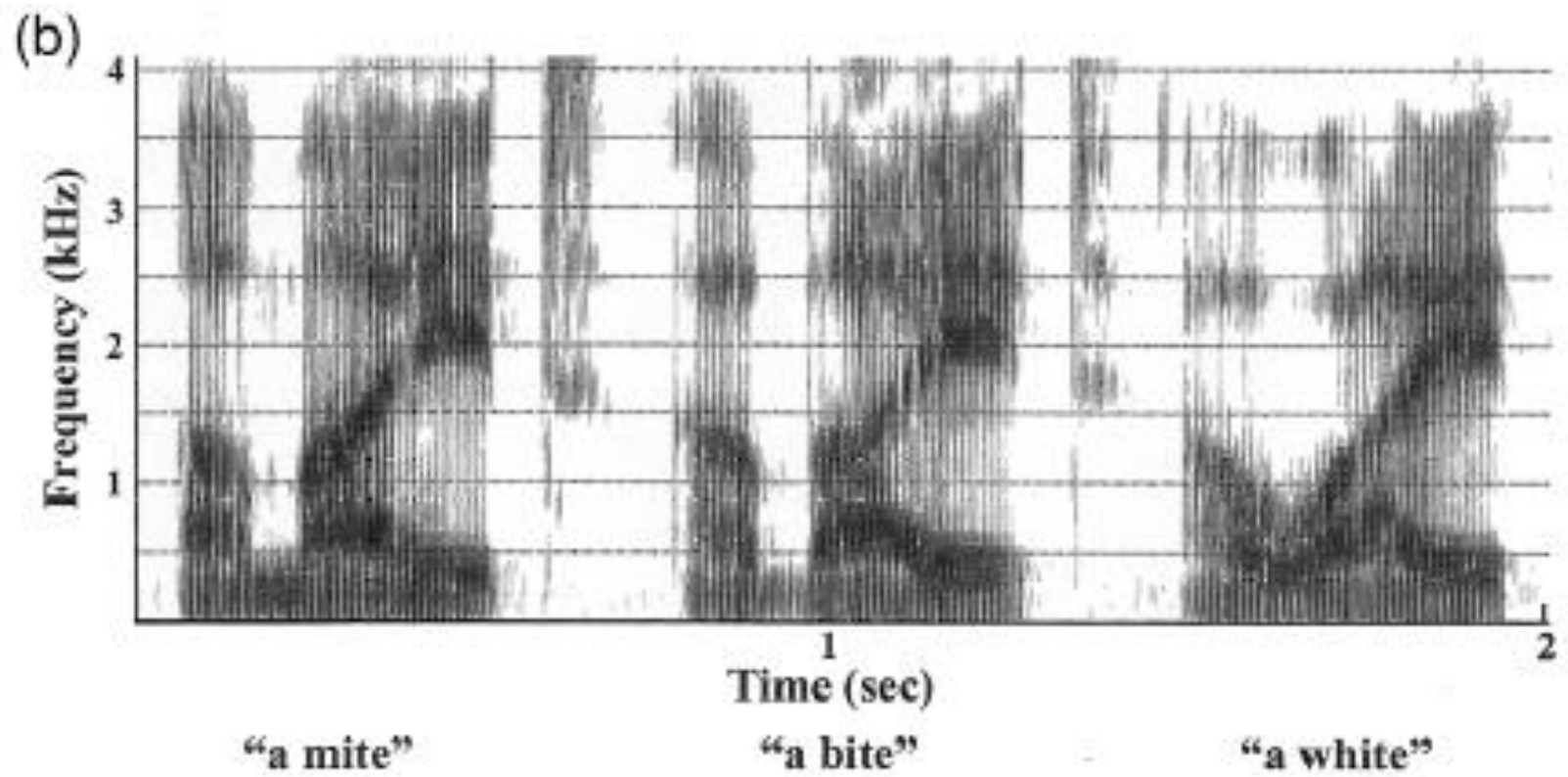
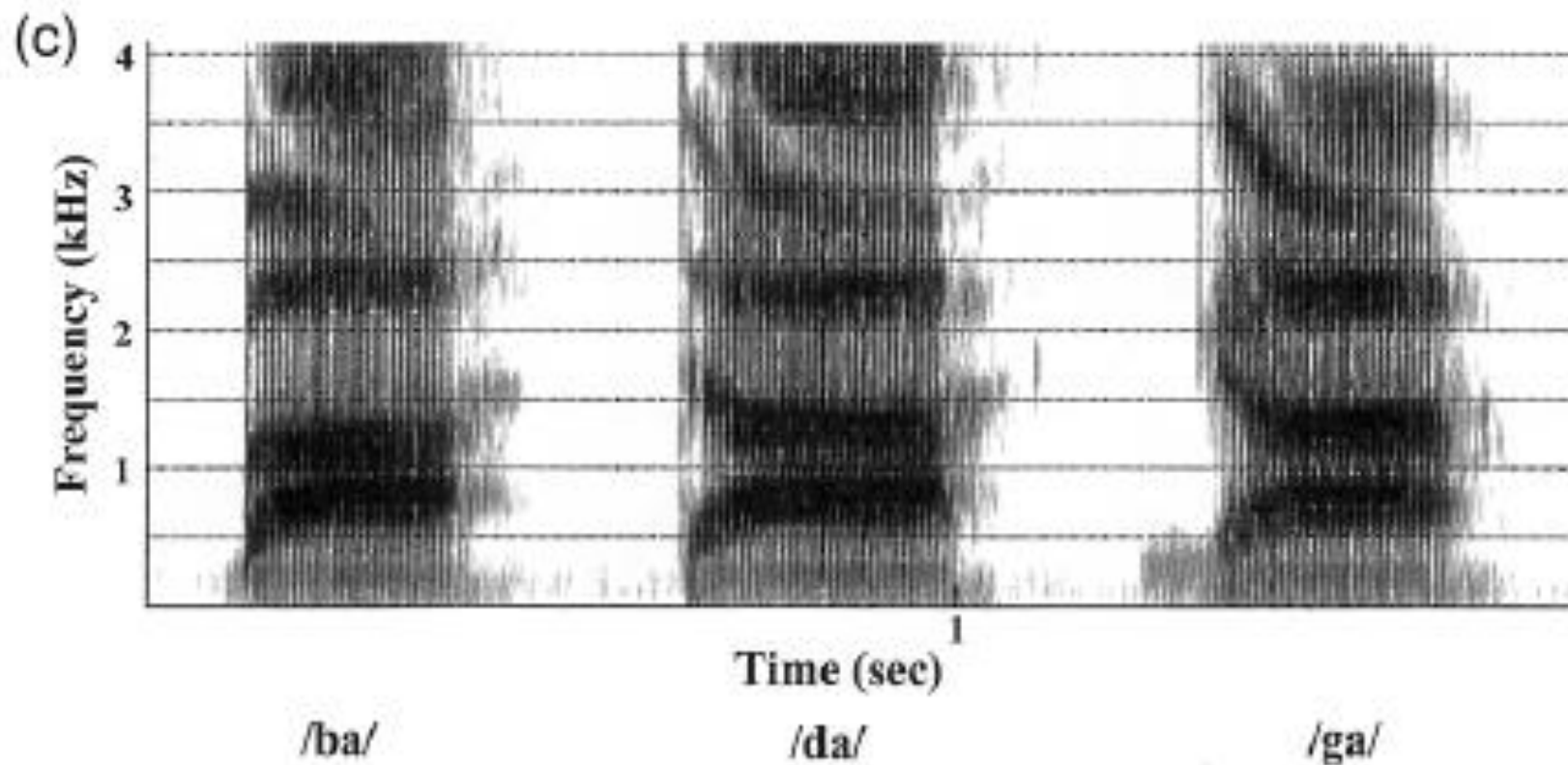


Fig. 10.8. Sound spectrogram of a female speaker saying "She had her dark suit in..." in (a), with the corresponding time-domain waveform signal in (b). (From [449])



- Diferenças entre “sh” e “ch” no espectro e na duração
- Pequeno silêncio antes de “ch”
- Bandas características das vogais (**formantes**):
 - “ei” de “say”: 500, 1700 e 2500 Hz;
 - “o” de “shop” ou “chop”: 750, 1250 e 2400 Hz.





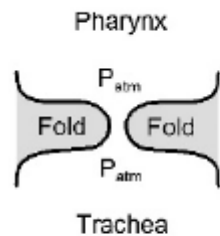
Para as 3 sílabas, a parte inferior do espectro da consoante aumenta para F1 da vogal, mas varia de forma diferente para as segunda e terceira formantes, o que permite distinguir os três sons.

“ba” – F2 e F3 aumentam para chegar à vogal;

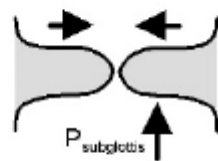
“da” – F2 e F3 caem para a vogal;

“ga” – F2 cai, enquanto F3 sobe para a vogal.

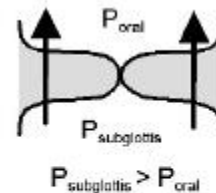
3 - Fonação (The acoustic buzzer – O zumbido acústico): o som laríngeo



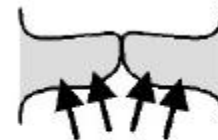
1. Folds at rest.



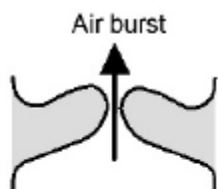
2. Folds move to midline because of muscle contractions (adduction).



3. Pressure below the glottis (subglottis) rises above the oral pressure.



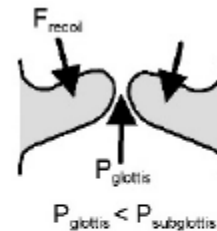
4. The pressure difference begins to force the folds open.



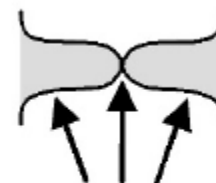
5. The folds open in an explosive manner, which rapidly releases a burst of air.



6. The air burst creates an overpressure which creates an acoustic shockwave that moves up the vocal tract at the speed of sound.



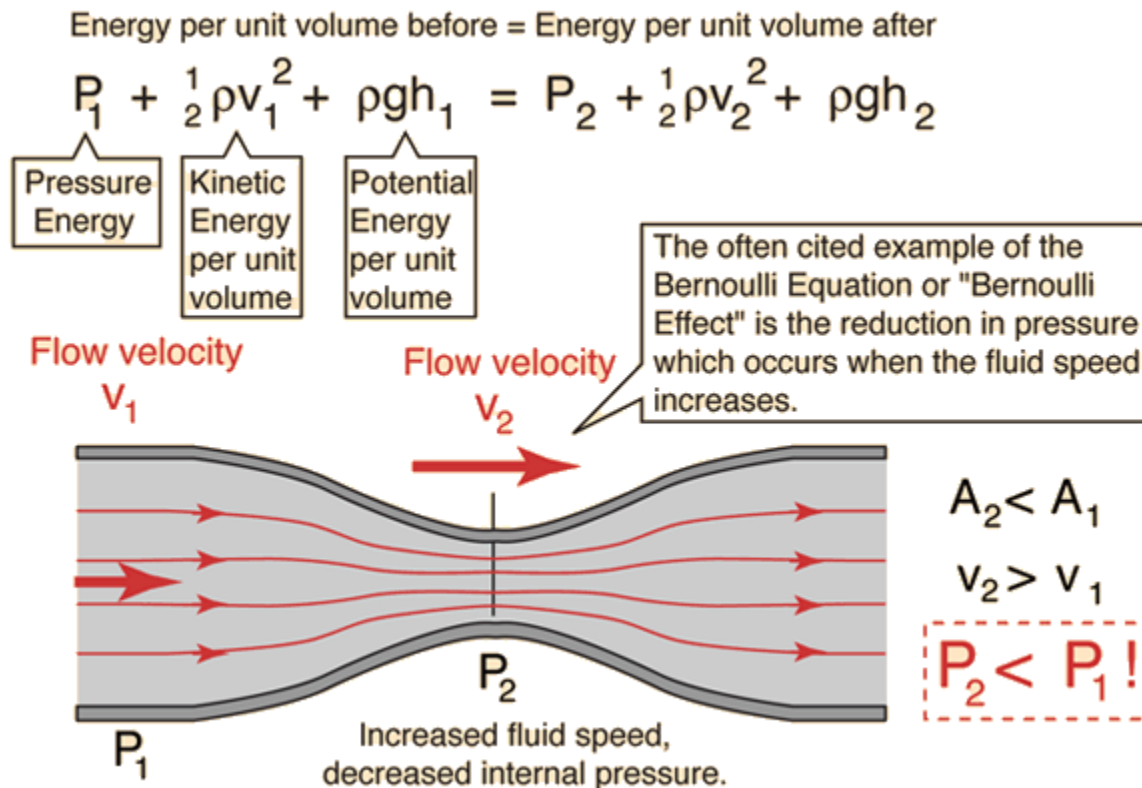
7. The folds begin to rebound because of muscle-induced recoil and the Bernoulli effect.



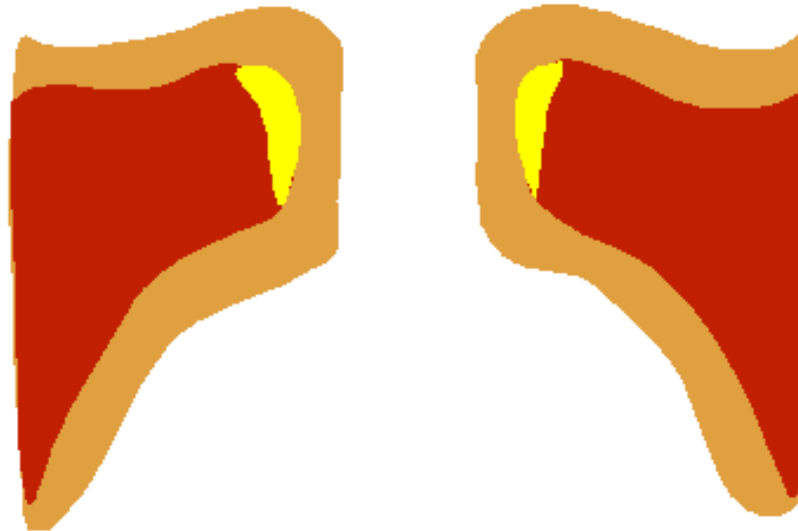
8. The folds close and the pressure beneath them begins to increase (as in part 3), the folds open again, and this cycle repeats over and over again.

Equação de Bernoulli

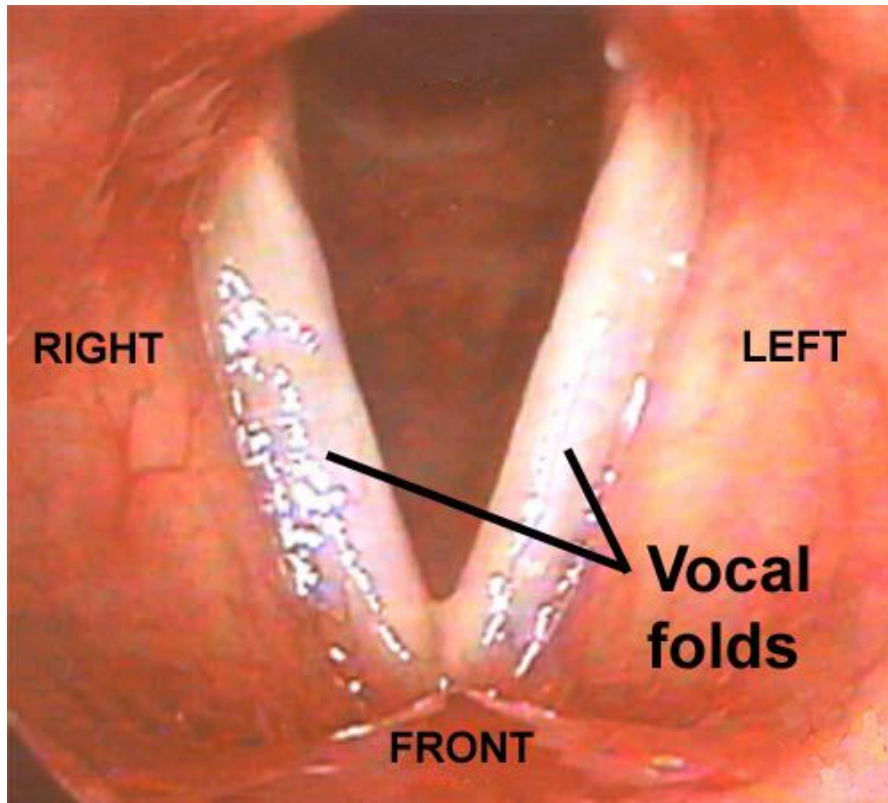
<http://hyperphysics.phy-astr.gsu.edu/hbase/pber.html#beq>



- As frequências do som laríngeo gerado dependem da massa, das dimensões e da tensão a que estão sujeitas as cordas vocais .
- Tipicamente, a fundamental é de ~ 125 Hz para os homens e ~ 250 Hz para as mulheres.

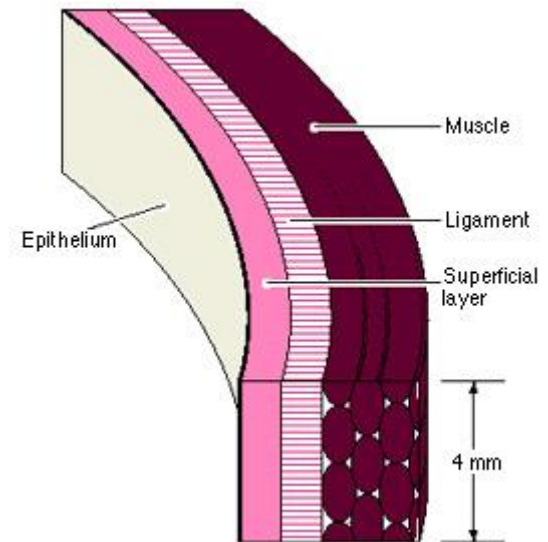
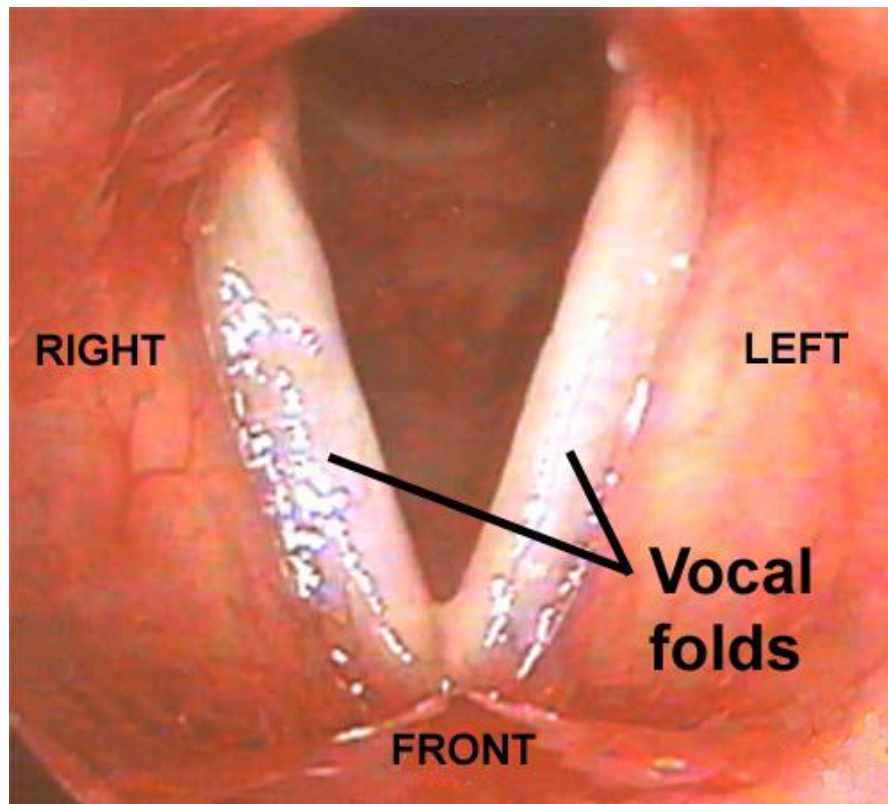


<http://www.singlikeastar.com/how-your-voice-works/>



The length of the vocal folds varies between 12.5mm and 17.5mm in females and 17mm and 25mm in males. These differences in length are partly responsible for the pitch range differences between men and women. The folds can also vary in thickness, which influences pitch range and quality.

Each vocal fold can be modeled as a two-dimensional ribbon-like object, with vibrations along the length of the fold and in a perpendicular direction across the height of the fold. To first-order, each of these vibrations can be treated with our one-dimensional models of vibration.



Cross-section through vocal fold

Fala: frequência fundamental

Valores médios da frequência fundamental (Titze, 1994): recém-nascidos, 500 Hz; criança de oito anos, 300 Hz; adultos fem., 200 Hz; adultos masc., 125 Hz.

Guimarães, 2002

Frequência fundamental medida em indivíduos adultos (18 a 65 anos) sem patologias vocais

	vogal			leitura em	
	/a/	/i/	/u/	voz alta	conversação
sexo feminino	211 Hz	221 Hz	225 Hz	190 Hz	187 Hz
sexo masculino	118 Hz	128 Hz	123 Hz	110 Hz	109 Hz

Um pouco da teoria de análise de Fourier

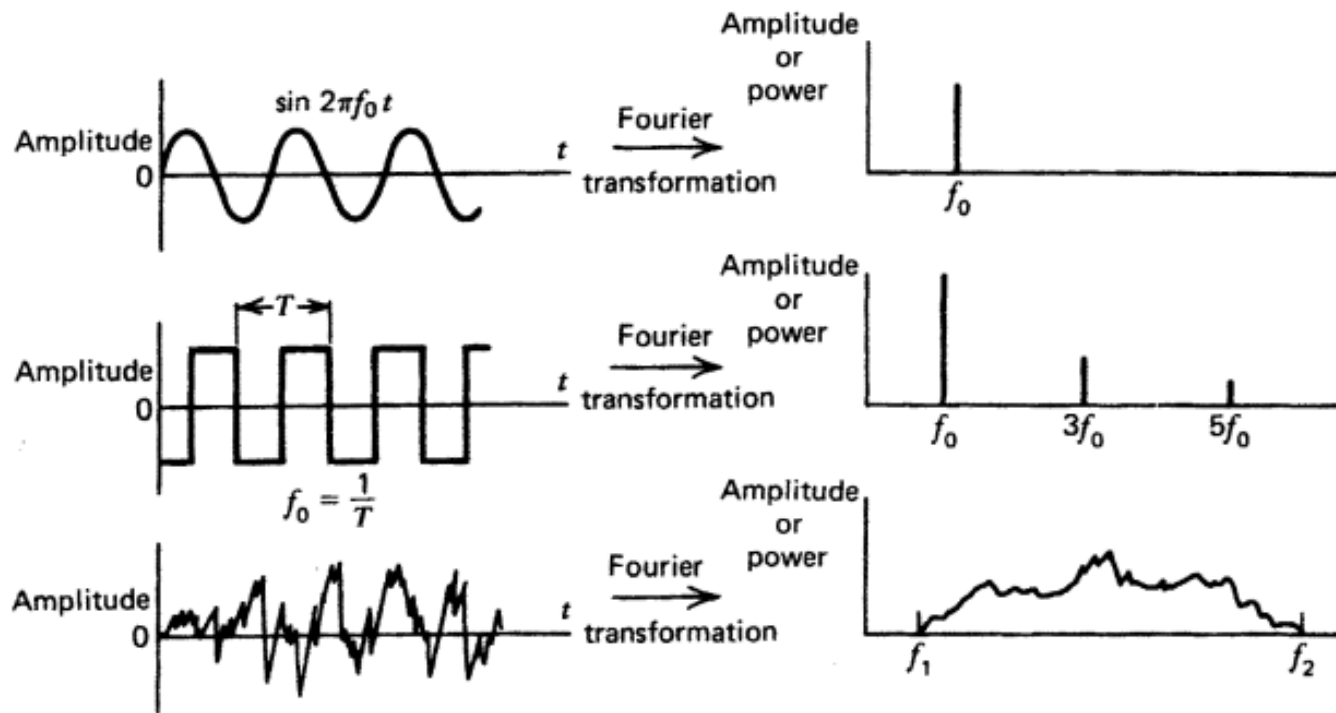


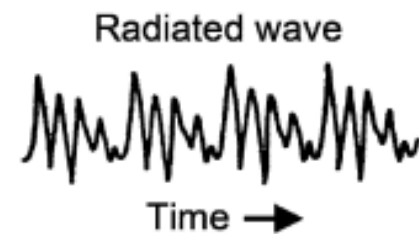
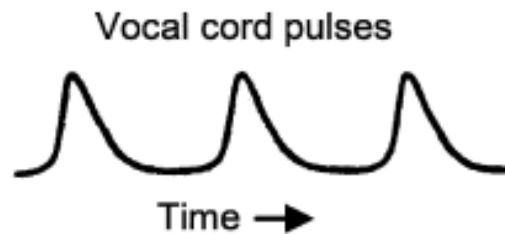
Fig. 10.17. Three waveforms on the left with their variations with time (the time domain), a sine (or cosine) wave, a square wave, and a more random pattern, with their respective frequency components (as Fourier analyzed and in the frequency domain). (From [507]. Reprinted with permission of Wiley)

$$I(t) = \sum_{n=1 \text{ to } \infty} I_n \sin(2\pi n f_{\text{fund}} t) \quad (10.42)$$

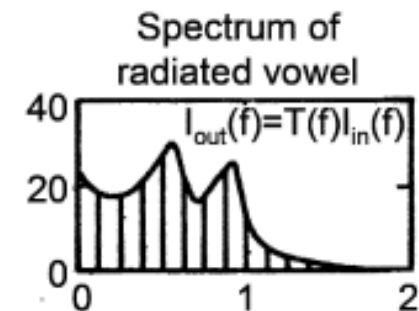
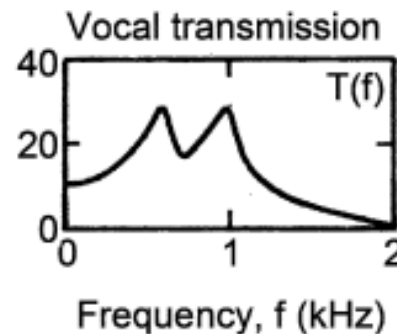
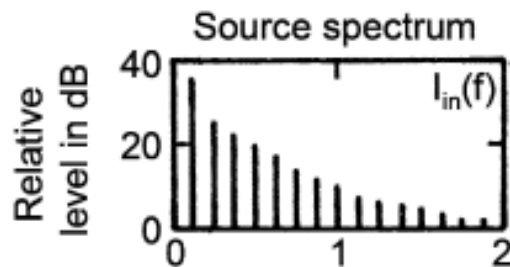
$$= I_1 \sin(2\pi f_{\text{fund}} t) + I_2 \sin(4\pi f_{\text{fund}} t) + I_3 \sin(6\pi f_{\text{fund}} t) + \dots \quad (10.43)$$

4 - Articulação (Voice-filtering theory) – O tracto vocal como filtro

Time domain



Frequency domain

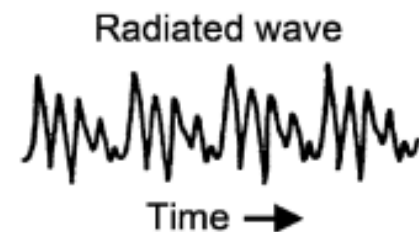
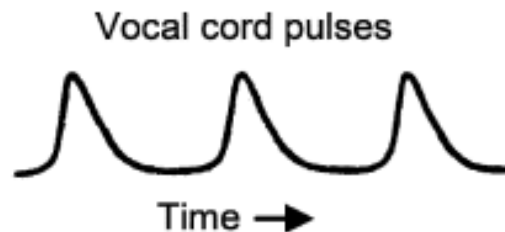


$$I(t) = \sum_{n=1 \text{ to } \infty} I_n \sin(2\pi n f_{\text{fund}} t) \quad (10.42)$$

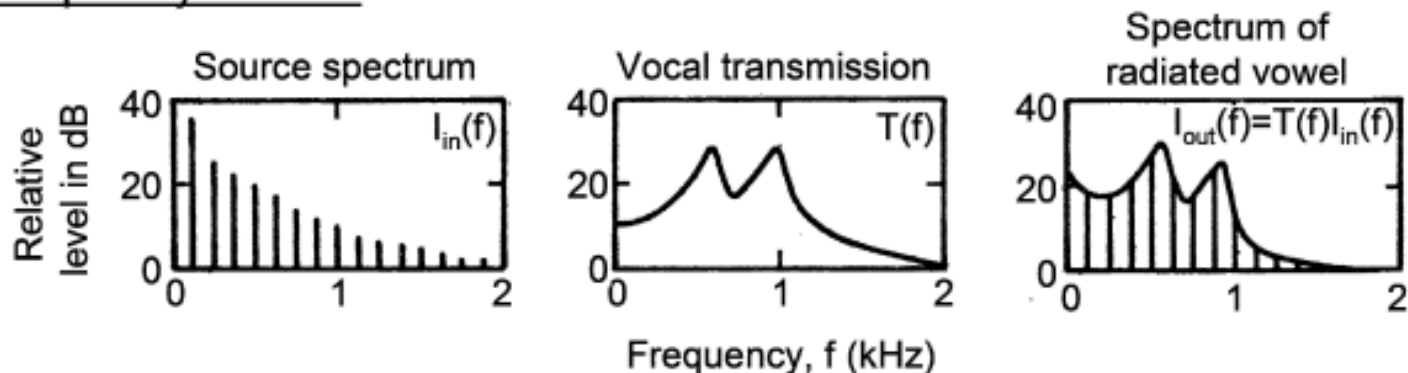
$$= I_1 \sin(2\pi f_{\text{fund}} t) + I_2 \sin(4\pi f_{\text{fund}} t) + I_3 \sin(6\pi f_{\text{fund}} t) + \dots \quad (10.43)$$

4 - Articulação (Voice-filtering theory) – O tracto vocal como filtro

Time domain



Frequency domain

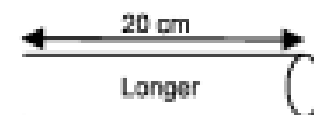


$$I_{\text{out}}(f) = T(f)I_{\text{in}}(f). \quad (10.44)$$

(d)



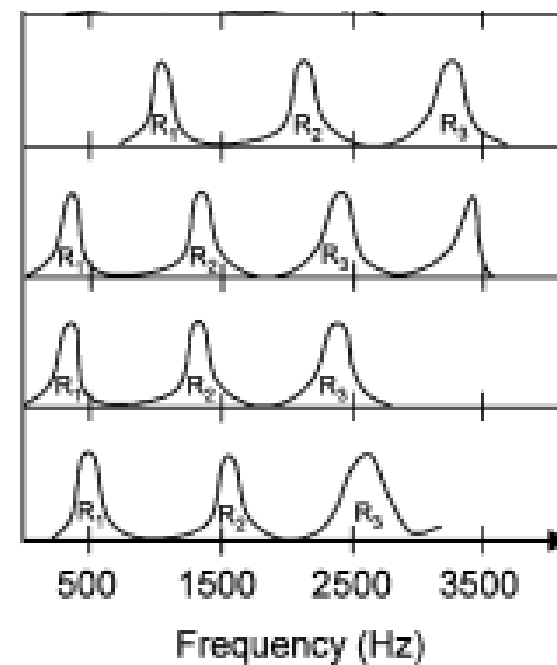
(e)

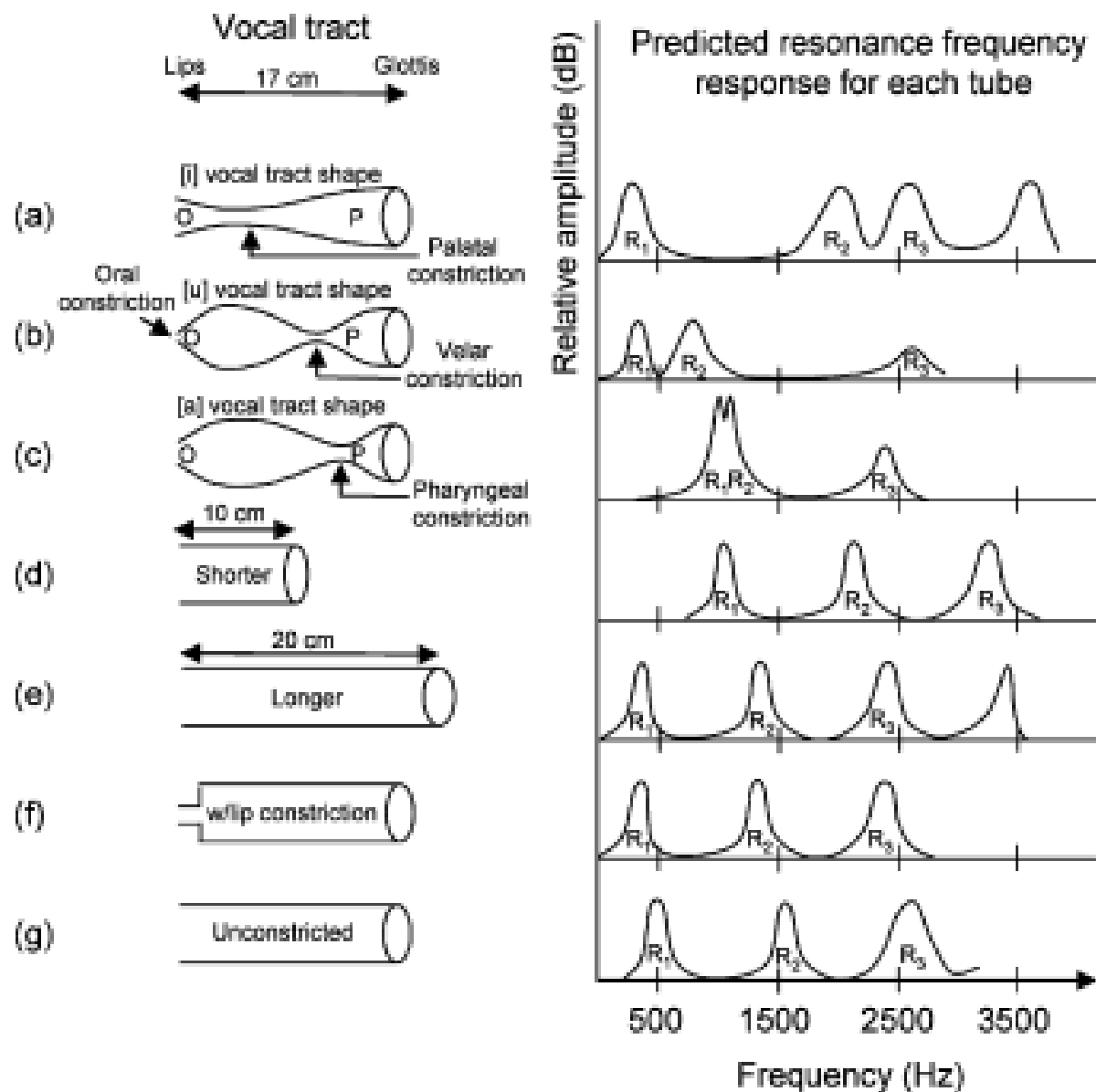


(f)



(g)



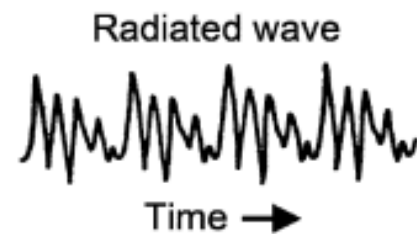
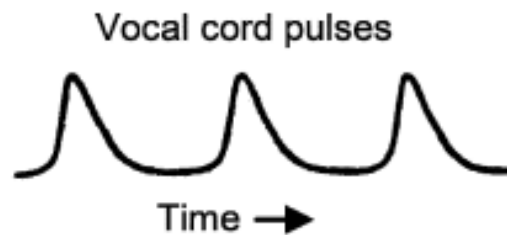


/i/

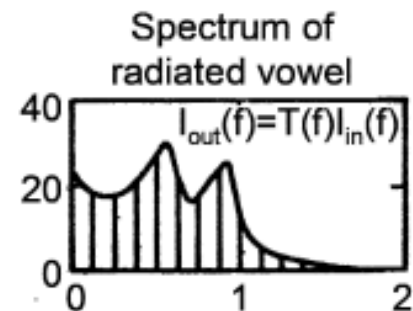
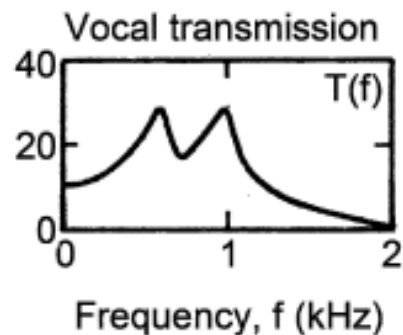
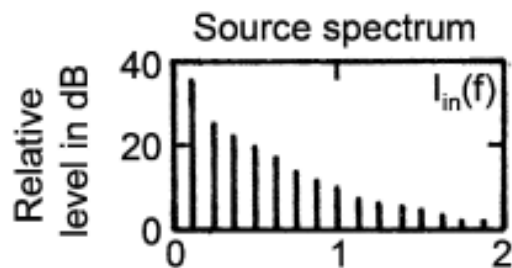
/u/

/a/

Time domain



Frequency domain



É essencial distinguir os papéis da frequência de oscilação das cordas vocais e as ressonâncias formantes do tracto vocal.

Dois tractos vocais / a mesma vogal

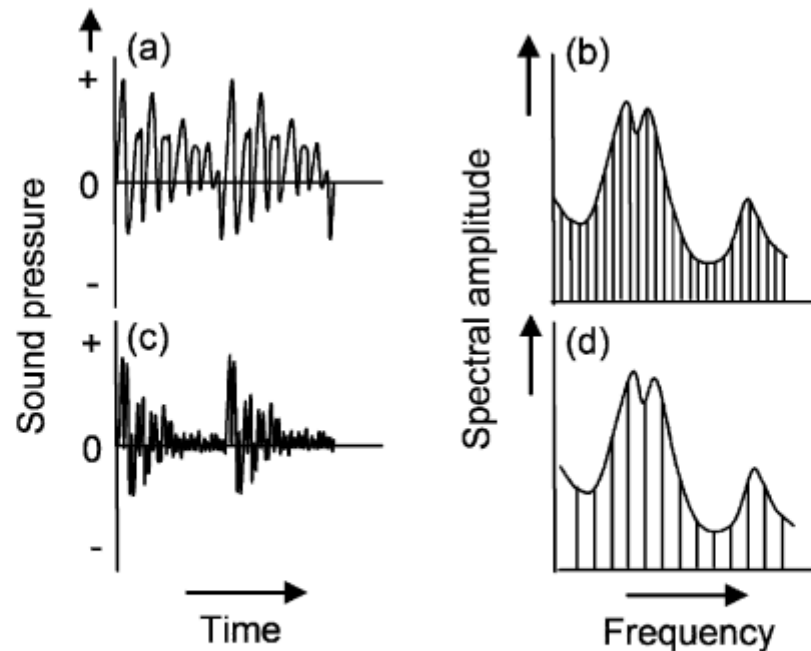


Fig. 10.21. Wave shapes (a,c) and corresponding spectra (b,d) for the same vowel “aw” as in “bought” and “awe,” but with two different vocal fold frequencies of (a,b) 90 Hz and (c,d) 150 Hz. (Based on [461])

Duas vogais / um tracto vocal

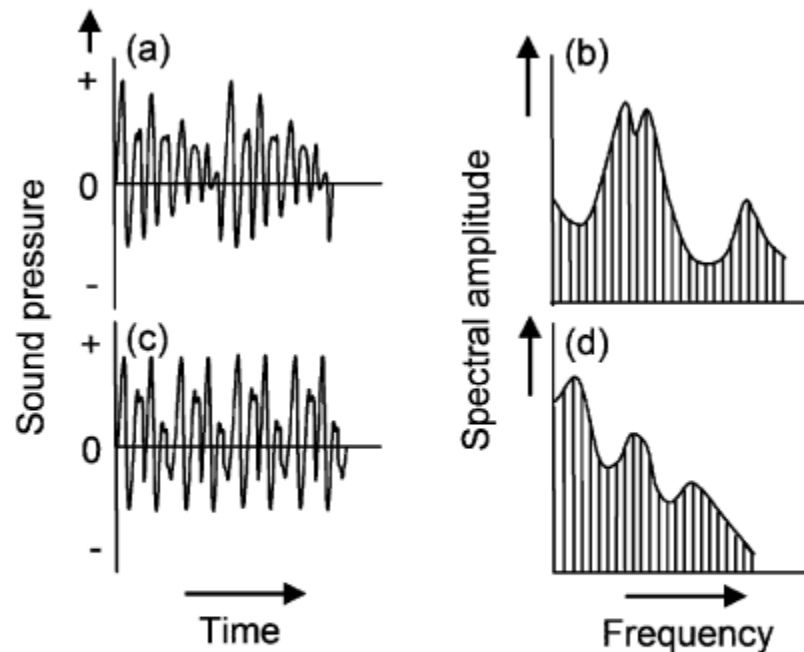


Fig. 10.22. Wave shapes (a,c) and corresponding spectra (b,d) for different vowels (a,b) “aw” as in “bought” and “awe” and (c,d) “uh” as in “but” and “about,” but with the same vocal fold fundamental frequency of 90 Hz. (Based on [461])

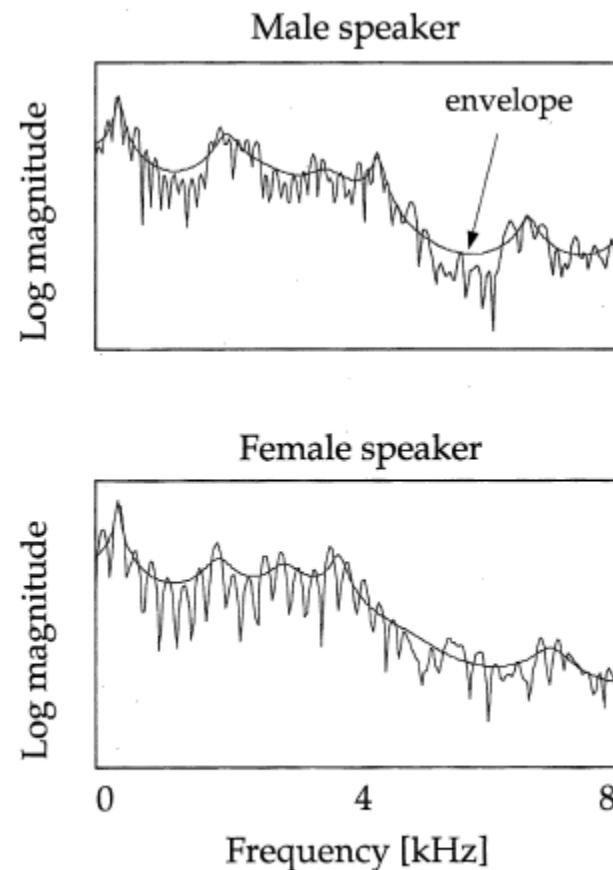
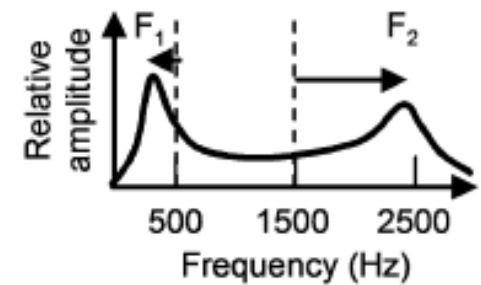


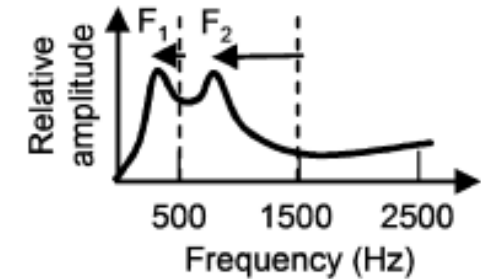
Fig. 10.23. The short-time spectra of the same linguistic pattern by a male and female speaker, with the spectral envelope shown. (From [449])

Figure 10.23 shows the spectral pattern of a man and woman saying the same message. The oscillation pattern for the man is more rapid than that for the woman, because the pattern consists of the harmonics of the fundamental buzzing frequency, which is lower in the man than the woman. However, the spectral envelope is the same for both, and this envelope carries the most important information for understanding speech.

(a) "ee" vowel



(b) "oo" vowel



(c) "ah" vowel

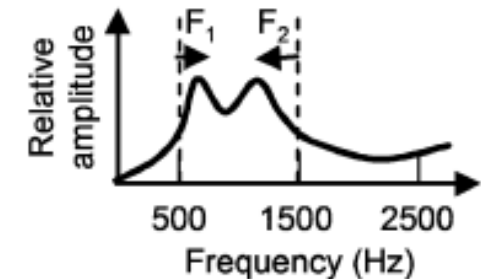


Figure 10.24 shows that the vocal tract has very different shapes for the three depicted vowels, and consequently these vowels have very different formant frequencies.

It also shows that this one-tube model gives $F_1 = 500$ Hz and $F_2 = 1500$ Hz for these vowels.

Calling $f_r = v_s/4L$, these correspond to f_r and $3f_r$.

However, the real formant frequencies are clearly shifted from these values.

- <http://www.singlikeastar.com/how-your-voice-works/>
- <http://www.fonologia.org/index.php>