

Módulo 09

Física da Fala e da Audição

A Voz Humana I

Produção da Fala

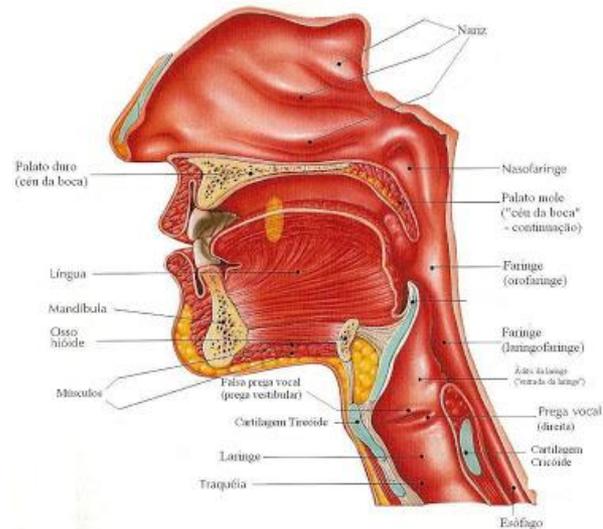
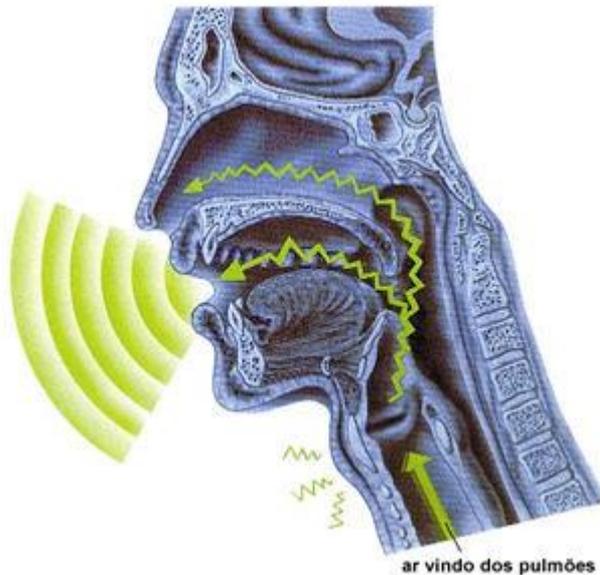
Prof. Edmilson Manganote

Instituto de Física Gleb Wataghin (IFGW)

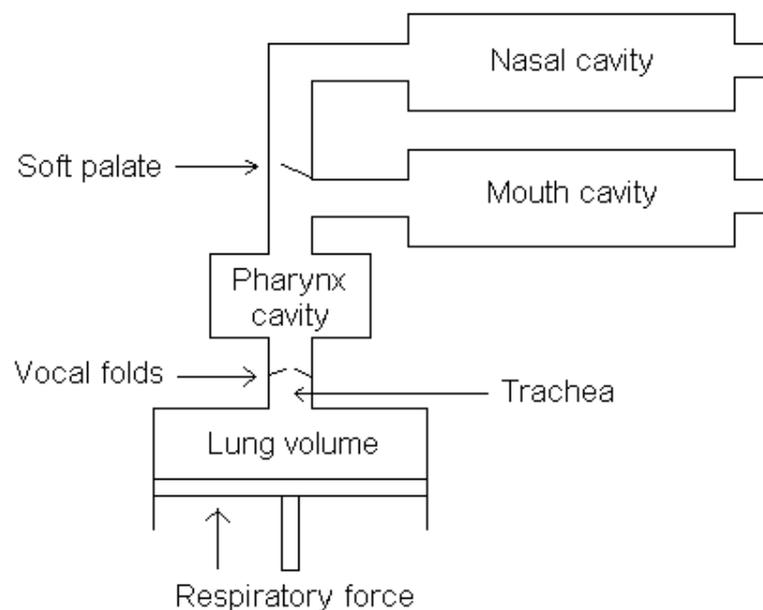
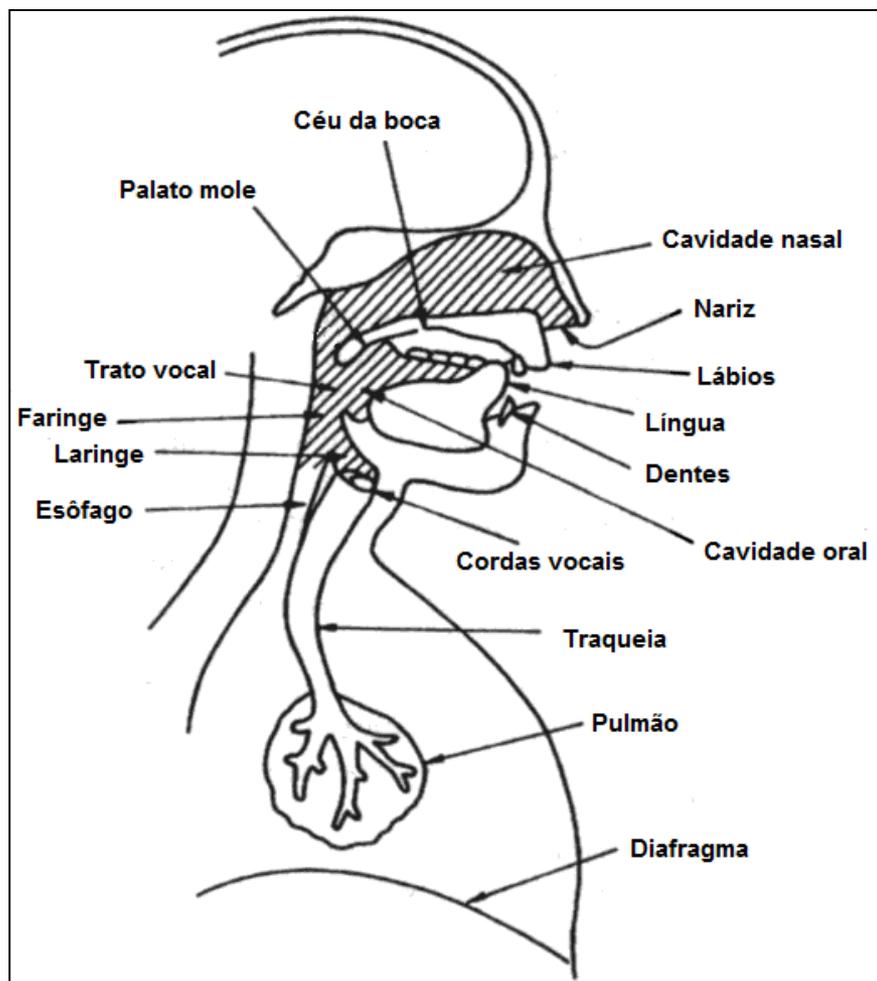
Universidade Estadual de Campinas (UNICAMP)

mangano@ifi.unicamp.br

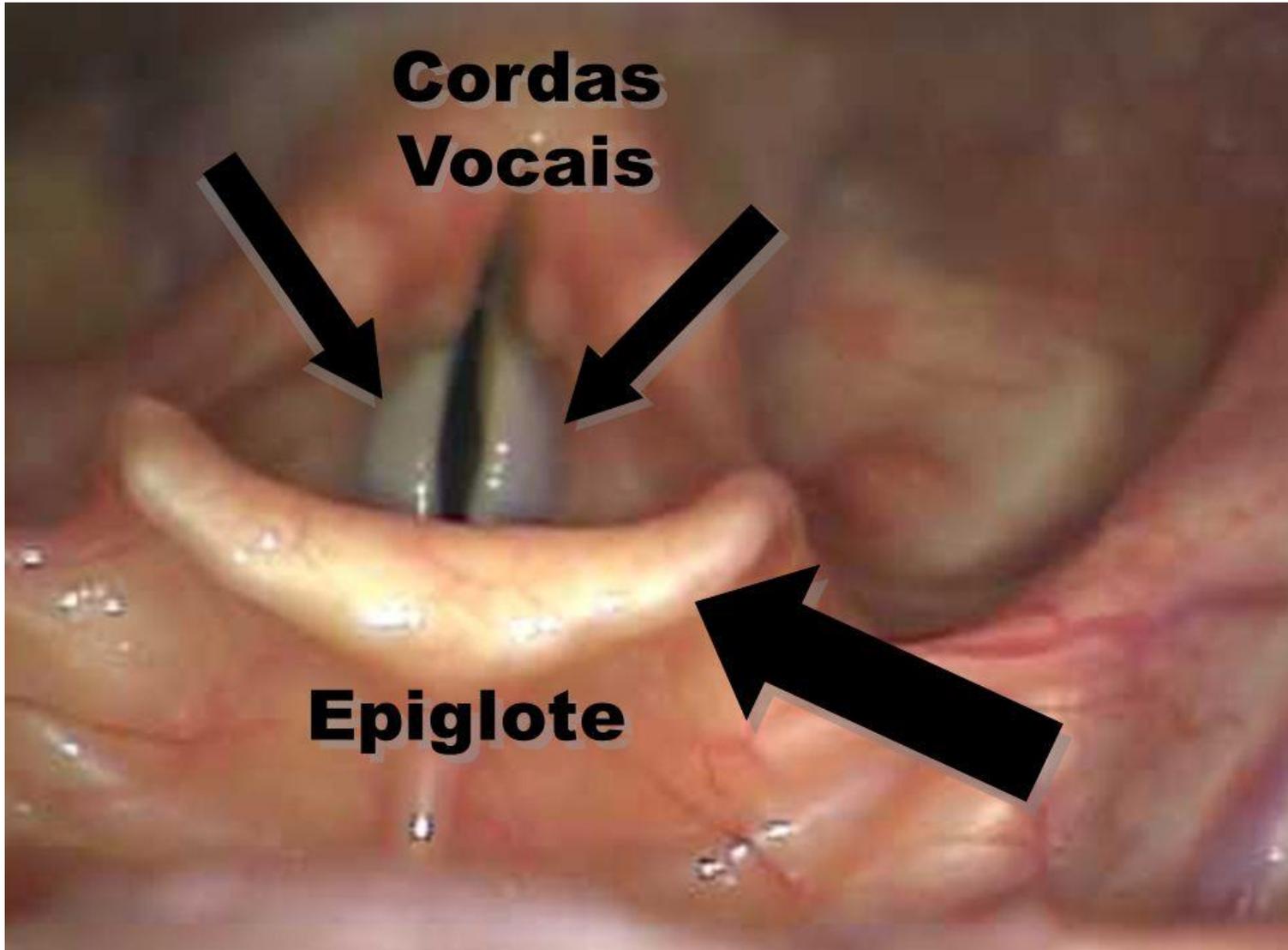
O ouvido tem sua sensibilidade máxima no intervalo de frequências de 1000 a 4000Hz. É neste intervalo que as ressonâncias (chamadas formantes) do trato vocal ocorrem.



Os Órgãos Vocais



Para produzirmos os sons da fala, o fluxo de ar precisa ser interrompido pelas cordas (pregas) vocais ou através de constrições no trato vocal (feitas usando-se a língua ou os lábios, por exemplo).



Os Órgãos Vocais - Como? Um Modelo...

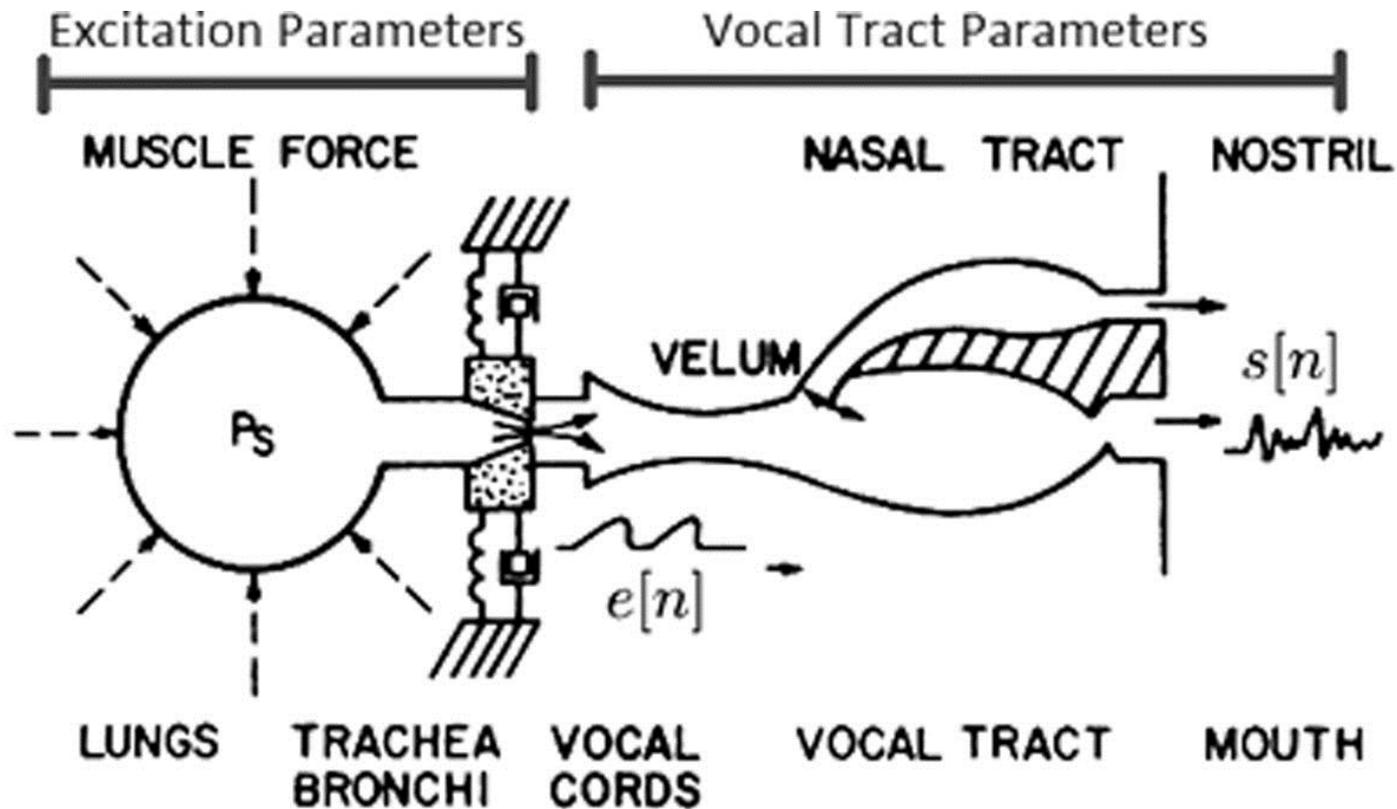


Figure 1: Speech Production System [J. L. Flanagan *et al.* "Synthetic Voices for Computers," IEEE Spectrum, vol. 7, pp. 22-45, Oct 1970]

Os Órgãos Vocais - Como? Um Modelo...

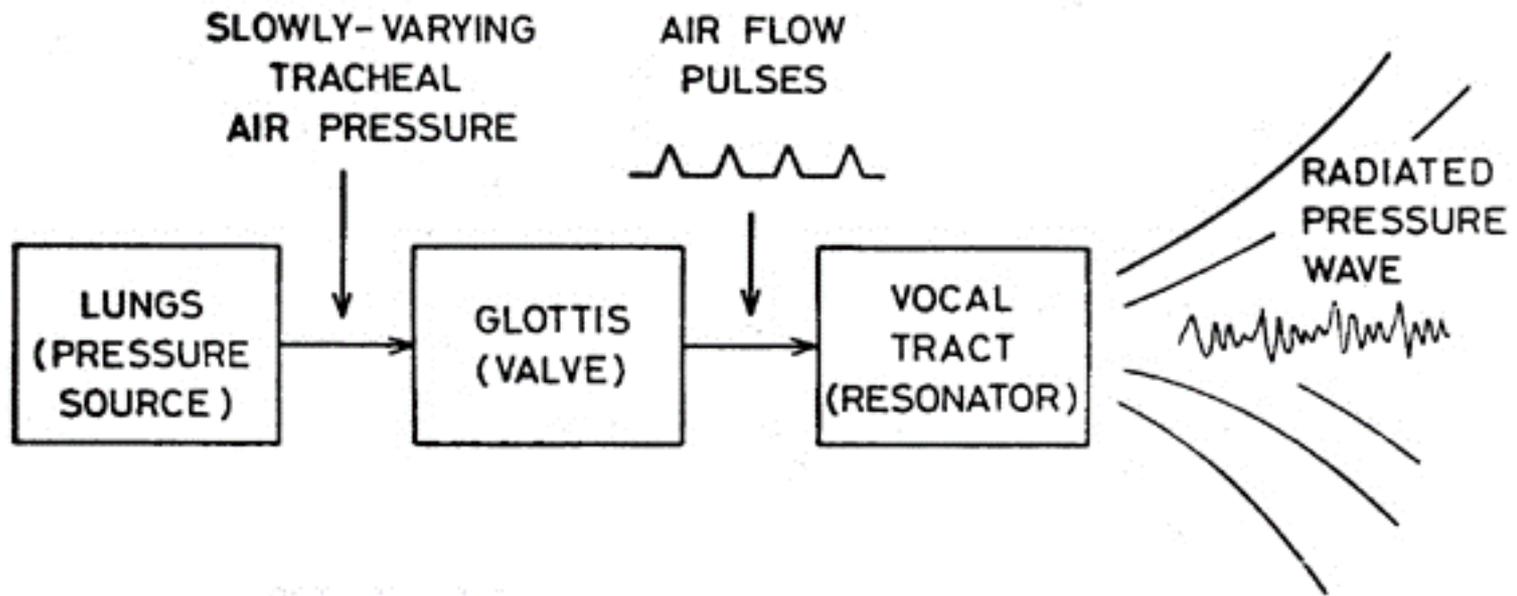


Figure 1. Non-interactive source-tract model for voice production.

Os Órgãos Vocais - Como? Um Modelo...

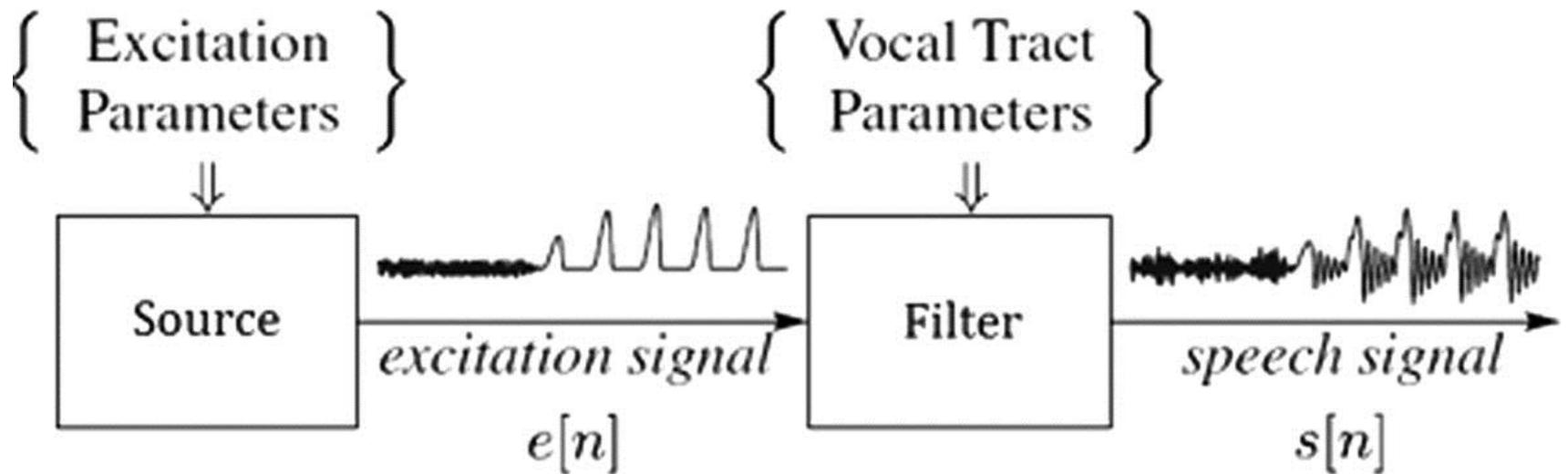
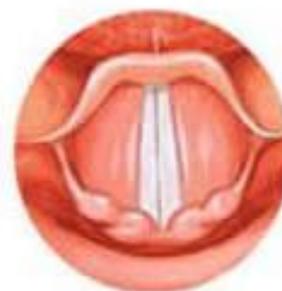
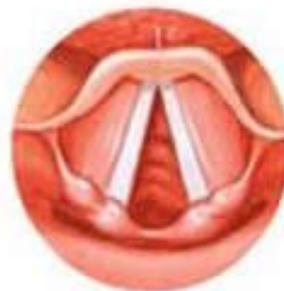
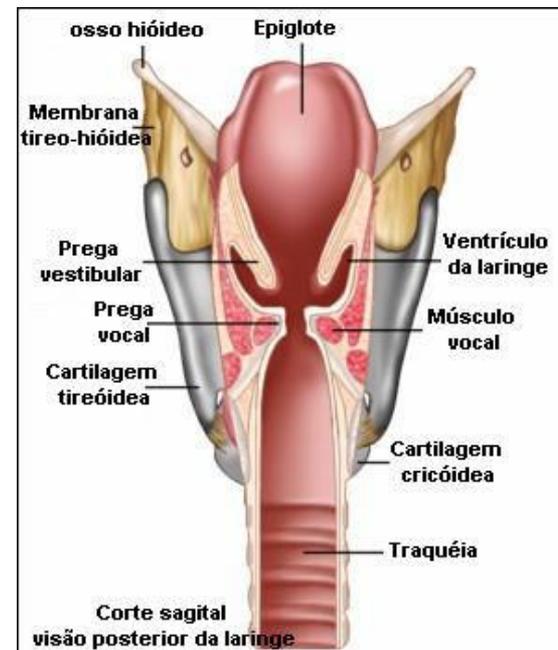
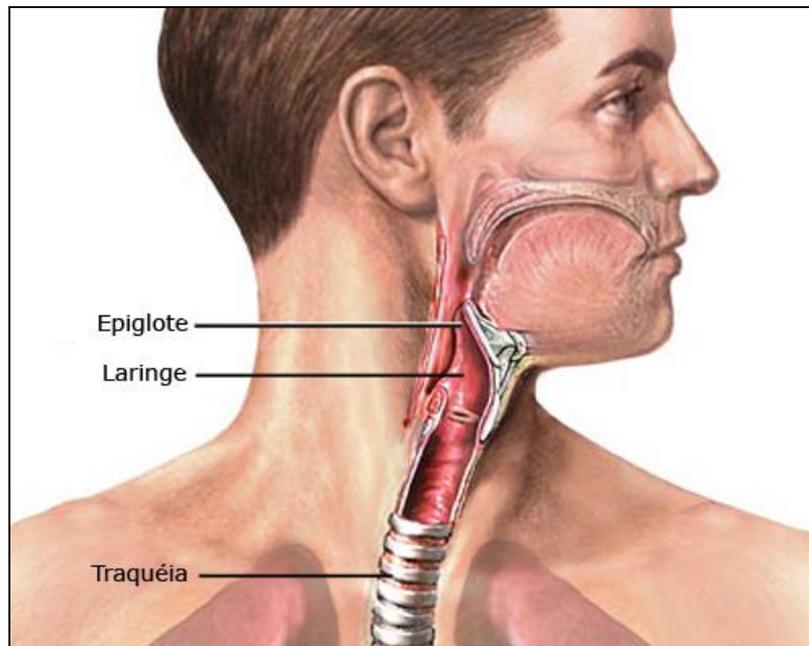


Figure 2: Source-Filter Model of Speech Production

A Laringe e as Cordas ou Pregas Vocais



A Laringe e as Cordas ou Pregas Vocais



As cordas vocais modulam o fluxo de ar, abrindo e fechando rapidamente a passagem pela laringe.

Sua taxa de vibração é determinada principalmente por sua massa e tensão, embora a pressão e velocidade do ar também contribuam, em menor escala.

Durante uma fala normal, a taxa de vibração das pregas vocais pode variar a uma razão de 2:1 (uma oitava).

Frequências típicas

Homem: 110 Hz; Mulher: 220 Hz; Crianças: 300 Hz.

A Laringe e as Cordas ou Pregas Vocais

Modos de Vibração:

- 1) Modo Normal, as pregas vocais se abrem e se fecham completamente durante o ciclo, gerando pulsos (sopros) de ar, que quando observados ao longo do tempo apresentam uma forma razoavelmente triangular.
- 2) Modo de Fase Aberta, as cordas não se fecham totalmente ao longo de seu comprimento; ou seja, o fluxo de ar não vai a zero (Sussurro...).
- 3) Um terceiro modo aparece quando um mínimo de ar passa pelas pregas vocais em sopros bem curtos, gerando uma voz "rangente".
- 4) Existe ainda um quarto modo de vibração, mas que não é utilizado na fala usual/cotidiana, é o *falsetto*.

A Laringe e as Cordas ou Pregas Vocais

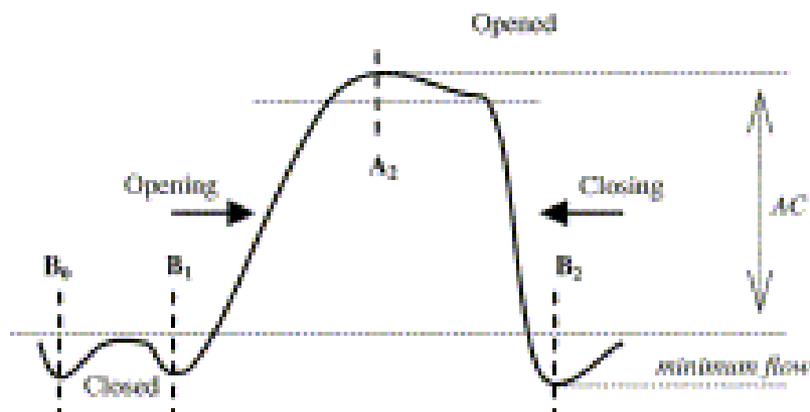
As cordas vocais se abrem por causa da pressão de ar na traquéia.

A amplitude de vibração das cordas vocais relaciona-se com a Loudness. Sons mais intensos necessitam de uma pressão extra de ar dos pulmões - a pressão dos pulmões varia de 400 a 2000 N/m².

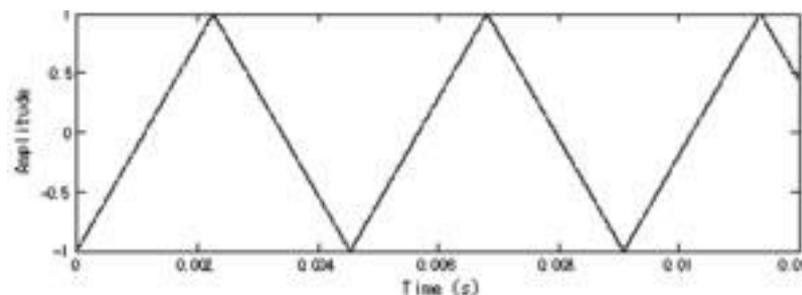
O fluxo de ar pela glote é razoavelmente proporcional à abertura da mesma. Para um esforço vocal normal, a forma de onda do fluxo de ar é aproximadamente triangular, repetindo-se a cada 8ms - o som resultante é como um zumbido, cheio de sobretons.

OBS: As cordas vocais não são as únicas fontes de sons. Sons como "f" e "s" são produzidos pelo fluxo turbulento de ar através de uma constricção em algum lugar do trato vocal. Consoantes como "p" e "t" são conseguidas através da liberação repentina de pressão em algum ponto.

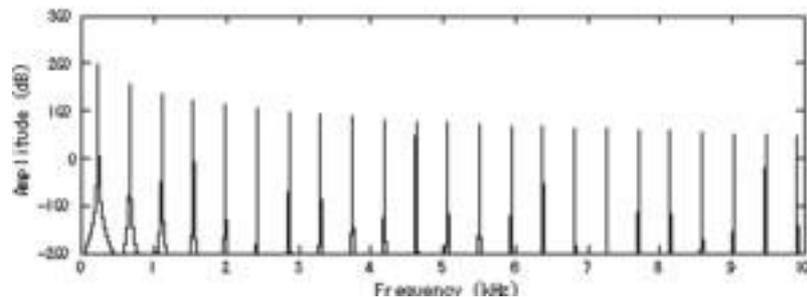
A Laringe e as Cordas ou Pregas Vocais



Visão esquemática do fluxo de ar através da glote.



Uma onda triangular no domínio do tempo (topo) e o gráfico de espectro correspondente (embaixo).

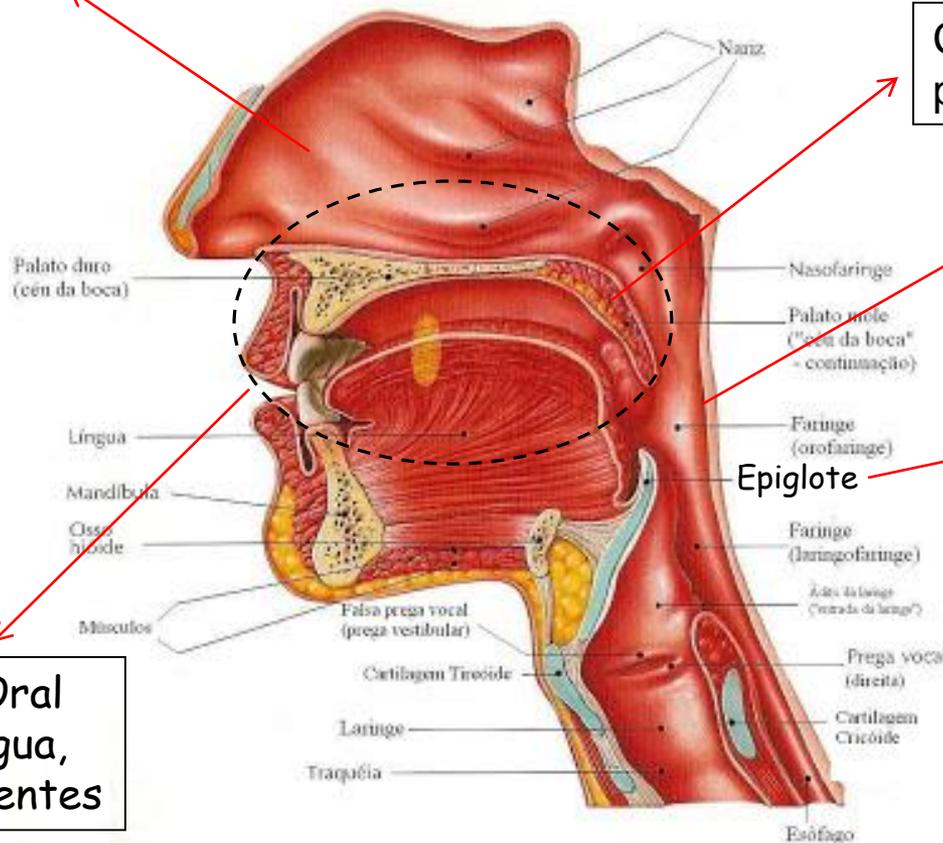


A frequência fundamental é de 220 Hz (Lá₂). Cada linha vertical indica a amplitude de uma das frequências componentes da onda.

O Trato Vocal

O **trato vocal** transforma os zumbidos das pregas vocais nos sons da fala. Isto é realizado através de mudanças na sua forma, o que produz variações nas ressonâncias acústicas.

Homem adulto
17 cm de comprimento
60 cm³ de volume



Controla o fluxo de ar da faringe para a cavidade nasal

Conecta a laringe com a cavidade oral

Funciona com uma válvula, evitando a passagem de alimentos para a laringe

Cavidade Oral
Palato, Língua,
Lábios e Dentes

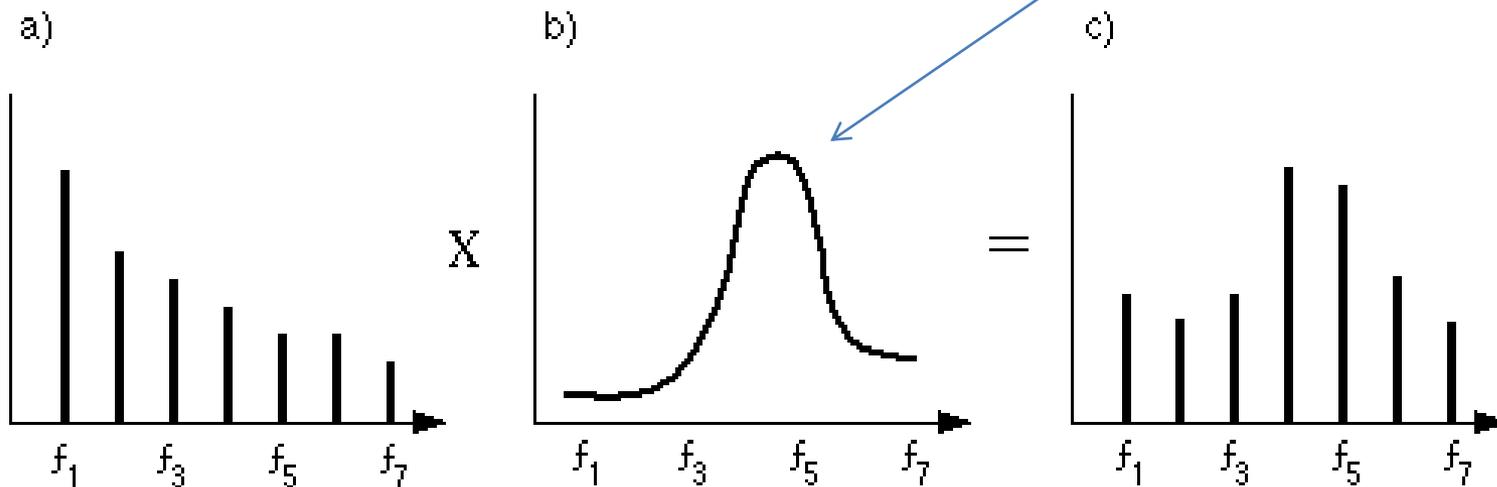
Som da Fala = Fonte x Função Filtro x Eficiência de Radiação

Ressonâncias do Trato Vocal - Formantes

Os formantes podem ser definidos como picos de energia em uma região do espectro sonoro. As parciais que se encontram nessa região de ressonância serão realçados.

Enquanto o espectro de cada nota de um instrumento (fonte) pode variar consideravelmente com a altura, as regiões dos formantes permanecem estáveis, seja qual for a frequência da nota.

Portanto, os formantes funcionam como uma espécie de assinatura de uma determinada fonte sonora.



Ressonâncias do Trato Vocal - Formantes

TABLE 15.3 Formant frequencies and amplitude of vowels averaged for 76 speakers

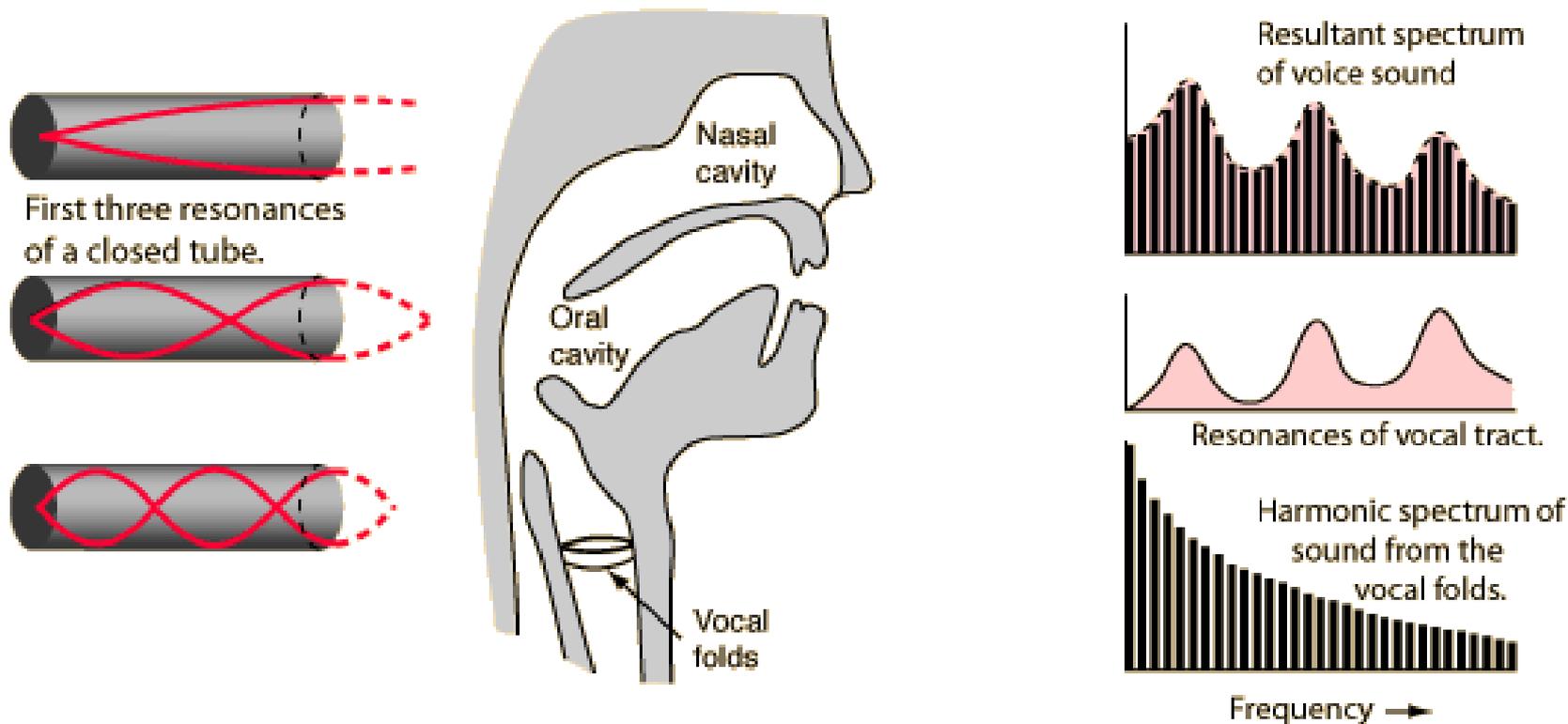
		/i/	/I/	/ɛ/	/æ/	/ɑ/	/ɔ/	/ʊ/	/u/	/ʌ/
Formant frequencies (Hz)		(ee)	(i)	(e)	(ae)	(ah)	(aw)	(û)	(oo)	(u)
F_1	M	270	390	530	660	730	570	440	300	640
	W	310	430	610	860	850	590	470	370	760
	Ch	370	530	690	1010	1030	680	560	430	850
F_2	M	2290	1990	1840	1720	1090	840	1020	870	1190
	W	2790	2480	2330	2050	1220	920	1160	950	1400
	Ch	3200	2730	2610	2320	1370	1060	1410	1170	1590
F_3	M	3010	2550	2480	2410	2440	2410	2240	2240	2390
	W	3310	3070	2990	2850	2810	2710	2680	2670	2780
	Ch	3730	3600	3570	3320	3170	3180	3310	3260	3360
Formant amplitudes (dB)		-4	-3	-2	-1	-1	0	-1	-3	-1
		-24	-23	-17	-12	-5	-7	-12	-19	-10
		-28	-27	-24	-22	-28	-34	-34	-43	-27

Peterson, G.E. and H.L. Barney (1952). "Control Methods Used in a Study of Vowels". *J. Acoust. Soc. Am.* **24**:175.

Articulação da Fala

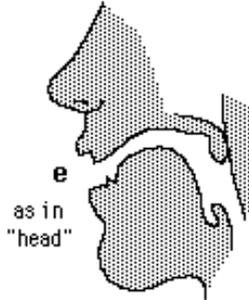
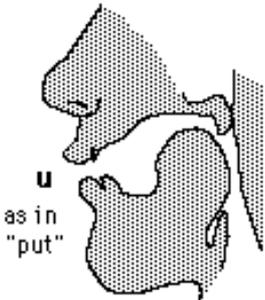
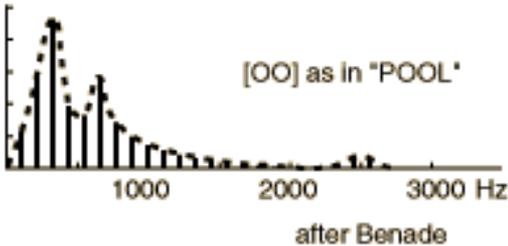
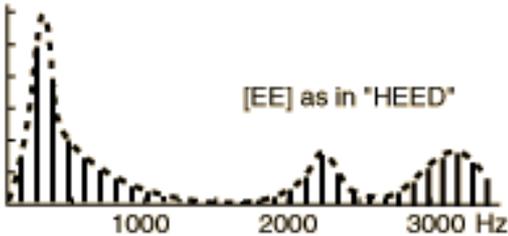
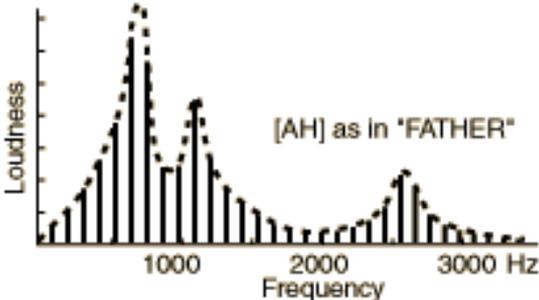
Para produzir sons distinguíveis, como as vogais, o mecanismo vocal deve controlar de algum modo as ressonâncias no trato vocal que produzem as formantes características.

Considerando o trato vocal como uma cavidade ressonante, a posição da língua, a área de abertura da boca e qualquer mudança que afeta o volume na cavidade vão fazer com que a frequência de ressonância seja alterada.



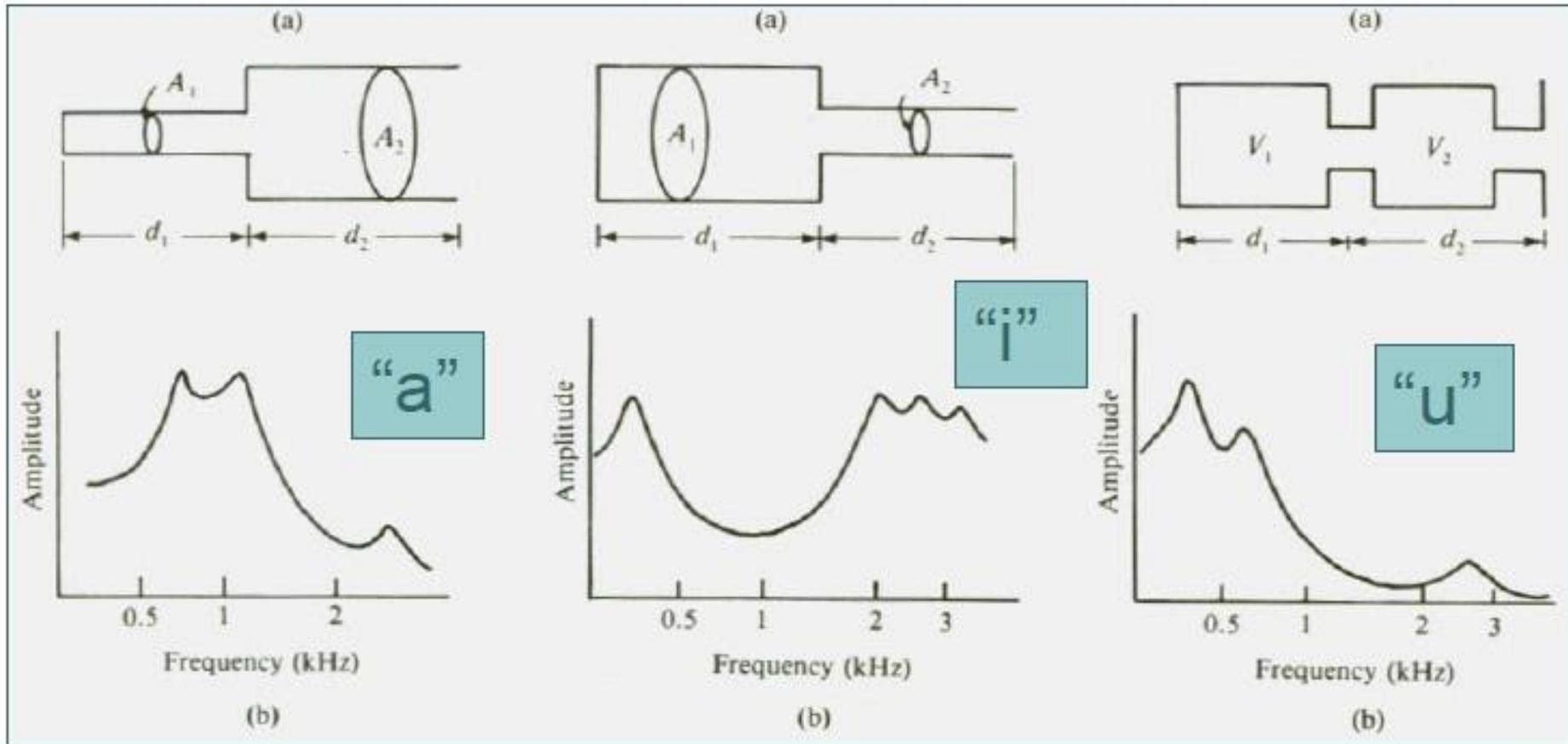
Articulação da Fala

Exemplos das mudanças na forma do mecanismo vocal no processo de articulação, formando diferentes vogais.



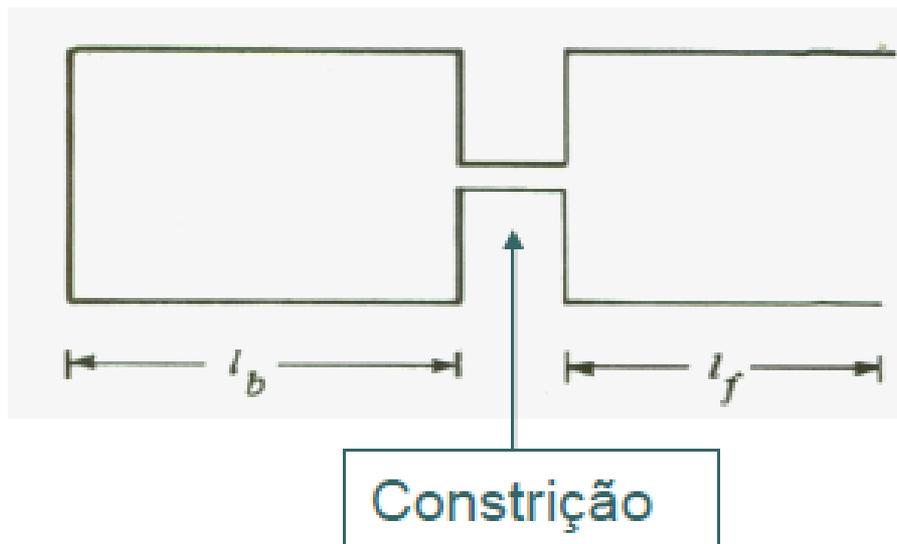
Modelos para o Trato Vocal

Aproximação, para o trato vocal, em dois tubos, para as vogais altas.



Modelos para o Trato Vocal

Modelagem da configuração do trato vocal para produção de consoantes.



A "constricção" (redução brusca na área de ligação entre os volumes de trás e frontal) é ajustada em diferentes posições modificando a relação entre os comprimentos l_b da cavidade de trás e l_f da cavidade frontal). O comprimento da constricção l_c é ~ 3 cm. Tipicamente $l_b + l_f + l_c = 17$ cm.

Stevens (1972). "The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data" in *Human Communication: A unified View*. Eds. E.E. David and P.B. Denes. New York: McGraw-Hill.

A mecânica das Pregas Vocais

A altura é aumentada ao se aumentar a tensão nas pregas vocais, o que pode ocorrer tanto externamente como internamente.

$$f_1 = \frac{1}{2L} \sqrt{\frac{\sigma}{\rho}}$$

Exemplo:

$$\sigma \approx 12kPa$$

$$\rho \approx 1040kg/m^3$$

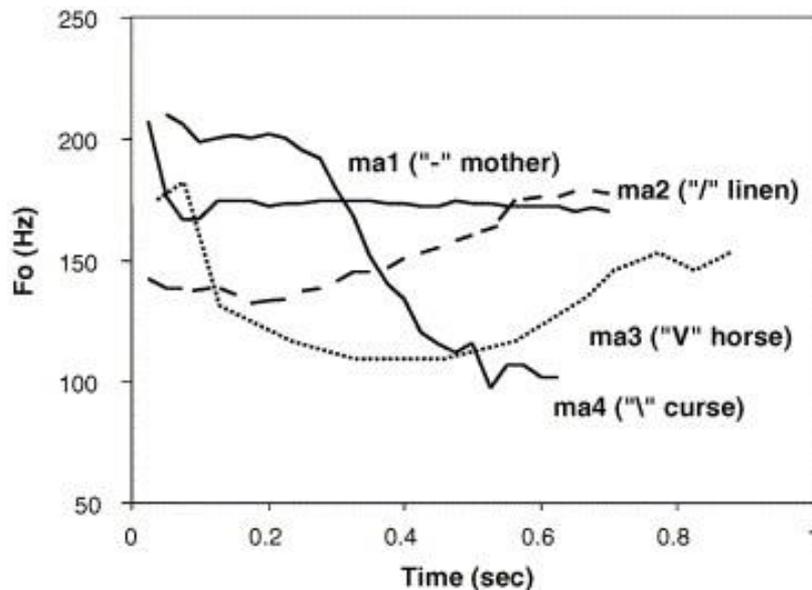
A natureza da tensão nas pregas vocais (interna ou externa) permite a fonação com diferentes registros.

Durante a adolescência as pregas vocais se alongam e, conseqüentemente, a voz fica mais grave (pitch diminui).

Características Prosódicas da Fala

Prosódia (originário do grego προσῳδία) é o estudo do ritmo, entonação e demais atributos correlatos na fala. Ela descreve todas as propriedades acústicas da fala que não podem ser preditas pela transcrição ortográfica (ou similar), em resumo, cuida da correta acentuação tônica das palavras.

Mudanças da frequência com o tempo par tons no Mandarin



As características prosódicas tendem, também, a indicar o estado emocional de quem fala (alegria, tristeza, excitação, etc.)

Wei, C.G; Cao, K. and Zeng, F.G. (2004) "Mandarin tone recognition in cochlear-implant subjects". *Hearing Research.* , **197**:1-2, pp. 87-95.