

# Módulo 10

## Física da Fala e da Audição

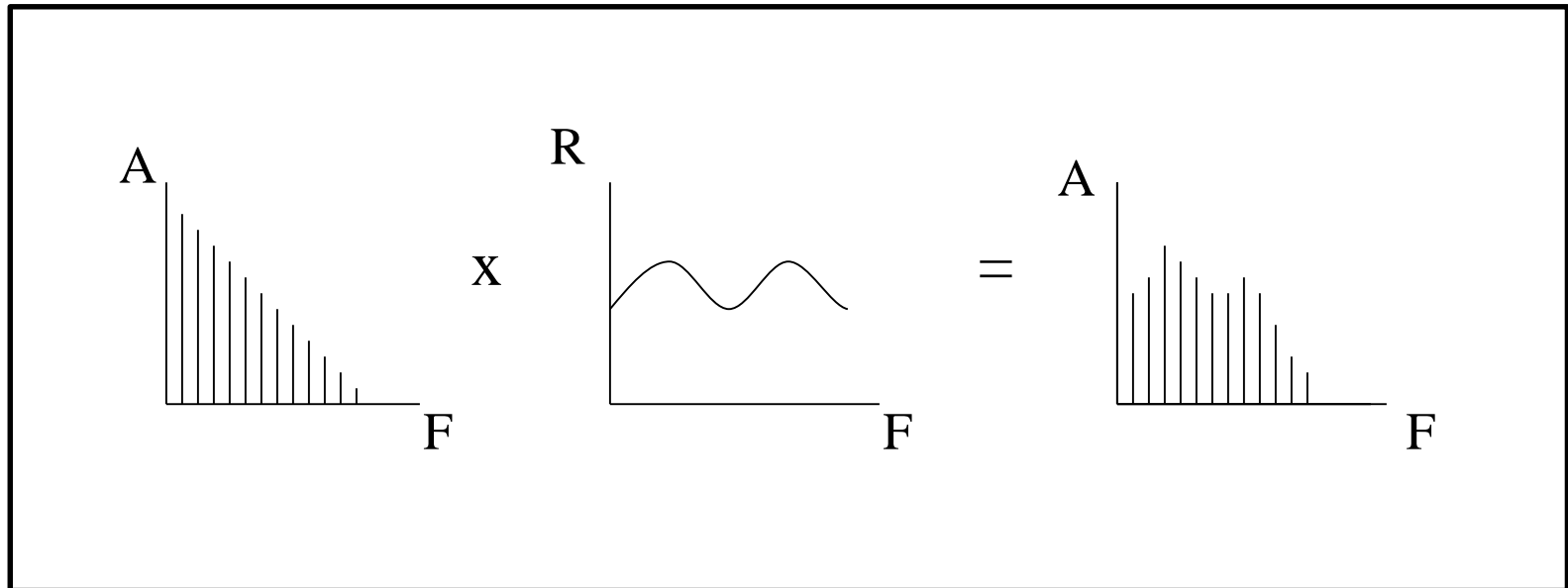
### A Voz Humana II

### Reconhecimento da Fala

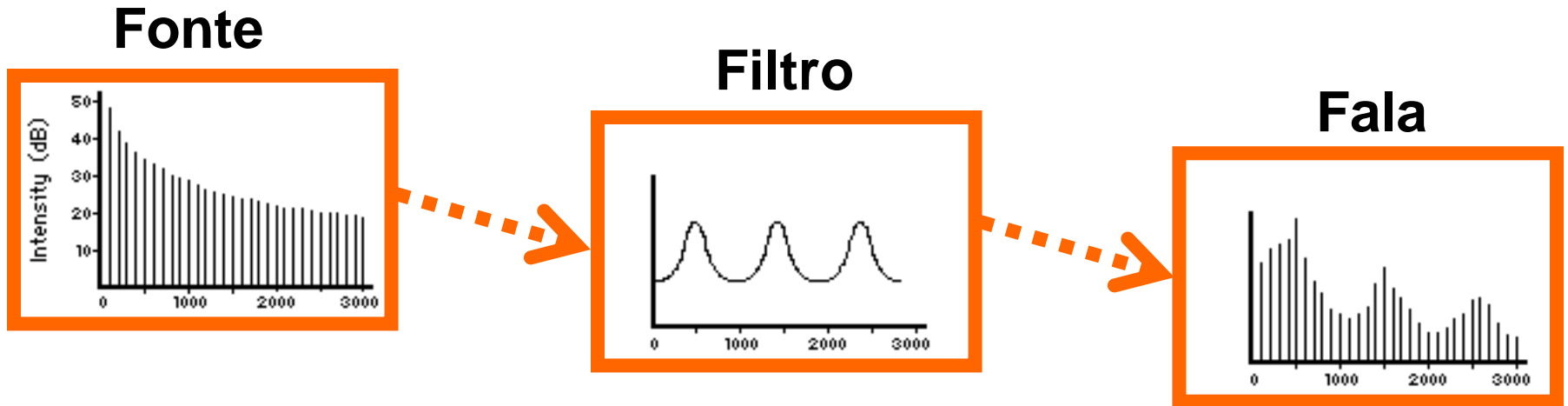
**Prof. Edmilson Manganote**  
Instituto de Física Gleb Wataghin (IFGW)  
Universidade Estadual de Campinas (UNICAMP)  
[mangano@ifi.unicamp.br](mailto:mangano@ifi.unicamp.br)

# Modelo Fonte-Filtro

O trato vocal dá forma ao espectro que é gerado na laringe. O espectro de saída tem características tanto da fonte quanto do filtro.



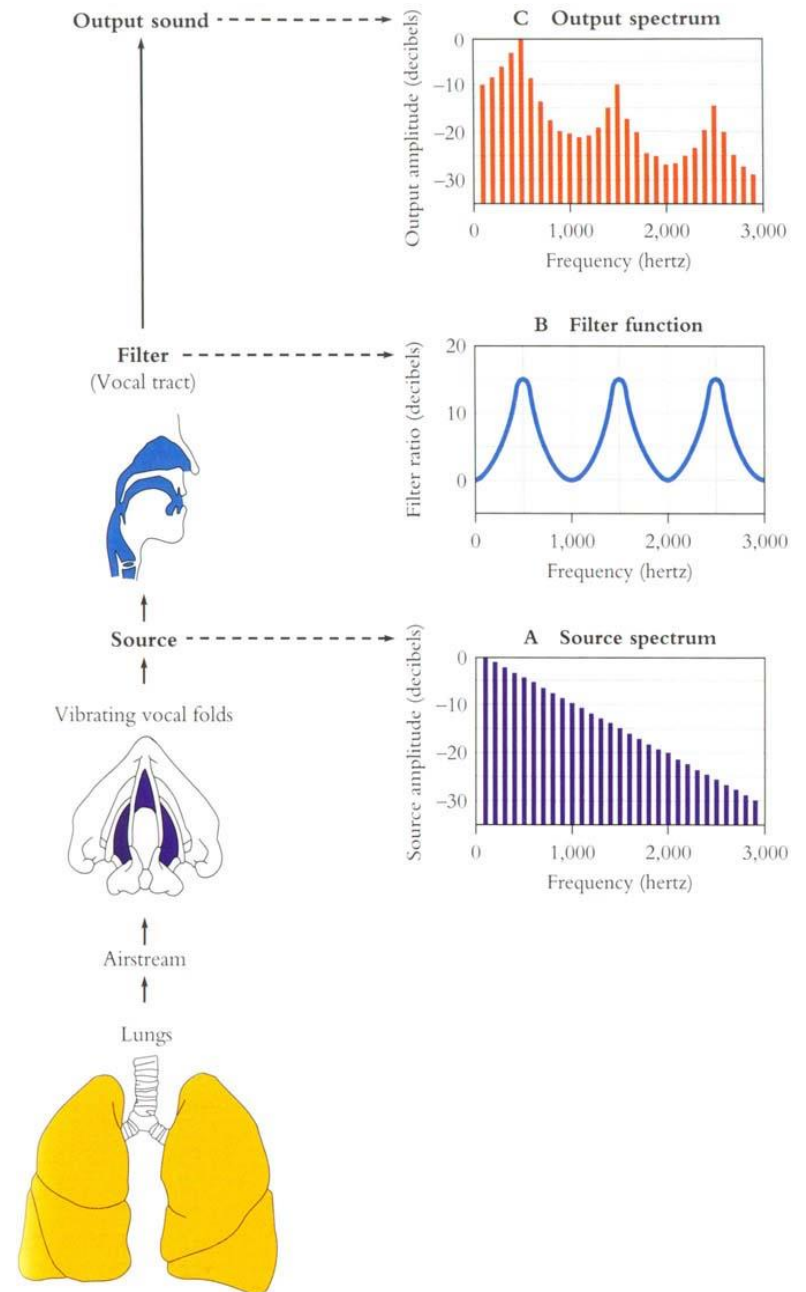
# Teoria fonte-filtro da produção da fala



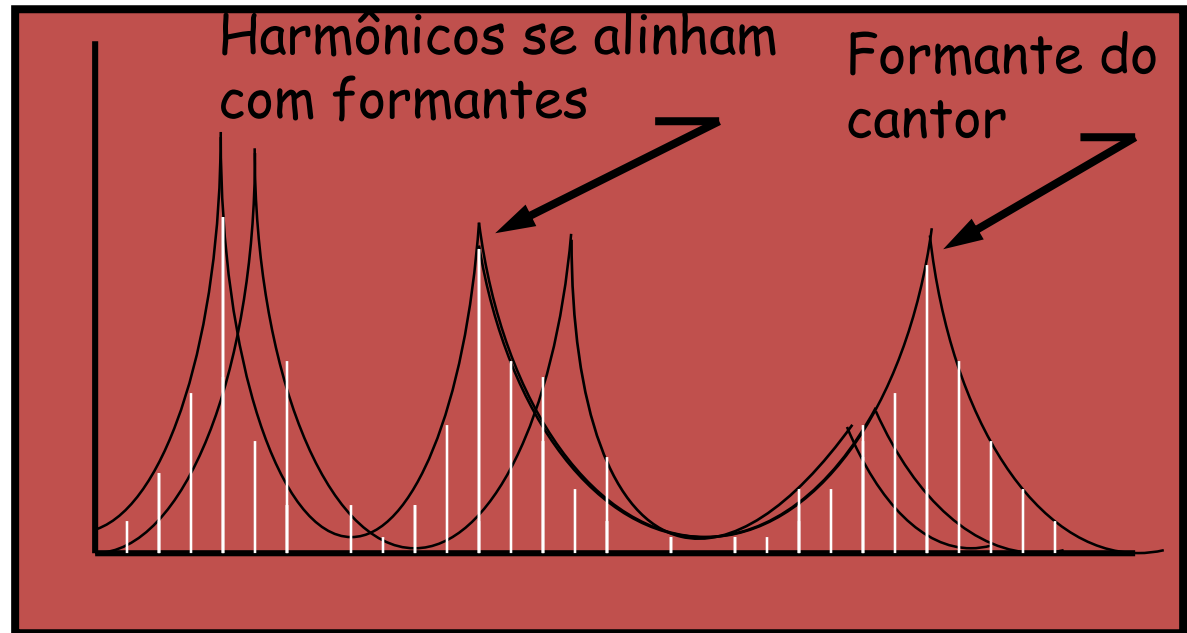
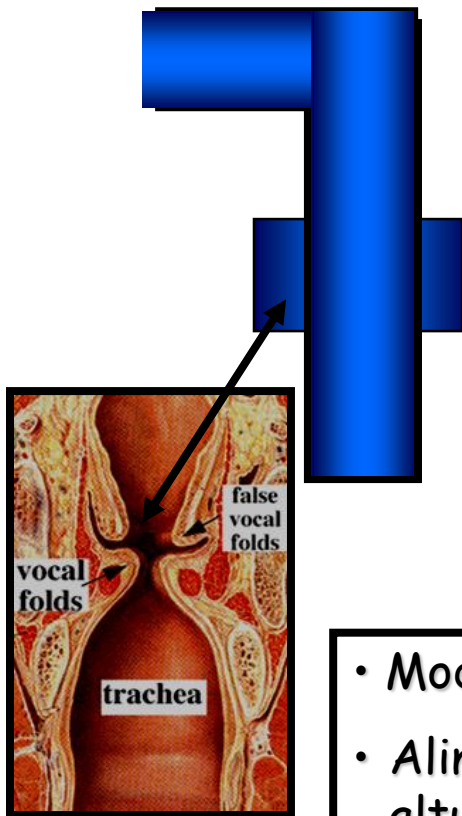
Produção da fala pode ser dividida em duas partes independentes:

- Fontes de som (sinais) - laringe
- Filtros que modificam o sinal (sistemas) - trato vocal

# Teoria fonte-filtro da produção da fala



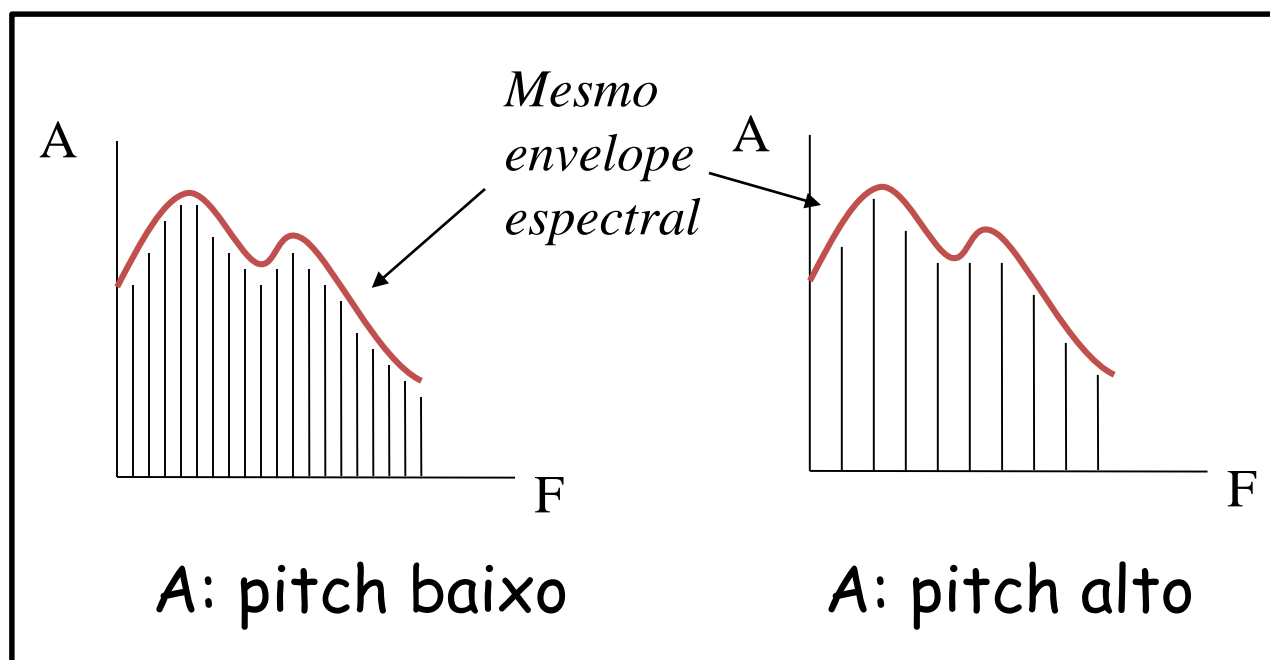
# Formantes e Canto



- Modificação da vogal desloca os formantes.
- Alinhamento dos formantes com harmônicos intensifica a altura.
- Dilatação do trato vocal causa o "Formante do Cantor" (esta técnica recebe nomes como impostação ou projeção)

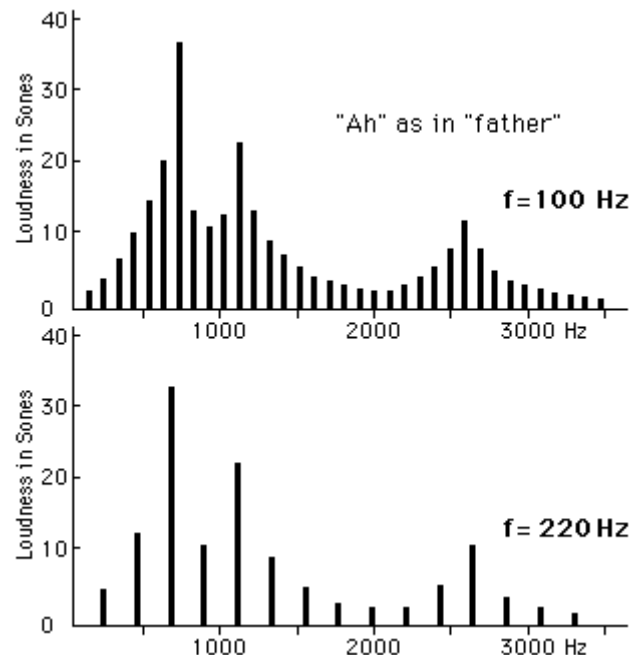
# Percepção de vogais

- Altura das vogais depende da frequência de repetição (=frequência fundamental).
- Essa é uma propriedade da **fonte**.



# Percepção de vogais

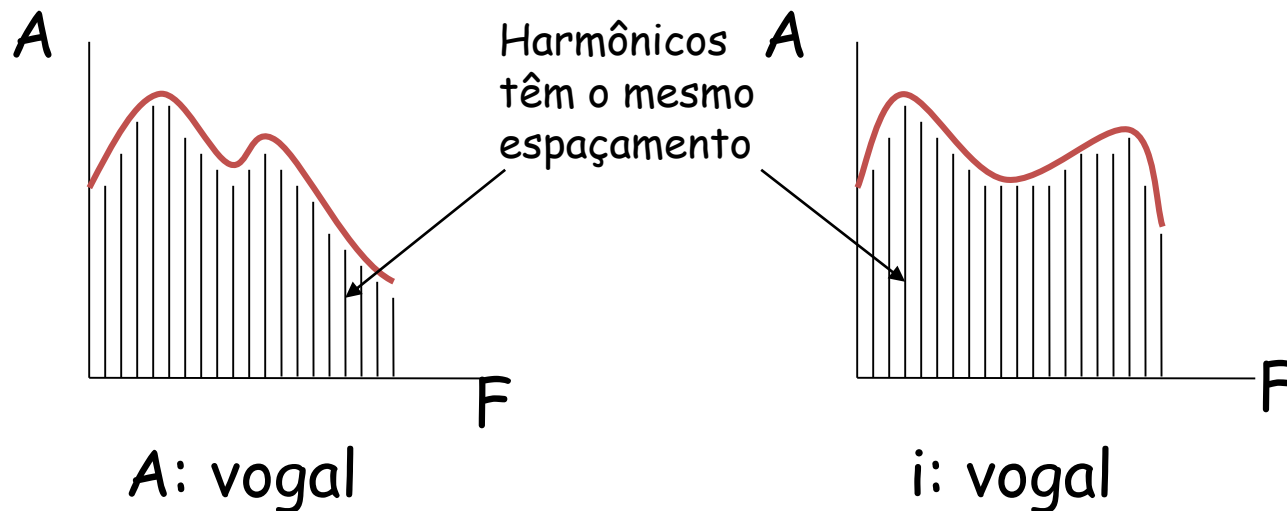
## Mesma Vogal, pitch diferente



Para explicar como o ouvido reconhece um som como uma mesma vogal, mesmo se falada com alturas diferentes, a idéia dos formantes deve ser invocada. Os dados mostram uma vogal "A" que apresenta um envelope similar, mesmo em diferentes alturas.

# Percepção de vogais

- O timbre das vogais depende da amplitude relativa dos harmônicos.
- Essas amplitudes são fortemente afetadas pelo **filtro**: em particular, suas frequências ressonantes.



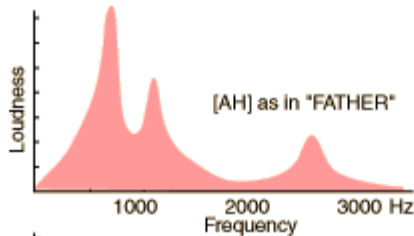


# Percepção de vogais

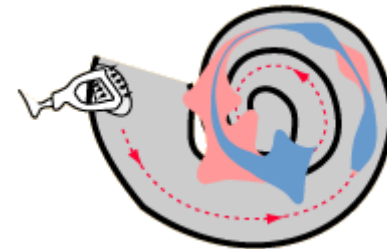
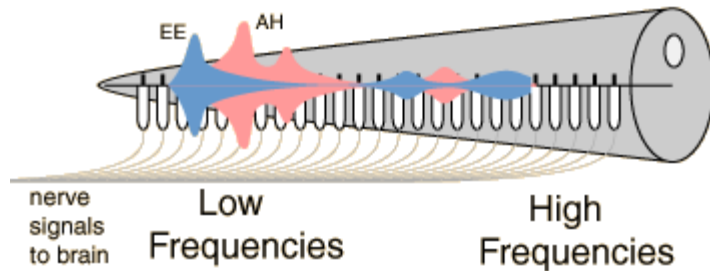
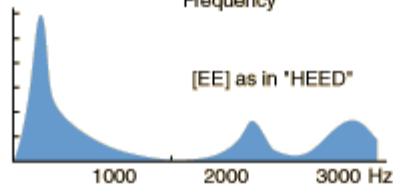
- Resumo

- O espectro da fonte é o mesmo para todas as vogais com uma dada frequência fundamental.
- É o filtro que produz mudanças no timbre, ou qualidade vogal.
- O filtro que dá forma às vogais pode ser caracterizado por frequências ressonantes, ou formantes do trato vocal.

# Percepção de vogais

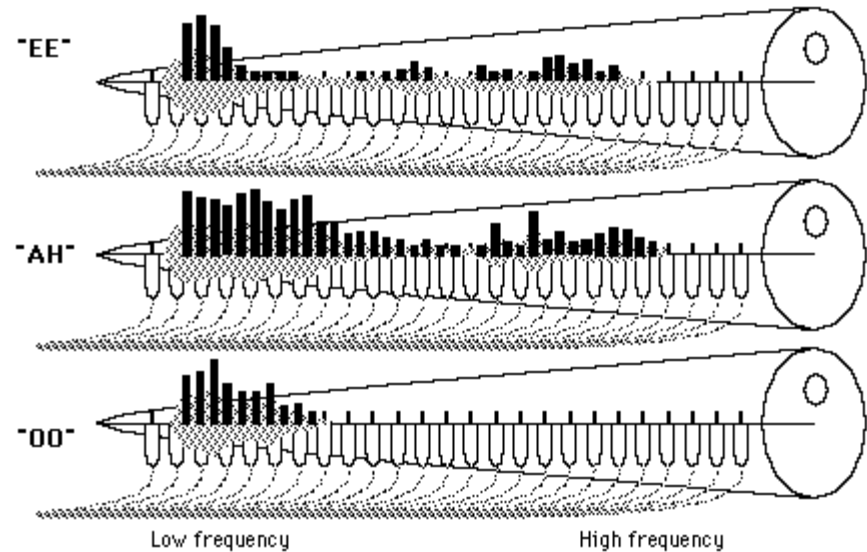


A conceptualization of how the ear distinguishes vowel sounds.



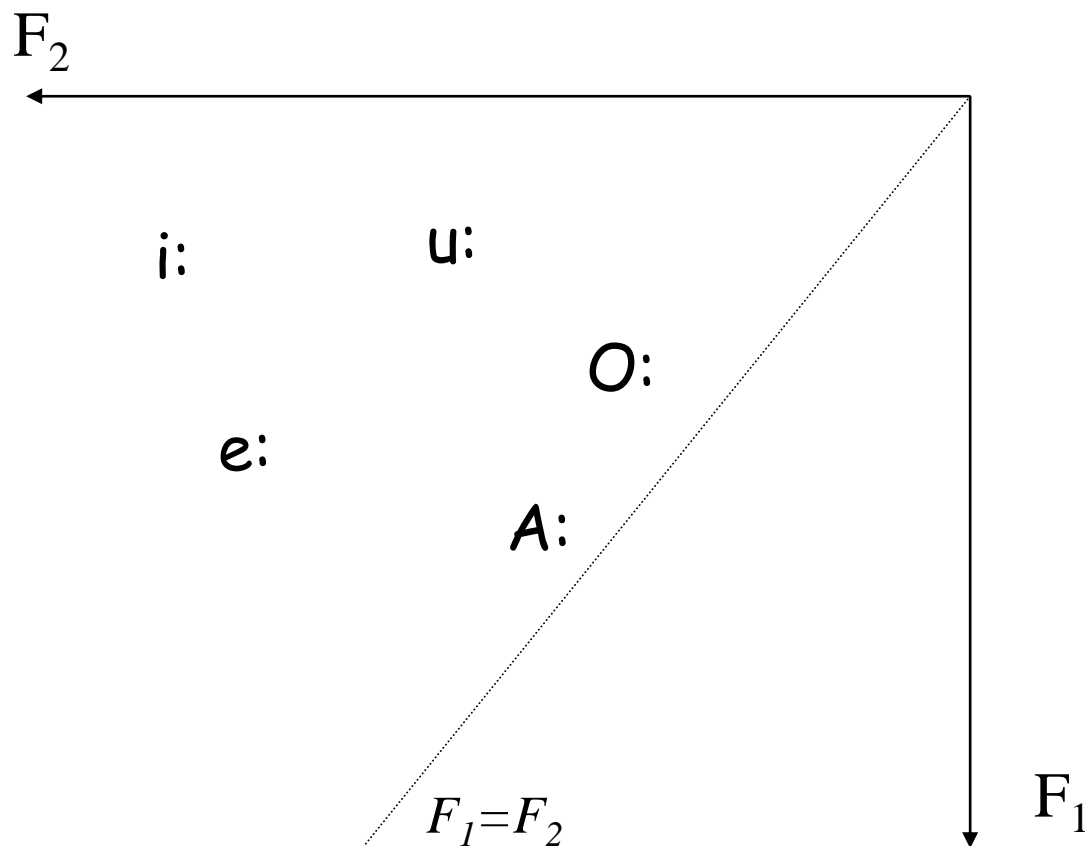
Different vowel sounds, different patterns of excitation of the inner ear.

Conceptual only!  
No actual scaling to ear was done.



# Plano $F_1$ - $F_2$ para vogais

A maior parte da qualidade distinguível das vogais pode ser atribuída a variações relativas das duas primeiras freqüências formantes.



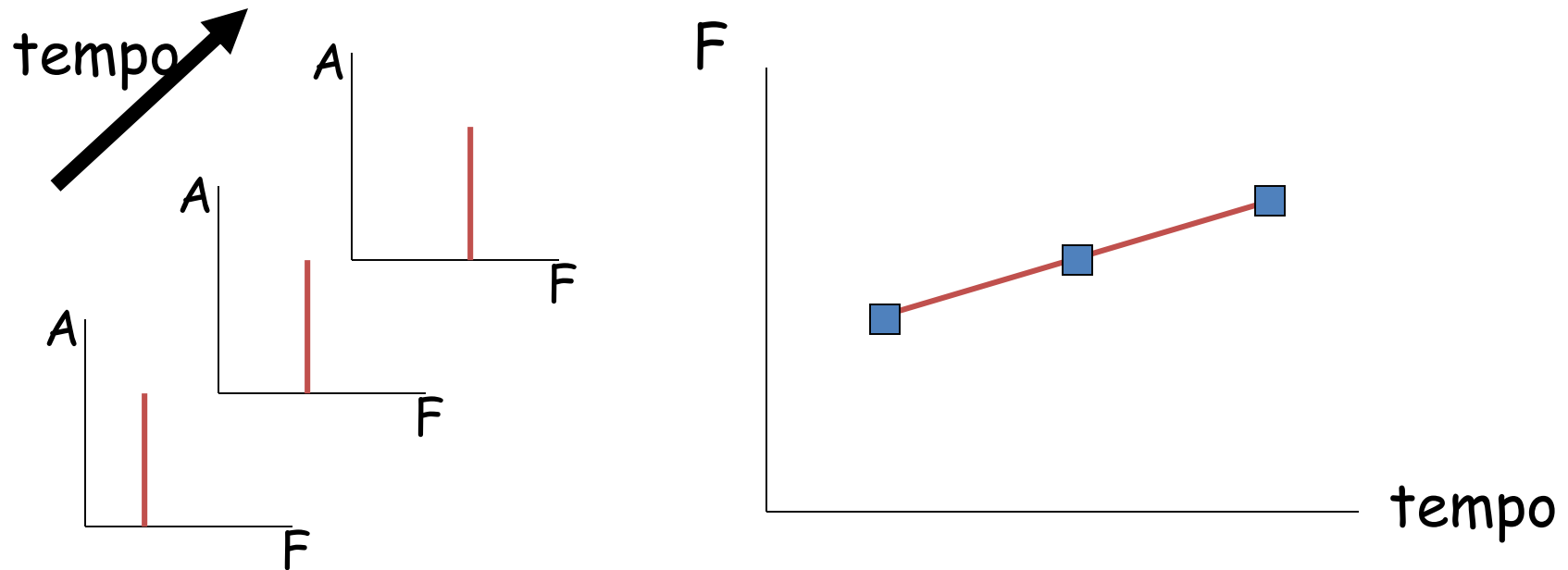
Note-se que a forma acústica das vogais varia bastante por causa de:

- Acentos
- Gestos do orador
- Tamanho do trato vocal
- Contexto prosódico
- Meio consonantal
- Ruído

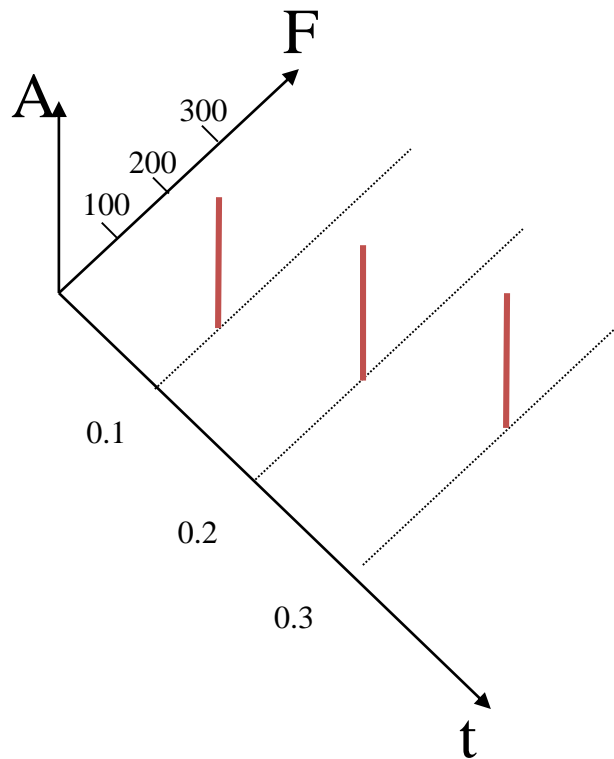
# Espectrografia

Adiciona um dimensão temporal à análise espectral.

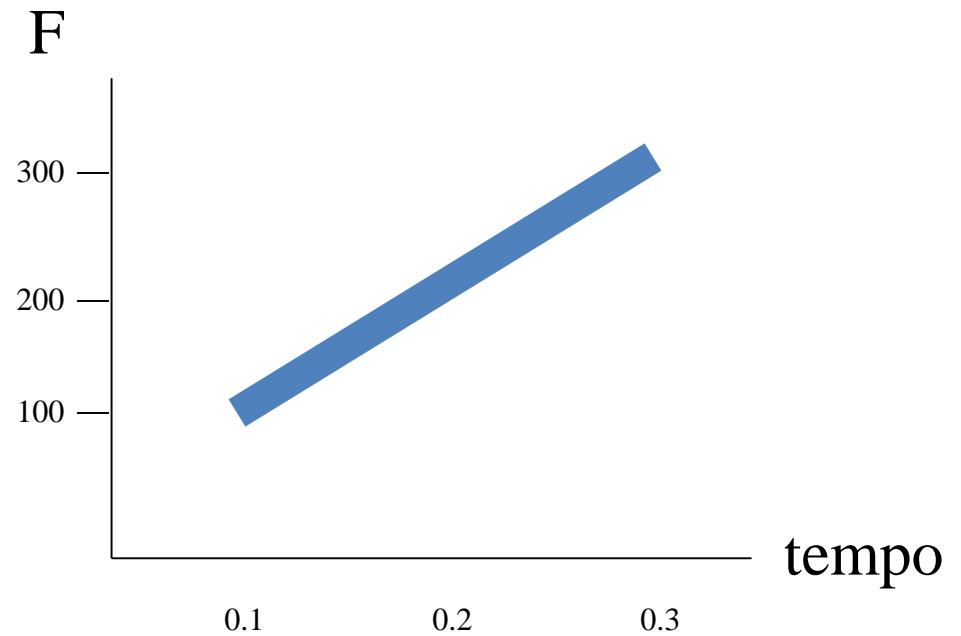
ex. Uma única onda senoidal aumentando em frequência



# Espectrografia

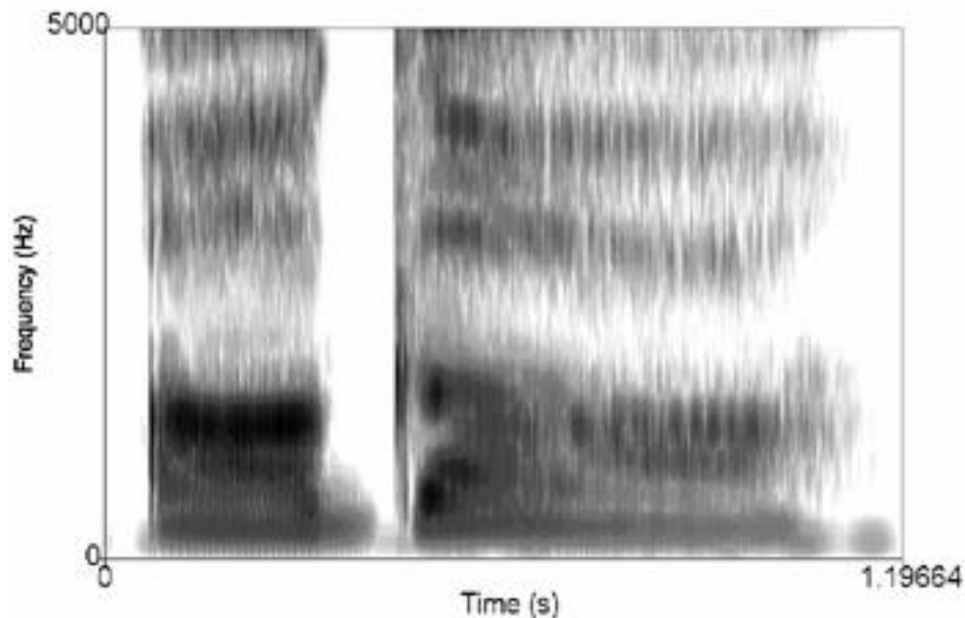


## Espectrograma

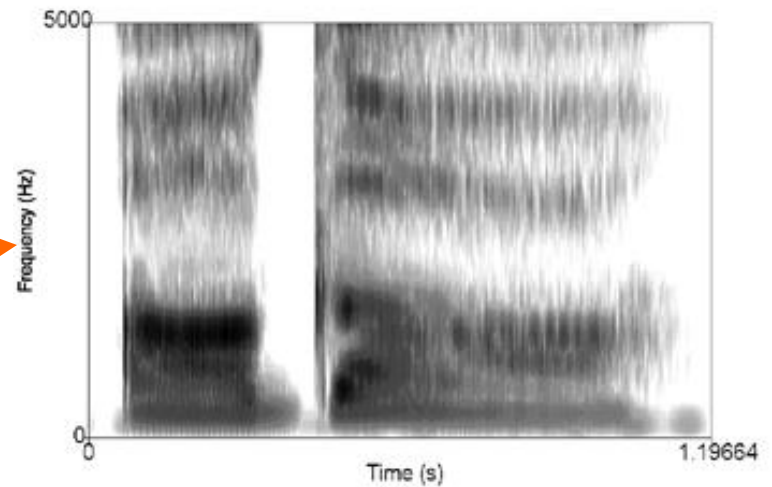
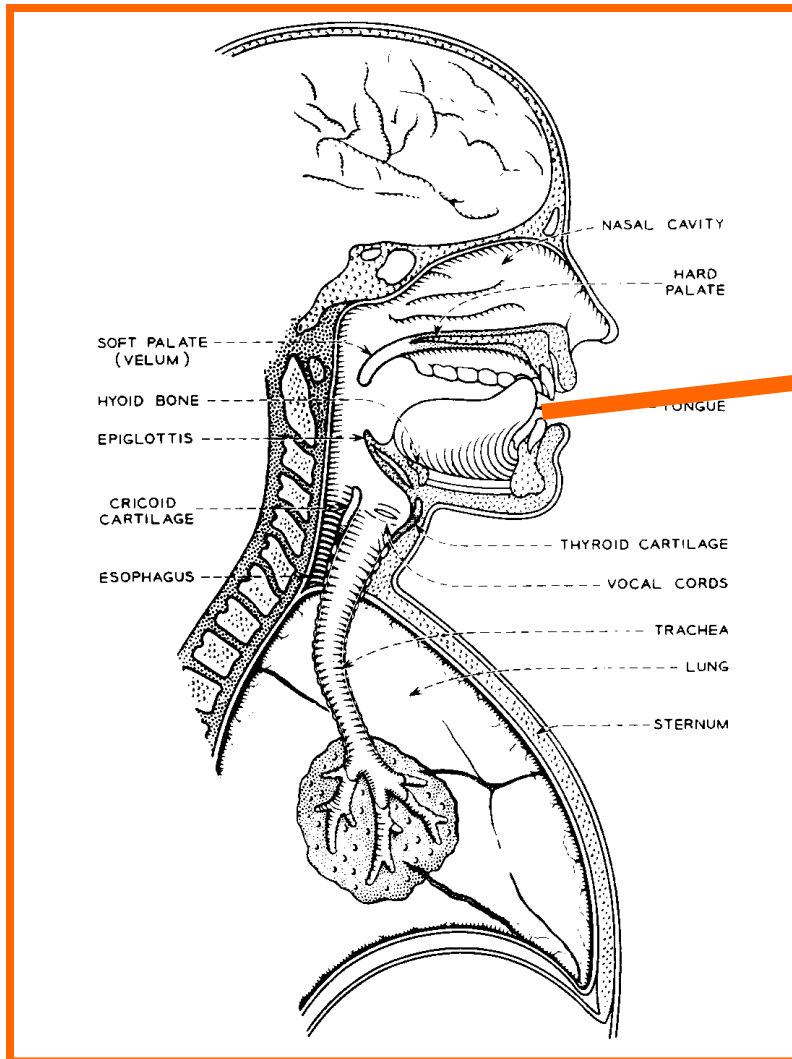


# Espectrograma

- É um gráfico do conteúdo em frequência de um sinal na forma de um gráfico em função do tempo.
- O eixo horizontal é o tempo
- O eixo vertical é a frequência
- A amplitude de cada componente presente no sinal em um dado tempo e frequência é mostrado numa escala de cinza (branco=pouco, preto=muito)



# A Acústica da Produção da Fala



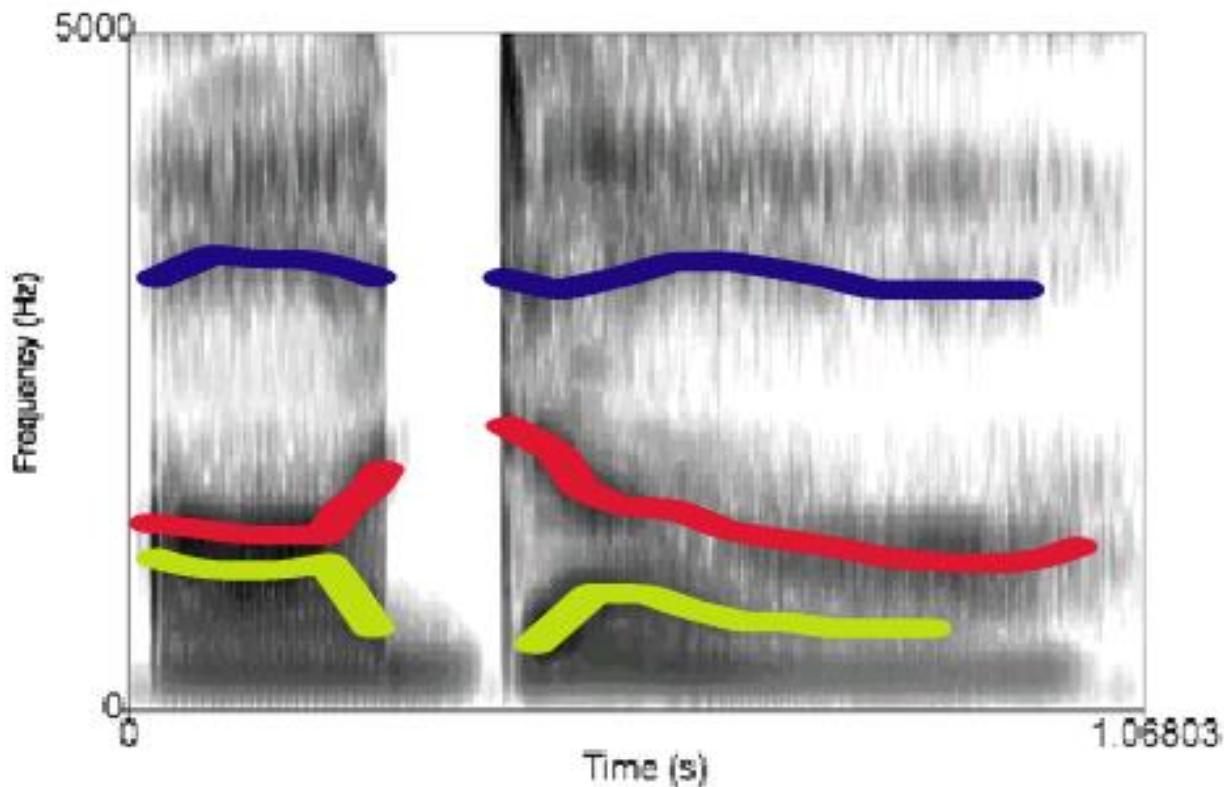
# O que ver em um espectrograma?

- Fala
  - Os pulsos periódicos da vibração das pregas vocais aparecem como estrias verticais.
- Ressonâncias Formantes de Vogais
  - Aparecem como regiões escuras entre pulsos.
- Fricativos (ffff...)
  - Aparecem como regiões "salpicadas", tipicamente em altas frequências



# Frequências Formantes

- O primeiro formante ( $F_1$ ) - Afetado pelo tamanho da constricção
- O segundo formante ( $F_2$ ) - Afetado pelo local da articulação
- O terceiro formante ( $F_3$ ) - Afetado pelo local da articulação

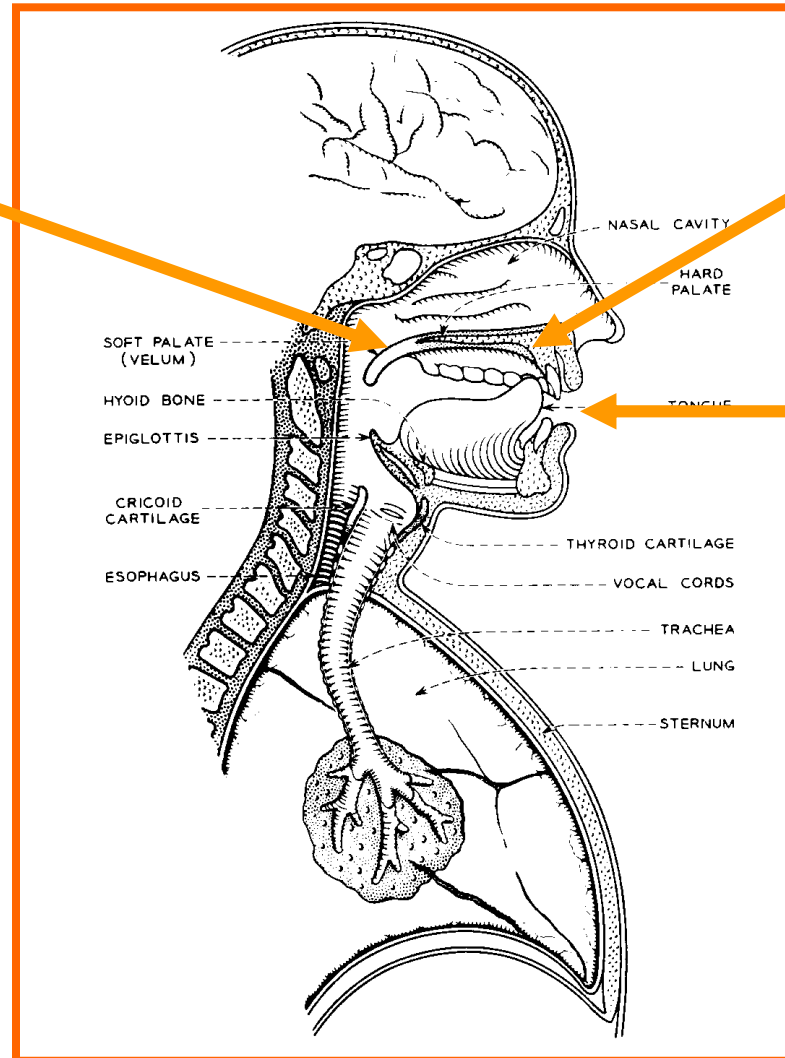


# Local de Articulação

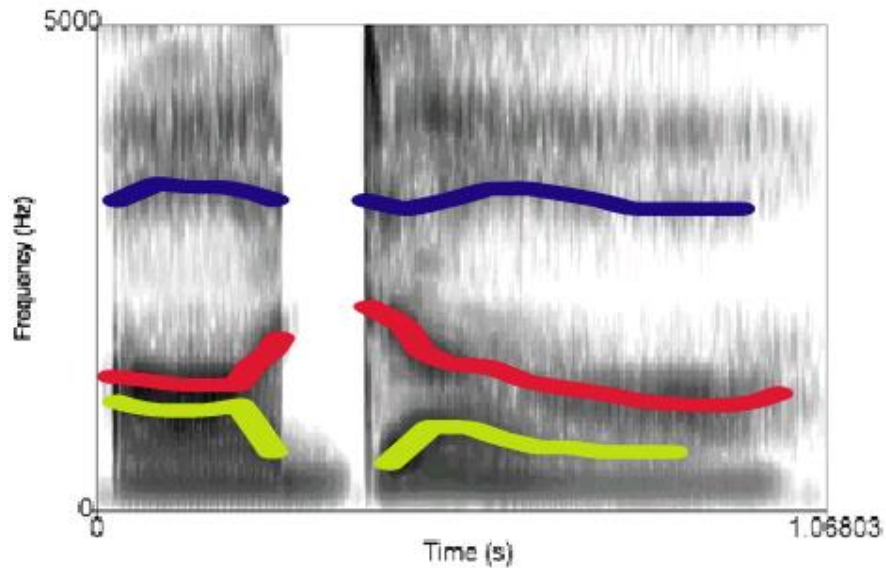
Velar

Alveolar

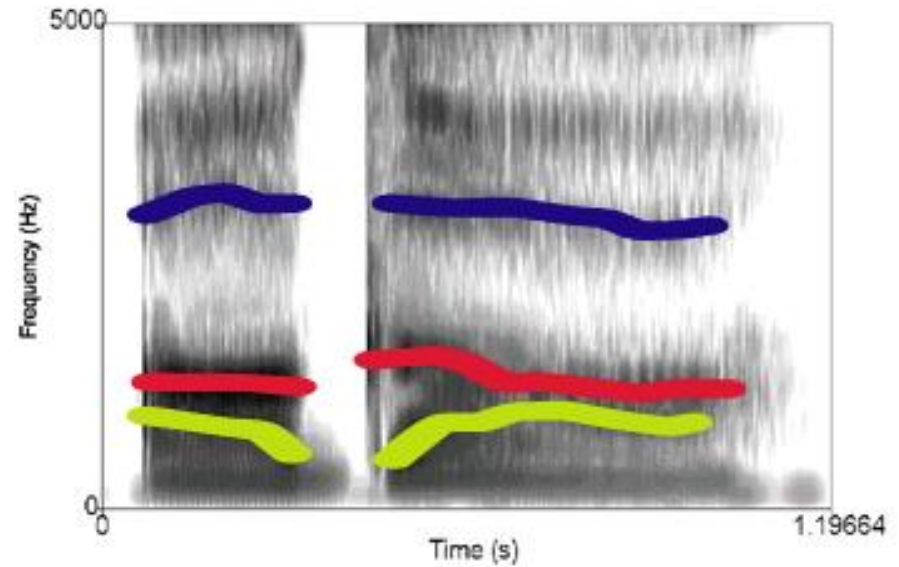
Bilabial



# Local de Articulação



/AdA/



/AgA/