
Noções básicas de acústica e voz com o uso do Praat

Linguistics Summer School 2024 - LiS 2024
Universidade Federal de Juiz de Fora

Aula 3

Prof.^a Dra. Aline Alves Fonseca
Me. Bianca Chaves Leite Lignani

Cronograma

Dia 13/03 (quarta-feira, das 9h às 12h)

- Unidades prosódicas e entoacionais.
- Qualidade de voz.
- Análise das unidades prosódicas e da qualidade de voz no ***Praat***.

Fonologia Prosódica

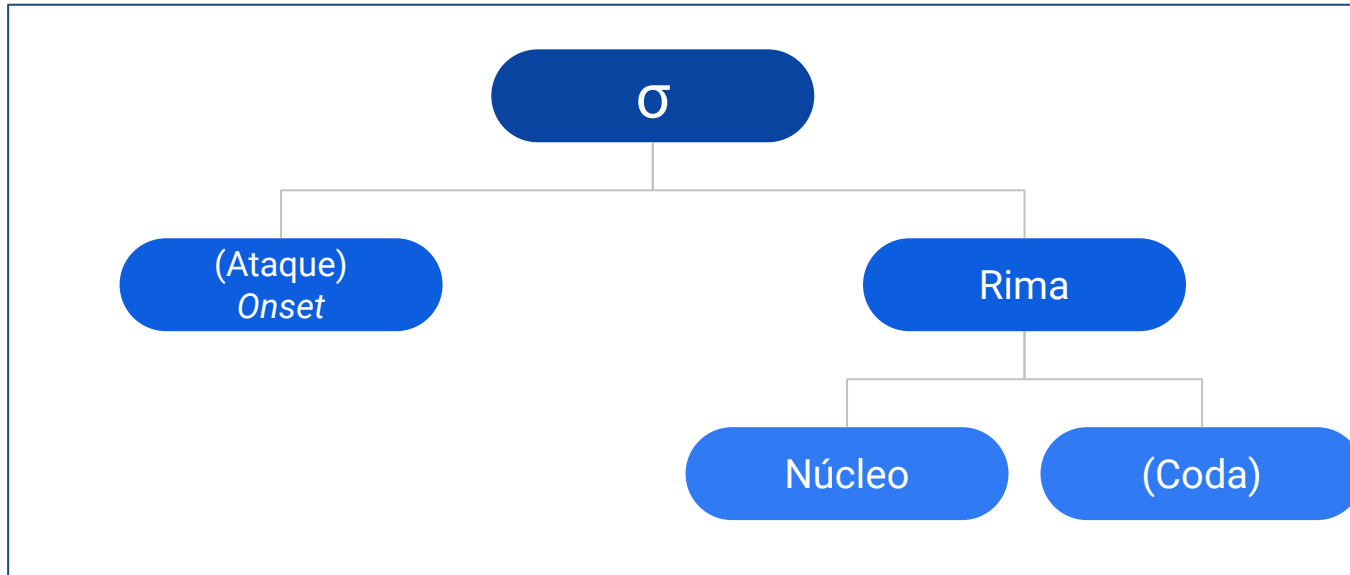
- A **Prosódia** é a disciplina que abrange os **fenômenos suprasegmentais** da fala - **entoação, acento, duração, ritmo, intensidade, pausa e tom** - em interdisciplinaridade com características que estão envolvidas na composição de segmentos (Barbosa, 2019).
- Em linhas gerais, o estudo prosódico de um dado enunciado permite que seja analisada **sua melodia (contorno melódico, entoacional ou prosódico)**.
- Isso implica desvendar características acústicas, fonético-fonológicas, psicológicas e/ou culturais sobre o falante que produziu o enunciado, sobre a língua na qual ele foi produzido e/ou sobre o seu contexto.

Sílabas

- Unidade básica da fala - vogais e consoantes só podem ocorrer dentro de sílabas.
- A sílaba canônica é a CV (consoante-vogal);
 - Não há línguas que não apresentem sílabas CV;
 - Sílabas CV são as primeiras adquiridas durante a aquisição de língua;
 - Mesmo em casos severos de afasias motoras, sílabas CV são preservadas;
 - Processos diacrônicos de reestruturação silábica no curso de evolução de uma língua tendem a criar sílabas CV.
- As **sílabas** podem ser analisadas sob o ponto de vista **fonológico** (fonemas) ou sob o **fonético** (fones).

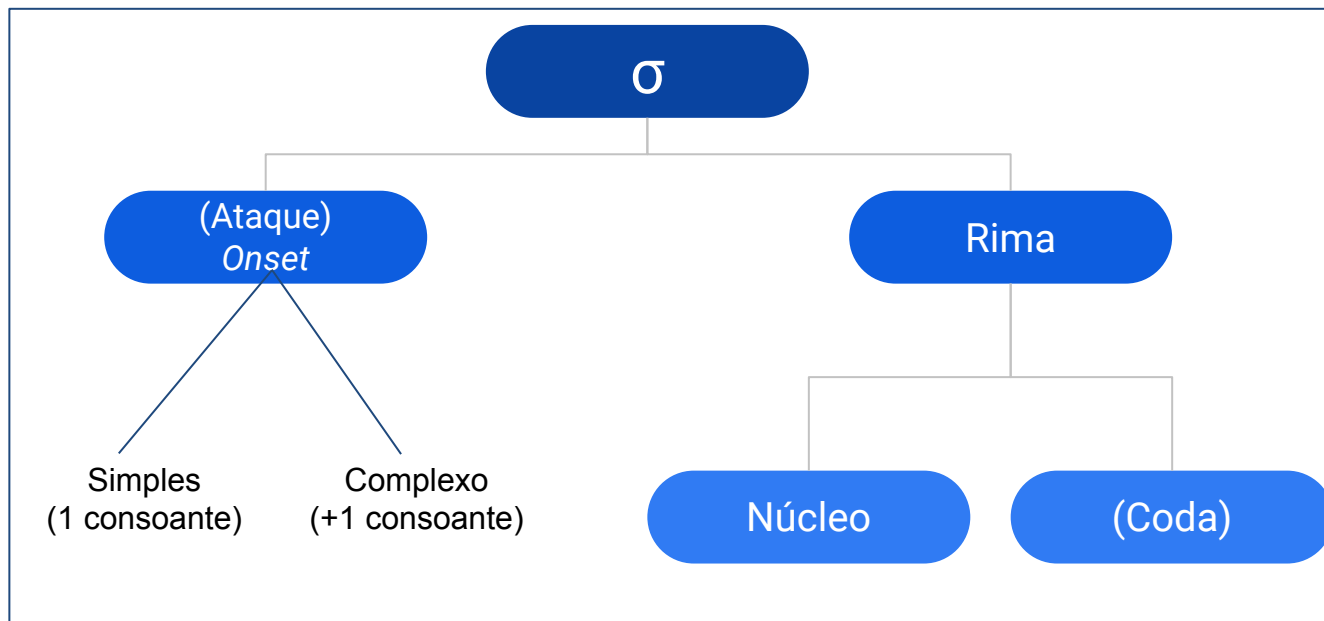
Sílaba Fonológica

- A ‘**sílaba fonológica**’ pode ser decomposta em um **ataque** (*onset*) e em uma **rima** (Barbosa, 2019).



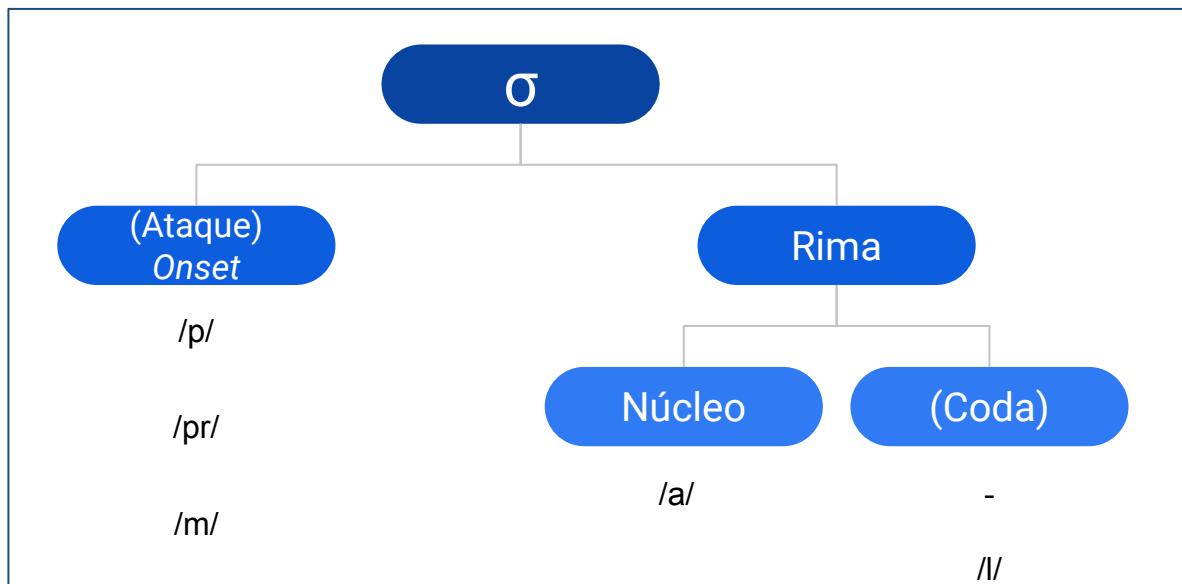
Sílaba Fonológica

- O ataque pode conter uma ou duas consoantes, classificando-se, respectivamente, como **simples** ou **complexo**.



Sílaba Fonológica

- A **rima** pode ser subdividida em **núcleo** e **coda**. No português, o núcleo pode ser formado apenas por vogais e a coda, por vogais, consoantes ou grupos consonantais.



Sílaba Fonológica

- Considerando-se a estrutura da **sílaba fonológica**, no Português existem vários tipos distintos de sílabas:
 - CV (consoante-vogal): pá; ca.sa;
 - CVC: pas.ta
 - V: a.mor
 - VC: ar.ma
 - CCVC: cruz
 - CCVCC: trans.por.te
 - CVV: má.goa
 - CVVC: pais
 - CCVVC: claus.tro
 - CCV: tra.ma
 - VV: eu


Sílaba Fonética

- A **sílaba fonética** trata-se da realização de uma sílaba fonológica na cadeia da fala.
- Considerando a execução acústica de uma sequência de fonemas, pode haver um *mismatch* entre os segmentos que representam a sílaba fonológica (fonemas) e a sílaba foneticamente realizada. Esse mismatch se dá por apagamentos ou pela inserção de vogais epentéticas que vão simplificar a construção silábica da palavra.
- Esse é o caso de muitas palavras do Português Brasileiro (PB), como:
 - Advogado [dSi]
 - Leite (1 ou 2 sílabas fonéticas)
 - Leit[l] ou Lei[tS]
 - xícara [kra]

Palavra Fonológica (ou Prosódica)

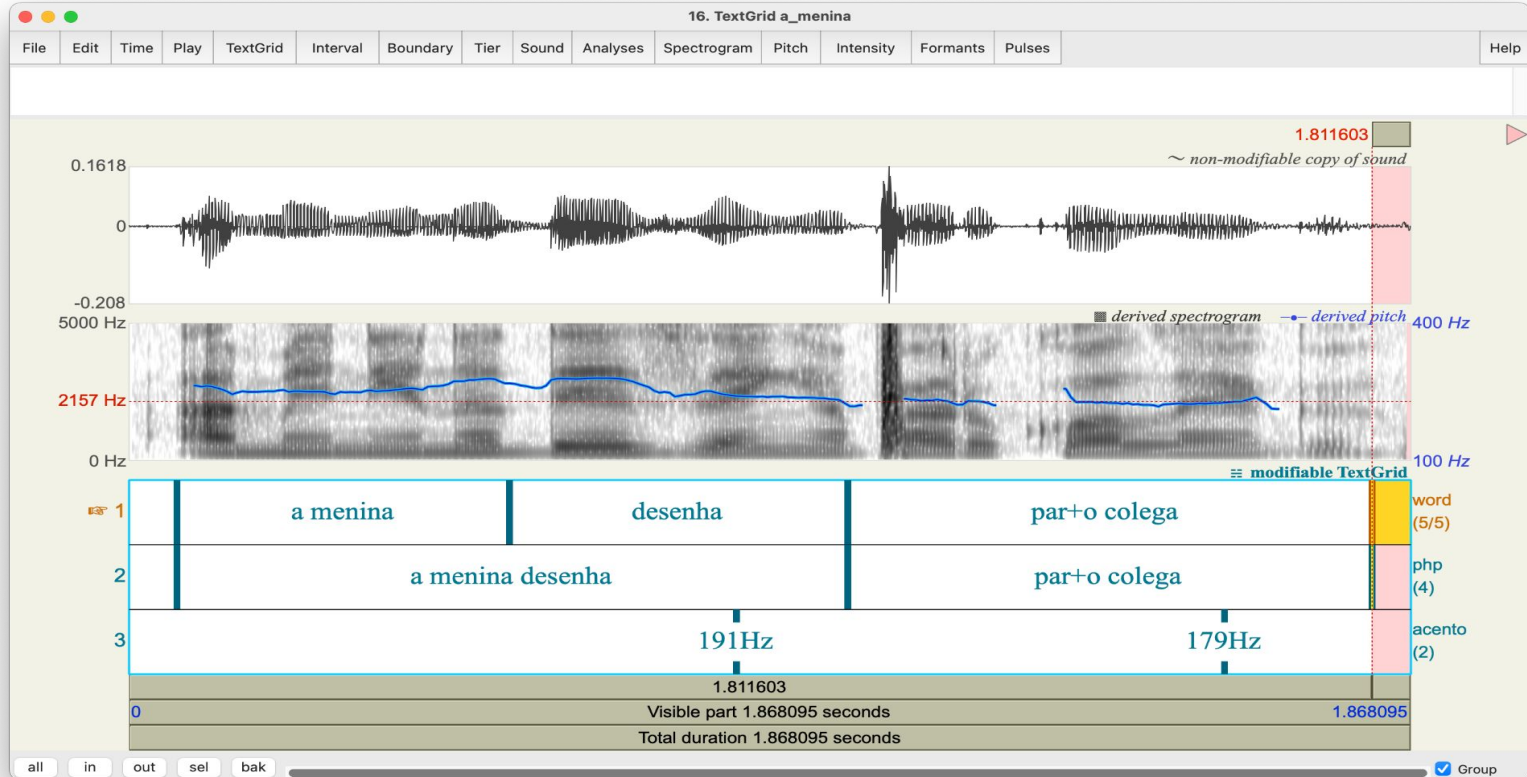
- A **palavra fonológica ou prosódica** é um conjunto de sílabas, no qual somente uma delas porta acento primário e as demais são átonas associadas (Nespor e Vogel, 2007).
- Por exemplo, no enunciado '**A menina desenha para o colega**', temos:
 - **palavras funcionais (clíticos):** 'a', 'para' e 'o';
 - **palavras fonológicas (prosódicas):** 'a menina', 'desenha', 'para o colega'.
- Note que os clíticos são sempre formados por **sílabas átonas**, que podem sofrer (e frequentemente sofrem) processo de redução durante a realização fonética.
 - Ex: "para o" (pro) colega.

Sintagma fonológico

- Um sintagma fonológico (PhP) é formado por um núcleo lexical (N, V, A) e palavras adjacentes (Nespor e Vogel, 2007) :
 - sintagma fonológico com 1 palavra fonológica: o acento da sílaba acentuada é proeminente no sintagma.
 - sintagma fonológico com 2 ou mais palavras fonológicas: somente um dos acentos é proeminente no sintagma (o que estiver mais a direita).
- Em '**A menina desenha para o colega**', podemos ter 2 sintagmas fonológicos:
 - 'A menina desenha' (reestruturado) e 'para o colega'. 
- O **acento de sintagma** do primeiro PhP é na sílaba 'se', de 'desenha' e o acento secundário, na sílaba 'ni', de 'a menina'. No segundo PhP, o acento de sintagma está em 'le' de colega.

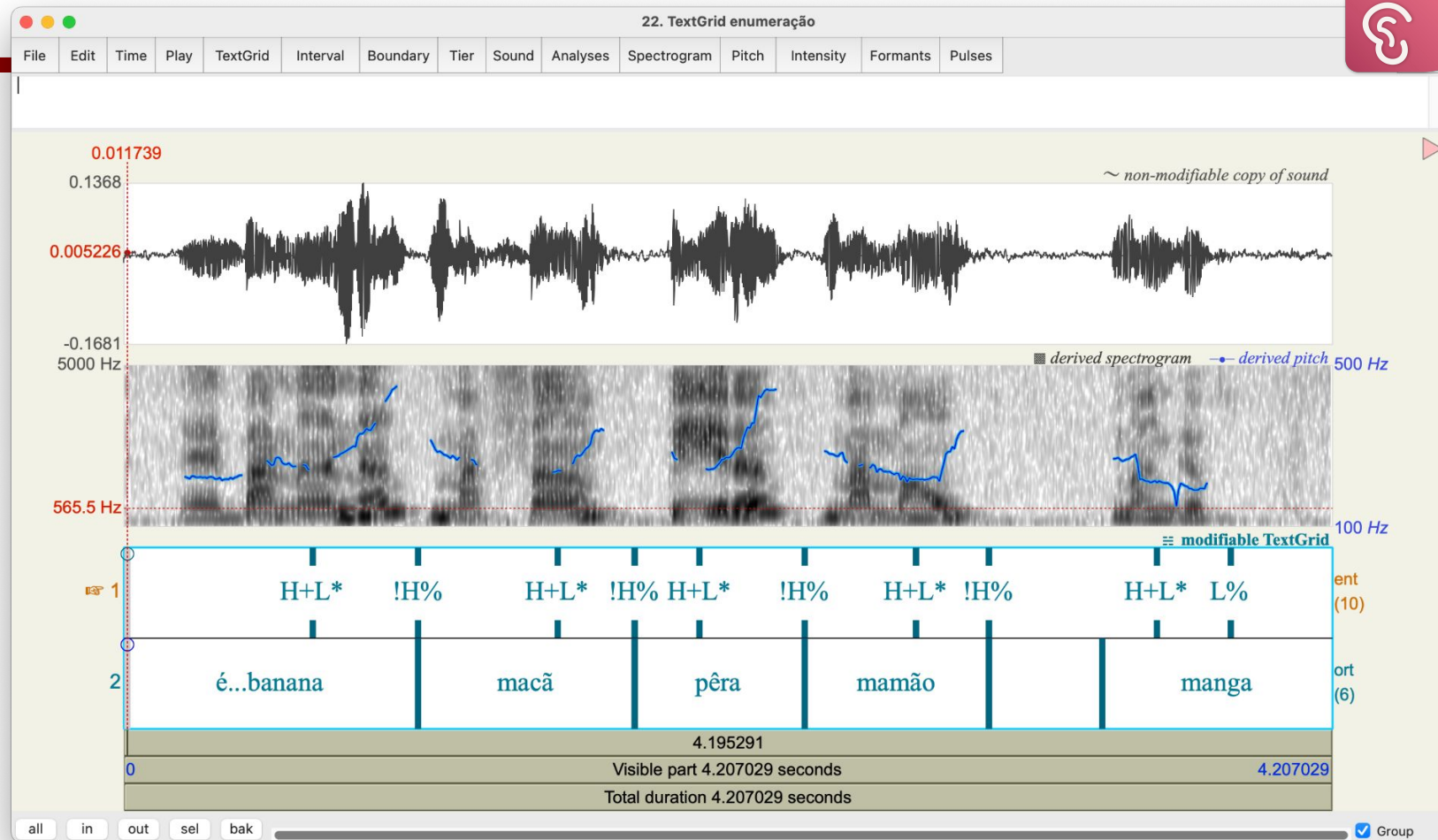
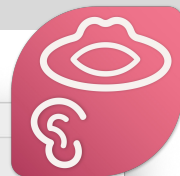


Sintagmas fonológicos



Sintagma entoacional (*Intonational Phrase*)

- Os **sintagmas entoacionais** são grupamentos mais complexos do que os sintagmas fonológicos, à medida que utilizam como critério de delimitação a própria curva melódica (Nespor e Vogel, 2007).
- Essa curva corresponde ao contorno de Pitch, que pode assumir um **padrão ascendente** ou um **padrão descendente**.
- Os movimentos de uma curva melódica, bem como pausas silenciosas, sinalizam as fronteiras entre os sintagmas entoacionais e podem fornecer pistas sobre o tipo de enunciado que está sendo produzido.



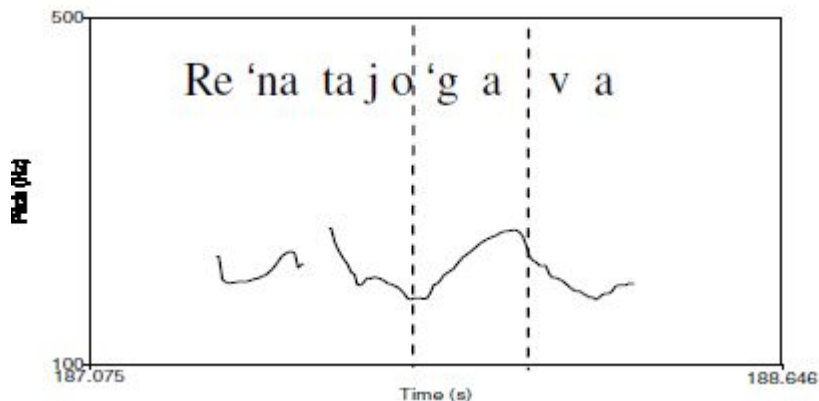
Contornos Prosódicos

- A unidade de análise do campo da **Prosódia** mais estudada é o **contorno prosódico (ou melódico)**, que corresponde a uma curva entoacional representante das diversas variações de frequência e de "tom" (notação entoacional H e L),
- No ***Praat***, esse contorno é visível quando selecionada a opção: **Pitch > Show Pitch** na janela de visualização do som (seleção default).
- O desenho que o **contorno melódico** imprime em recursos que o tornam visível (como o ***Praat***) permite a análise desses correlatos acústicos e prosódicos também sob uma perspectiva perceptual, interpretando variações de *pitch*, por exemplo, como noções de movimento grave/agudo durante a fala.
- Um contorno prosódico pode ser subdividido em diferentes tipos de **unidades prosódicas**.

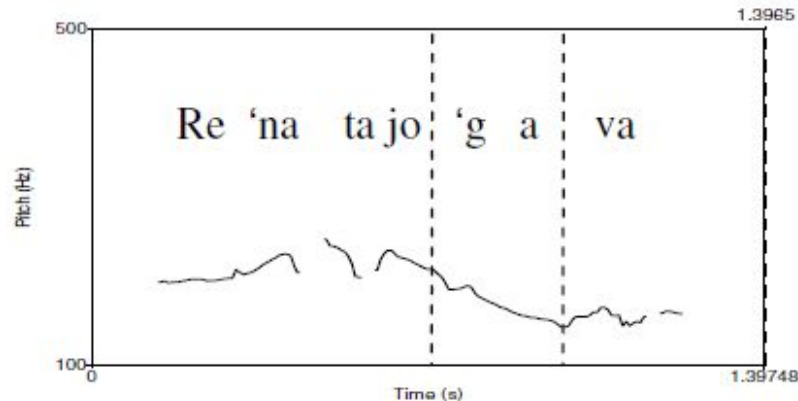
Contornos prosódicos

- Diferentes padrões de contornos prosódicos podem ser associados a diferentes tipos de enunciados, em diferentes línguas.
- No PB, por exemplo, associa-se um padrão com final ascendente para perguntas-QU (1) e um padrão com final descendente para enunciados afirmativos (2) (Moraes, 2008).

(1)



(2)



Sistema Entoacional

- O sistema entoacional apresenta uma série de dificuldades de análise que estão relacionadas com o carácter abstrato deste constituinte linguístico.
 - O alinhamento tom-sílaba é variável, de acordo com o material fonético usado na fala.
 - Os tons não são referenciais. Não se pode contrastar dois contornos tonais a não ser de forma experimental.
 - É muito complicado separar padrões entoacionais de variações estilísticas.

Sistema Entoacional

- Inicialmente, havia a competição entre duas abordagens de abstração dos padrões entoacionais :
 - Trager and Smith (1951) e Pike (1945), a melodia era decomposta em termos de níveis tonais, formando 4 níveis foneticamente distintos: L, LM, HM, H.
 - Bolinger (1951, 1958) (British School), a melodia era decomposta em variações da F0 ou em trajetórias melódicas (fall, rise, low, high e suas combinações)

Sistema Entoacional

Pierrehumbert	British-style
H* L L%	fall
H* L H%	fall-rise
H* H L%	stylised high rise
H* H H%	high rise
L* L L%	low fall
L* L H%	low rise (narrow pitch range)
L* H L%	stylised low rise
L* H H%	low rise
L+H* L L%	rise-fall
L+H* L H%	rise-fall-rise
L+H* H L%	stylised high rise (with low head)*
L+H* H H%	high rise (with low head)
L*+H L L%	rise-fall ('scooped')
L*+H L H%	rise-fall-rise ('scooped')
L*+H H L%	stylised low rise*
L*+H H H%	low rise†
H+L* L L%	low fall (with high head)
H+L* L H%	low rise (with high head)
H+L* H L%	stylised low rise (with high head)
H+L* H H%	low rise (high range)
H*+L H L%	stylised fall ('calling contour')
H*+L H H%	fall-rise (high range)

Ladd (1996), p. 91

Sistema Entoacional

- Modelo de 4 tons: desvantagens
 - É difícil relativizar as variações tonais dentro de espectros de Frequência que também são variáveis por fatores externos como: altura da voz do indivíduo, estado emocional, mudanças de estilo etc. (Ladd, 1996, p.14)
 - O modelo falha por não poder categorizar mais do que 4 variações de tons em um mesmo enunciado.
- Modelo de variação de F0: desvantagens
 - As combinações de low, high, rise e fall não são suficientes para descrever (contrastivamente) todos os padrões melódicos e acaba por necessitar de descrições adicionais.

Sistema Entoacional

- A solução encontrada para esses problemas foi reduzir o inventário de 4 tons para um modelo binário.
- O mais bem sucedido modelo de 2 tons para a análise entoacional de padrões tonais do Inglês foi o de Pierrehumbert (1980) e posteriores.
- O trabalho de Pierrehumbert teve como referência o trabalho de Bruce (1977) para o Sueco.

Sistema Entoacional

- A notação de Pierrehumbert para o Inglês se divide em acentos tonais e tons de borda (tom fronteira e acentos de sintagma).
- Os acentos tonais podem ser bitonais. Um deles será o central e será marcado por um asterisco (tom estrelado) o outro poderá ser anterior ou posterior ao tom estrelado. Usa-se um sinal de + para ligar os dois tons que compõem o acento tonal (ex. L+H* ou L*+H).
- O tom fronteira é um tom simples marcado pelo símbolo de porcentagem (H% ou L%).
- O acento de sintagma também é um tom simples, marcado pelo sinal de – sobrescrito (H⁻ ou L⁻).

Sistema Entoacional

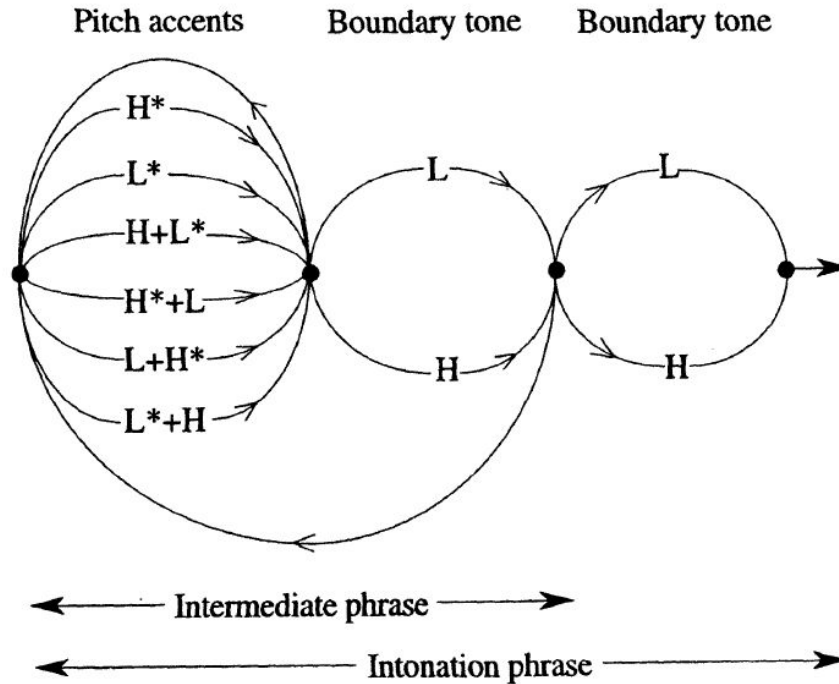


Figure 6. The grammar of English intonation patterns, according to Beckman and Pierrehumbert (1986).

Sistema Entoacional

- Três condições contextuais são importantes para a análise entoacional seguindo o modelo proposto por P&B (1986):
 - A faixa de F0 do falante.
 - A regra de downstep (escalonamento decrescente de tons H)
 - A regra de upstep (o não abaixamento após um tom H pode ser considerado um tom L)

Sistema Entoacinal

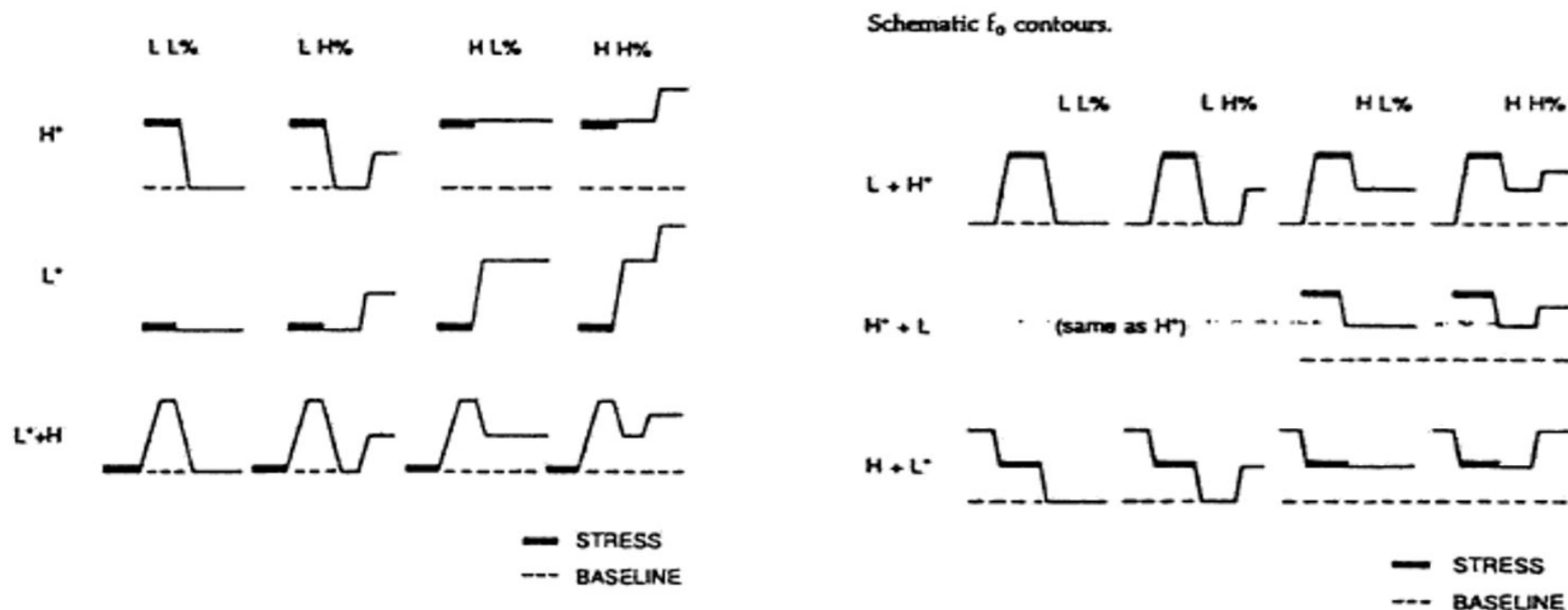
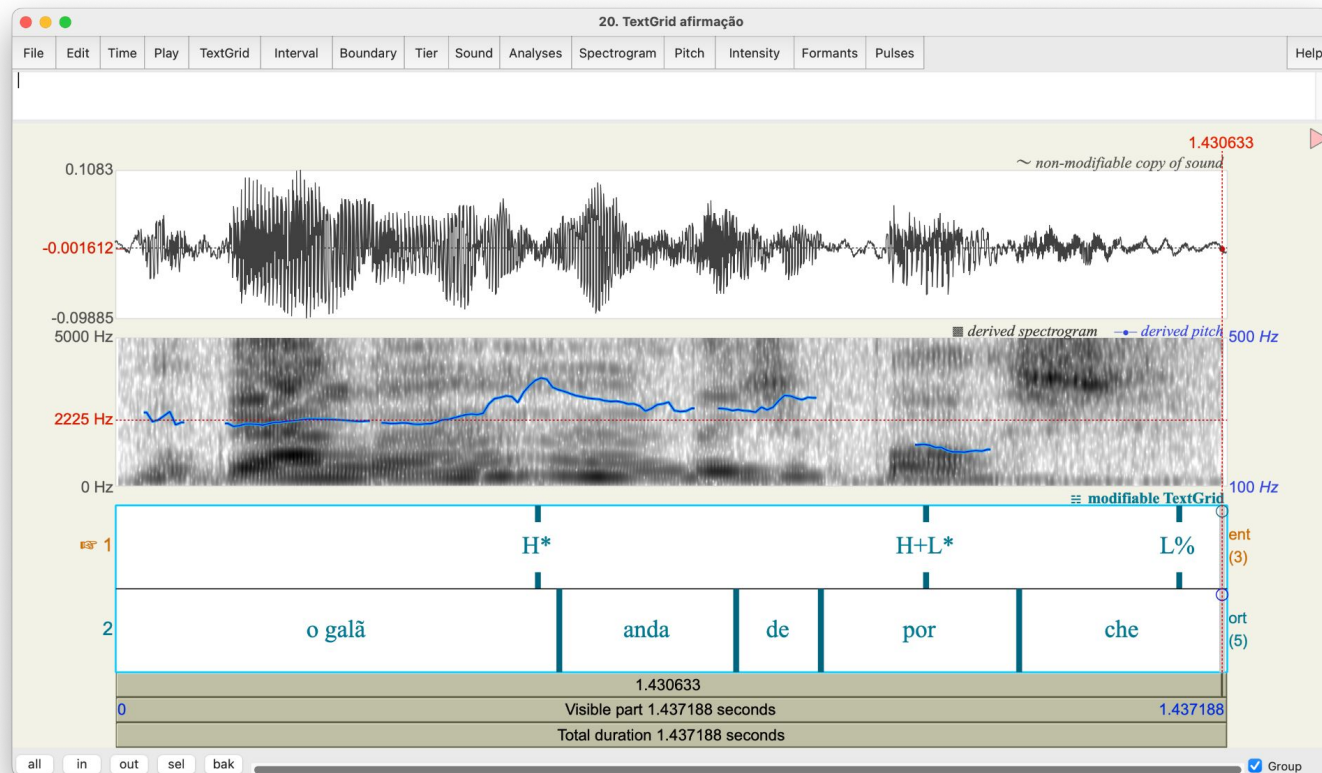


Figure 8. The full inventory of nuclear accents in combination with phrase-final tones, as schematized in Pierrehumbert and Hirschberg (1990).

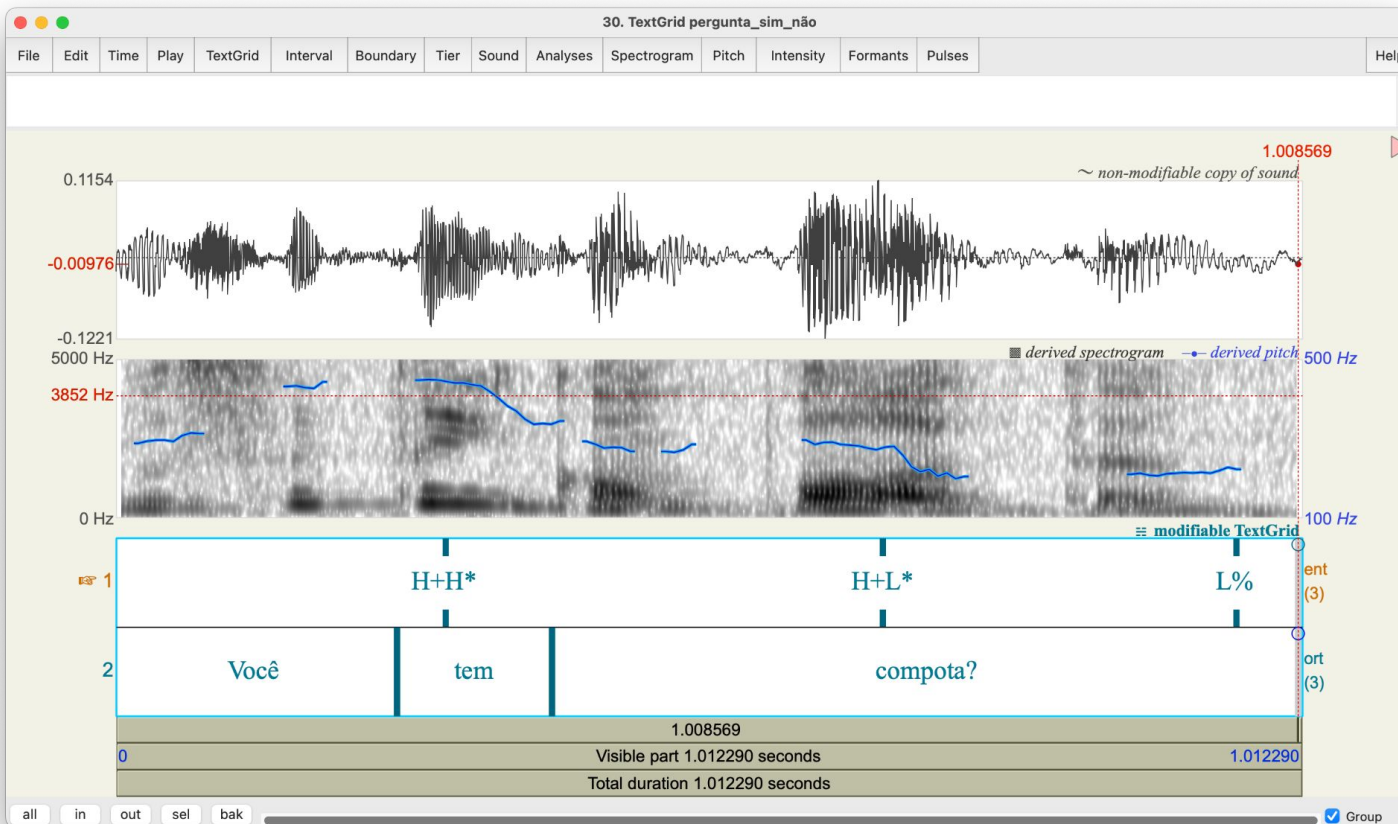


Exemplos do PB - afirmação



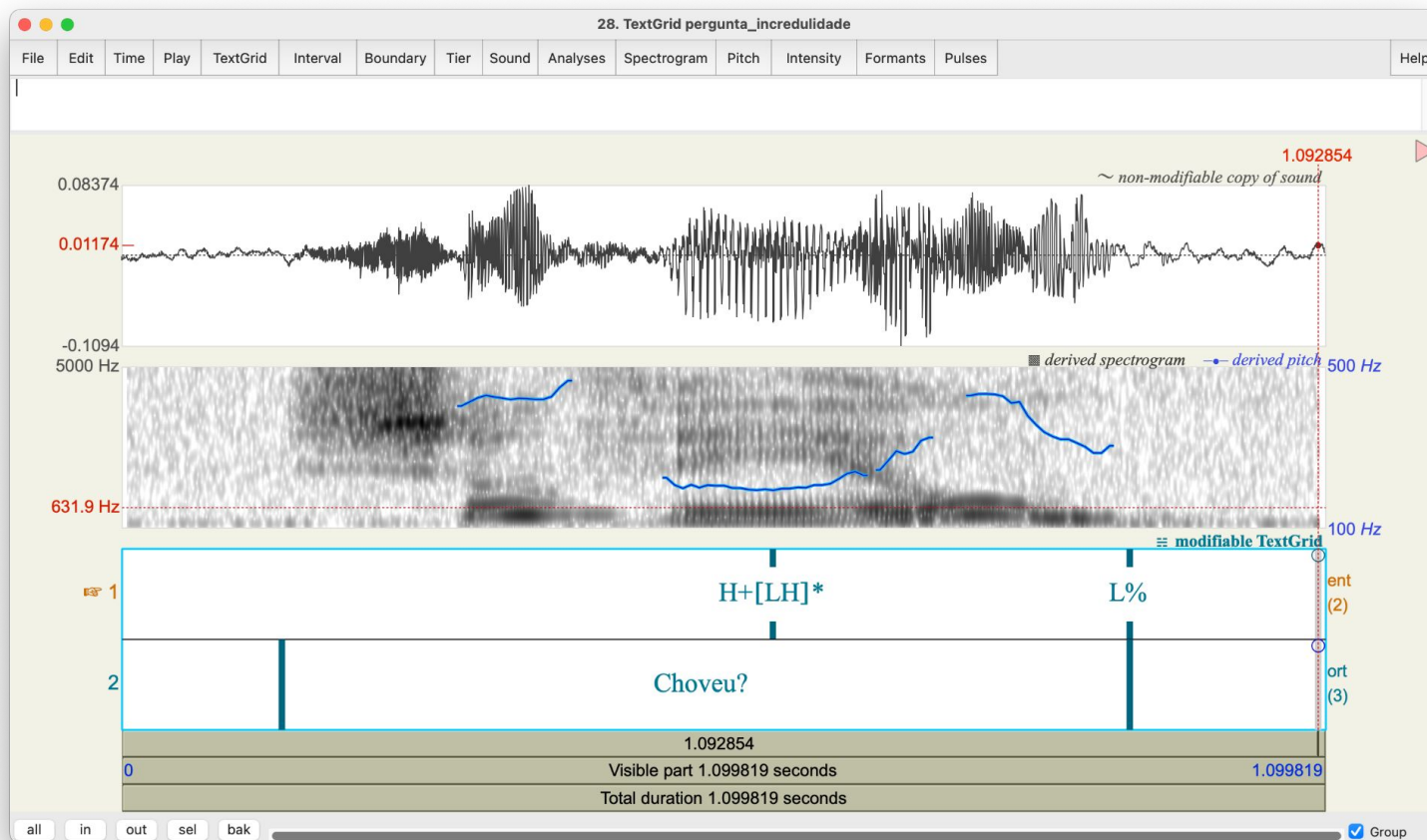


Exemplos do PB - pergunta total





Exemplos do PB - pergunta de incredulidade



Atividade de fixação - Parte 1

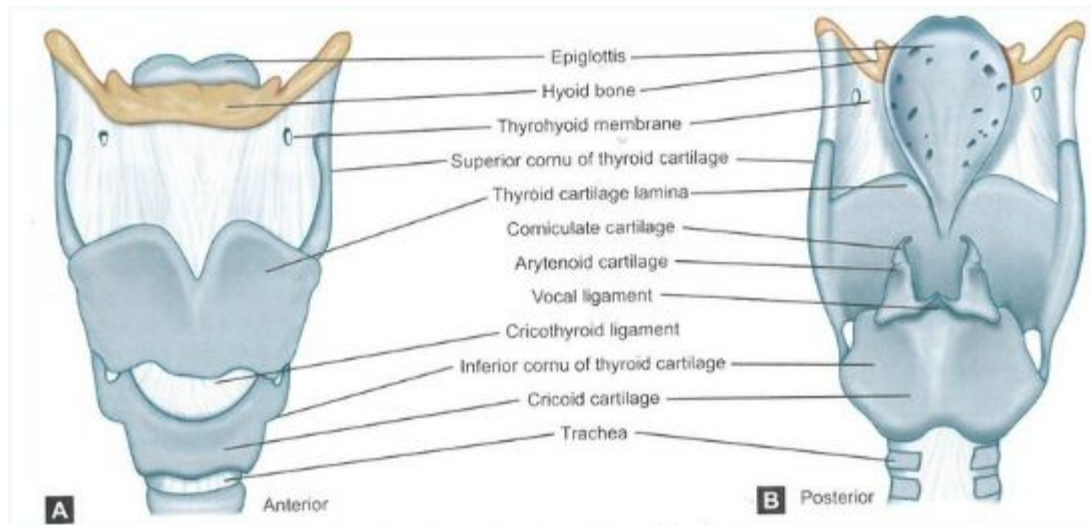


Exercício

1. Grave três enunciados no *Praat*:
 - a. 'Camila perdeu a caneta?', com entoação interrogativa.
 - b. 'Camila perdeu a caneta', com entoação de afirmativa.
 - c. 'Camila perdeu a caneta!', com entoação de exclamação.
2. Para cada enunciado, faça um TextGrid de 4 camadas: 2 camadas de intervalo (sil e pal) e 2 camadas de ponto (Pitch e Tonal)
3. Na camada sil, separe o enunciado em sílabas.
4. Na camada pal, separe os enunciados em palavras.
5. Na camada Pitch, anote o valor do pitch na porção estável das vogais pré-tônica, tônica e pós-tônica da palavra 'caneta'.
6. Na camada Tonal, anote o tom (H ou L) encontrado nas três sílabas da palavra 'caneta'.

Qualidade de Voz

- Conforme visto anteriormente, **qualidade de voz** compreende aspectos do modo de fonação (configuração da laringe) durante a produção da fala.
- Em algumas abordagens, a laringe é tratada como articulador ativo (John Esling, 2019). Isso porque a laringe apresenta um conjunto de cartilagens e músculos que são capazes de modificar a e, por isso, fonação.



Qualidade de Voz

- As diferentes configurações que a laringe pode assumir alteram os aspectos acústicos da voz.
- Esses aspectos podem ser involuntários, relacionados ao porte físico e/ou ao estado emocional do falante, por exemplo.
- A **qualidade da voz percebida** é uma sensação auditiva relacionada a aspectos da voz que não têm relação com f_0 ou *pitch*, compreendidos com relativa facilidade quando evocados.
- Contudo, a intenção e a função comunicativa do falante também podem guiar escolhas voluntárias deste, com relação à qualidade da voz que será percebida pelo ouvinte.

Qualidade da voz percebida

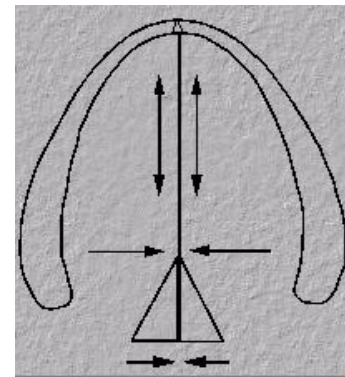
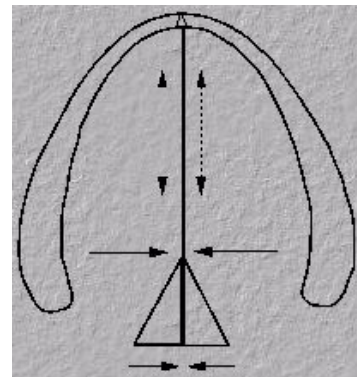
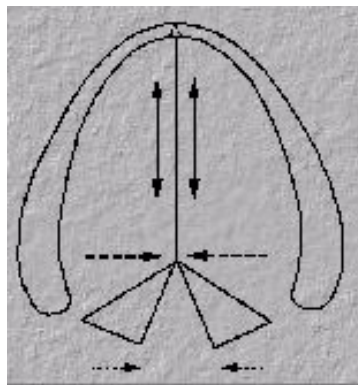
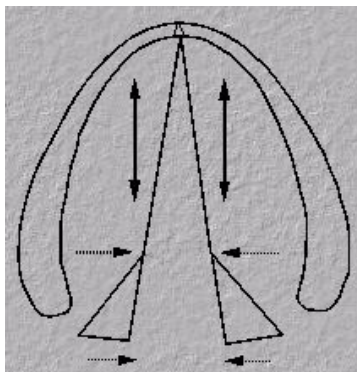
- A voz rouca ou murmurada/soprosa, e de alguns modos de fonação, como *creaky voice* (voz laringalizada ou crepitante) ou *harsh voice* (voz tensa), por exemplo, são tipos de vozes que diferem em termos de qualidade.
- Cada um desses tipos de voz implica um efeito distinto em quem os escuta.
- Ademais, diferentes qualidades de voz são relacionadas a fenômenos sociais distintos: a voz soprosa é socialmente associada ao convencimento e à intimidade, enquanto que a *creaky voice* é associada à dureza, à aspereza (Meirelles, 2020).

Qualidade da voz percebida

- Ouça um exemplo clássico de produção de [voz](#). Identifique as características e a provável finalidade da produção desse tipo de voz, no contexto em questão.
- Ouvindo esse tipo de [voz](#), descreva-o e de contraste-o com o tipo de voz ouvido anteriormente. Nomeie as sensações provocadas por ela em você.
- Ouça essa [voz](#). Descreva-a em termos acústicos e contraste-a com as duas vozes ouvidas anteriormente.
- Ouça essa [voz](#). Descreva-a em termos acústicos e contraste-a com o exemplo de tipo de voz acima.

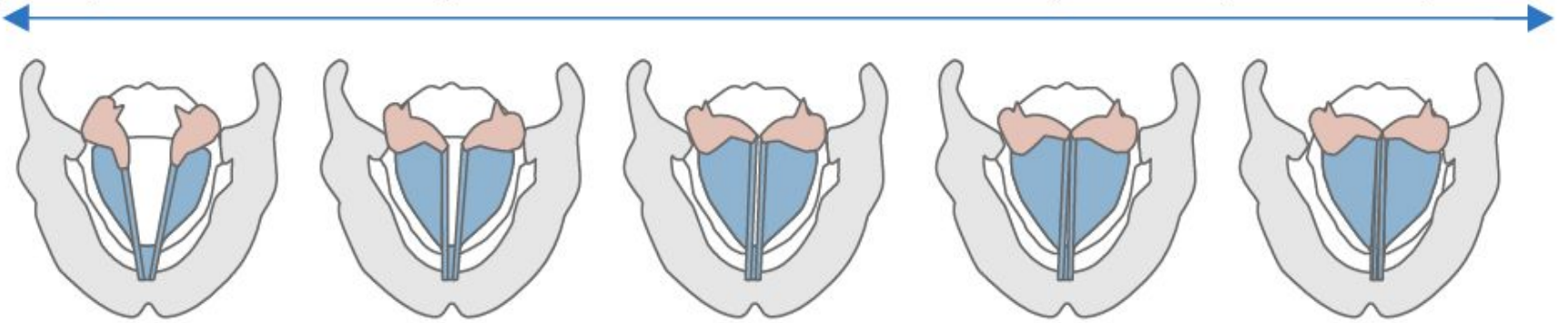
Qualidade da voz percebida

- Voz soprosa: menor compressão pelos músculos e cartilagens da laringe, maior passagem de ar, menor nível de esforço (1).
- Voz sussurrada: tensão de adução muito baixa, forças mediais e longitudinais relativamente altas (2).
- Voz rangente: tensão de adução muito alta, com baixas forças longitudinais, o que causam o espessamento das pregas vocais (3).
- Voz ríspida: tensão muito alta das pregas vocais (principalmente, devido às forças longitudinais e adutoras da laringe), com aproximação extrema entre elas (4).

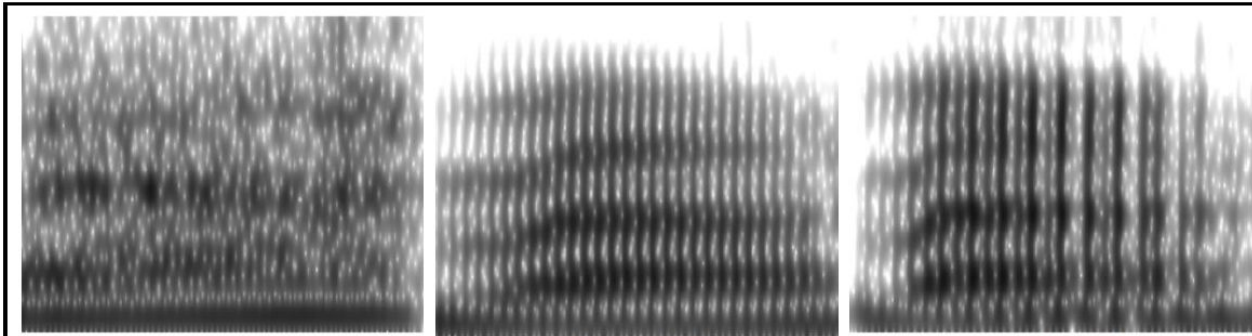


Qualidade da voz percebida

(spread) voiceless breathy voiced modal voiced creaky voiced (constricted) voiceless



Fonte: Wright *et al*, 2019



Breathy [æ]

Modal [æ]

Creaky [æ]

Fonte: Wright *et al*, 2019

Qualidade da voz percebida

Airstream Types

Ǝ	œsophageal speech	I	electrolarynx speech
Ю	tracheo-œsophageal speech	↓	pulmonic ingressive speech

Phonation types

V	modal voice	F	falsetto
W	whisper	C	creak
Ṽ	whispery voice (murmur)	Ṽ	creaky voice
Ṽ̤	breathy voice	Ç	whispery creak
V!	harsh voice	V!!	ventricular phonation
Ṽ!!	diplophonia	Ṽ!!	whispery ventricular phonation
Ṽ̤	anterior or pressed phonation	W	posterior whisper

Qualidade da voz percebida



- No *Praat*, a lista de comandos **Pulses**, dos arquivos de TextGrid, permitem a exibição de informações sobre a voz (e sobre a qualidade dela).
- **Pulses > Voice Report** mostra:
 - a mediana, a média e o desvio-padrão do *pitch*, além dos valores máximo e mínimo de pitch (em Hz);
 - o número de pulsos glotais, o número de períodos glotais, a média dos períodos glotais e o desvio-padrão do período;
 - a fração de porções desvozeadas, o número e a taxa de quebras da voz;
 - taxas de *jitter*;
 - taxas de *shimmer*;
 - harmonicidade de partes vozeadas.

Atividade de fixação - Parte 2



Exercício

1. Grave a súplica: “Você pode ir comigo?” no *Praat* em três qualidades de voz:
 - a. voz modal;
 - b. voz soprosa;
 - c. *creaky voice*.
2. Peça o relatório de voz (*Voice report*) no TextGrid e compare as taxas de *jitter* e *shimmer* para cada tipo de áudio.

Conclusão



Ao longo deste terceiro dia:

- ❑ estudamos noções básicas de Prosódia;
- ❑ identificamos e contrastamos diferentes tipos de unidades prosódicas;
- ❑ compreendemos a definição e alguns parâmetros de qualidade de voz;
- ❑ aperfeiçoamos nossas anotações no ***Praat***;
- ❑ treinamos a análise dos parâmetros: **Pitch**, **Intensity** e **Pulses** no ***Praat***.

Obrigada!
Até amanhã