

Geração de prosódia para o português brasileiro em sistemas text-to-speech

Felipe Cortez de Sá

Universidade Federal do Rio Grande do Norte

Junho de 2018

1 Introdução

2 Prosódia

3 TTS

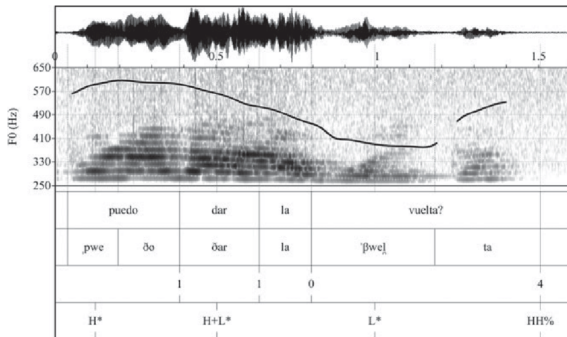
4 Implementação

- *Voice User Interfaces*
 - *Apple - Siri*
 - *Google Assistant*
 - *Microsoft - Cortana*
 - *Amazon - Alexa*
- *Acessibilidade*

- O que é prosódia?
 - pros, "verso"
 - odé, "canto"

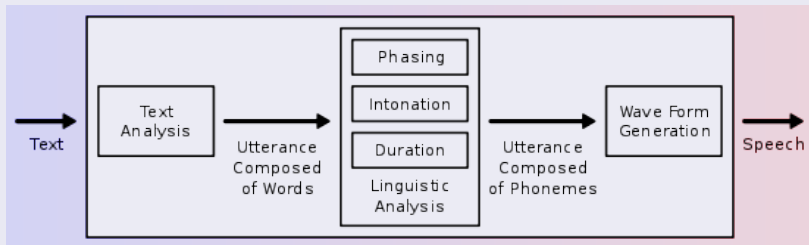
■ O que é prosódia?

- pros, "verso"
- odé, "canto"



- Afetiva (pura)
- Suprasegmental
- Aumentativa

Arquitetura genérica



- Prosódia neutra
- Prosódia aprendida
- SSML – *Speech Synthesis Markup Language*

- Comerciais

- CereProc
- Ivona (Amazon)
- Watson (IBM)

- Open-source

- MaryTTS
- espeak-ng
- Merlin

- Dífonos (MBROLA)
- Unit selection (MaryTTS)
- Síntese paramétrica (espeak-ng)
- DNN (Merlin)

- ToBI (Tone Breaks and Indices)
 - Usado por análises feitas por Moraes para pt-br
 - Usado pelo grupo de pesquisas de TTS da CMU / UoE
 - Pierrehumbert (modelo *Autosegmental-Metrical*)
- INTSINT
- IPO

- Python 3
- espeak-ng (front-end)
- MBROLA (back-end)
- Módulo gráfico

- Usar *Natural Language Understanding* para adivinhar prosódia
- Gerar prosódia a partir de marcação SSML