



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
DEPARTAMENTO DE INFORMÁTICA E MATEMÁTICA APLICADA
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO



Um gerador de prosódia para o português brasileiro

Felipe Cortez de Sá

Natal-RN
Julho de 2018

Felipe Cortez de Sá

Um gerador de prosódia para o português brasileiro

Monografia de Graduação apresentada ao Departamento de Informática e Matemática Aplicada do Centro de Ciências Exatas e da Terra da Universidade Federal do Rio Grande do Norte como requisito parcial para a obtenção do grau de bacharel em Ciência da Computação.

Carlos Augusto Prolo

Nome e titulação do(a) professor(a) orientador(a)

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE – UFRN
DEPARTAMENTO DE INFORMÁTICA E MATEMÁTICA APLICADA – DIMAP

Natal-RN

Junho de 2018

Monografia de Graduação sob o título *Um gerador de prosódia para o português brasileiro* apresentada por Felipe Cortez de Sá e aceita pelo Departamento de Informática e Matemática Aplicada do Centro de Ciências Exatas e da Terra da Universidade Federal do Rio Grande do Norte, sendo aprovada por todos os membros da banca examinadora abaixo especificada:

Dr. Carlos Augusto Prolo
Orientador(a)
Departamento de Informática e Matemática Aplicada
Universidade Federal do Rio Grande do Norte

Dr. Antônio Carlos Gay Thomé
Departamento de Informática e Matemática Aplicada
Universidade Federal do Rio Grande do Norte

The Third Man
Departamento
Universidade

Natal-RN, data de aprovação (por extenso).

Dedicado a várias pessoas

Agradecimentos

Obrigado várias pessoas

Some few people are born without any sense of time. As consequence, their sense of place becomes heightened to an excruciating degree. They lie in tall grass and are questioned by poets and painters from all over the world. These time-deaf are beseeched to describe the precise placement of trees in the spring, the shape of snow on the Alps, the angle of sun on a church, the position of rivers, the location of moss, the pattern of birds in a flock. Yet the time-deaf are unable to speak what they know. For speech needs a sequence of words, spoken in time.

Alan Lightman, *Einstein's Dreams*

Título do trabalho

Autor: Nome do aluno

Orientador(a): Titulação e nome do(a) orientador(a)

RESUMO

O resumo deve apresentar de forma concisa os pontos relevantes de um texto, fornecendo uma visão rápida e clara do conteúdo e das conclusões do trabalho. O texto, redigido na forma impessoal do verbo, é constituído de uma seqüência de frases concisas e objetivas e não de uma simples enumeração de tópicos, não ultrapassando 500 palavras, seguido, logo abaixo, das palavras representativas do conteúdo do trabalho, isto é, palavras-chave e/ou descritores. Por fim, deve-se evitar, na redação do resumo, o uso de parágrafos (em geral resumos são escritos em parágrafo único), bem como de fórmulas, equações, diagramas e símbolos, optando-se, quando necessário, pela transcrição na forma extensa, além de não incluir citações bibliográficas.

Palavras-chave: Palavra-chave 1, Palavra-chave 2, Palavra-chave 3.

A Prosody Model for Brazilian Portuguese

Author: Felipe Cortez de Sá

Advisor: Dr. Prolo

ABSTRACT

This is an abstract

Keywords: Prosody, Computational phonology, Machine Learning.

Lista de figuras

1	Diagrama	p. 19
---	--------------------	-------

Lista de tabelas

Lista de abreviaturas e siglas

Lista de símbolos

Sumário

1	Introdução	p. 14
1.1	Organização do trabalho	p. 14
2	Capítulo 2	p. 15
2.1	TTS	p. 15
2.1.1	Breve história?	p. 15
2.1.2	Fonemas e fones	p. 15
2.1.3	Abordagens	p. 15
2.1.3.1	Difones	p. 15
2.1.3.2	Unit selection	p. 15
2.1.3.3	HMM	p. 15
2.1.3.4	DNN	p. 15
2.1.4	TTS em português	p. 15
2.2	Prosódia	p. 16
2.2.1	Tipos de prosódia	p. 16
2.2.1.1	Aumentativa	p. 16
2.2.1.2	Suprasegmental	p. 16
2.2.1.3	Afetiva	p. 16
2.2.2	Elementos	p. 16
2.2.3	Prosódia como elemento extra-textual	p. 16
2.2.4	Prosódia no português brasileiro	p. 16
2.2.5	Modelos de prosódia	p. 16

2.2.6	Trabalhos semelhantes?	p. 16
3	Capítulo 3	p. 17
4	Capítulo 4	p. 18
4.1	Implementação	p. 18
4.1.1	espeak-ng	p. 18
4.1.2	MBROLA	p. 18
4.1.2.1	Formato	p. 18
4.1.3	Arquitetura	p. 18
4.1.4	Módulo de prosódia	p. 18
4.1.4.1	Sintaxe	p. 18
4.1.4.2	Editor gráfico	p. 19
4.1.5	Sintaxe prosódica	p. 19
5	Capítulo 5	p. 20
5.1	Resultados	p. 20
5.1.1	Metodologia	p. 20
5.1.2	MOS	p. 20
5.2	Trabalhos futuros	p. 20
6	Considerações finais	p. 21
	Referências	p. 22
	Apêndice A – Primeiro apêndice	p. 23
	Anexo A – Primeiro anexo	p. 24

1 Introdução

1.1 Organização do trabalho

2 Capítulo 2

2.1 TTS

2.1.1 Breve história?

Modelos físicos, Bell Labs,

2.1.2 Fonemas e fones

2.1.3 Abordagens

2.1.3.1 Difones

2.1.3.2 Unit selection

2.1.3.3 HMM

2.1.3.4 DNN

Resultados realistas, mas não há como controlar parâmetros

2.1.4 TTS em português

LianeTTS (MBROLA), HMM-based (Maia et al), MaryTTS (FalaBrasil)

2.2 Prosódia

2.2.1 Tipos de prosódia

2.2.1.1 Aumentativa

2.2.1.2 Suprasegmental

2.2.1.3 Afetiva

2.2.2 Elementos

Intonational tune Downdrift Microprosódia

2.2.3 Prosódia como elemento extra-textual

Justifica abordagem do trabalho: considerando o texto como sequência de palavras, é difícil determinar prosódia afetiva. Gerar a prosódia certa é uma questão de Natural Language Understanding, isto é, é preciso entender o texto para gerar os contornos melódicos afetivos.

2.2.4 Prosódia no português brasileiro

Trabalhos de Moraes, Tenani, ...

2.2.5 Modelos de prosódia

British school, autosegmental metrical, Fujisaki, Tilt, INTSINT “The AM model is phonological, the INTSINT model phonetic and the Fujisaki and Tilt models acoustic.” INTSINT: IPA para prosódia (mais ou menos o que eu quero fazer, mas INTSINT é para análise) ref Moraes, Intonation Systems (20 languages) AM: Pierrehumbert, Moraes (pitch analysis by synthesis)

2.2.6 Trabalhos semelhantes?

3 Capítulo 3

4 Capítulo 4

4.1 Implementação

4.1.1 `espeak-ng`

Normalização de texto (citar o Martin Jurafsky) Grafema-fonema (G2P)

4.1.2 MBROLA

Baseado no algoritmo PSOLA (Pitch Synchronous Overlap and Add) (descrever PSOLA) Fonemas/fones simplificados, determinados pelo autor de uma voz, gerando saída com voz sintetizada. Síntese por dífonos.

4.1.2.1 Formato

Pausas ou fones!!! e duração seguido por um ou mais pares de porcentagem e frequência (Hz)

4.1.3 Arquitetura

Diagrama aqui

4.1.4 Módulo de prosódia

4.1.4.1 Sintaxe

O programa foi codificado em Python em sua versão 3.6.

4.1.4.2 Editor gráfico

Editor gráfico com Qt (ou HTML, CSS e JavaScript, não decidi ainda)

4.1.5 Sintaxe prosódica

A solução para melhorar a geração prosódica foi adicionar marcações à linguagem natural, denotando o contorno de acordo com o modelo autosegmental metrical. Parecido com SSML (Speech Synthesis Markup Language)

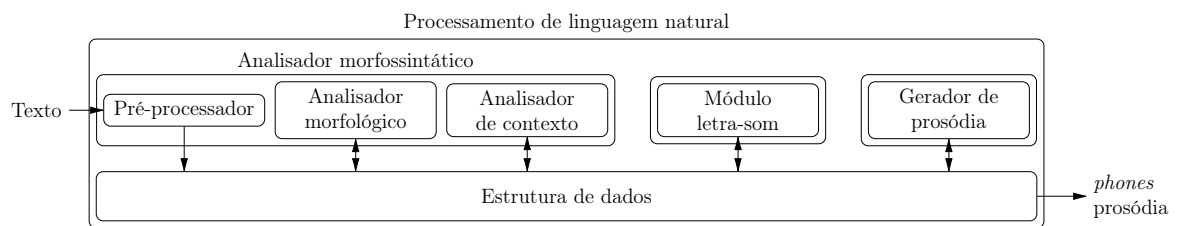


Figura 1: Diagrama

5 Capítulo 5

5.1 Resultados

5.1.1 Metodologia

5.1.2 MOS

5.2 Trabalhos futuros

6 Considerações finais

As considerações finais formam a parte final (fechamento) do texto, sendo dito de forma resumida (1) o que foi desenvolvido no presente trabalho e quais os resultados do mesmo, (2) o que se pôde concluir após o desenvolvimento bem como as principais contribuições do trabalho, e (3) perspectivas para o desenvolvimento de trabalhos futuros. O texto referente às considerações finais do autor deve salientar a extensão e os resultados da contribuição do trabalho e os argumentos utilizados estar baseados em dados comprovados e fundamentados nos resultados e na discussão do texto, contendo deduções lógicas correspondentes aos objetivos do trabalho, propostos inicialmente.

Referências

APÊNDICE A – Primeiro apêndice

Os apêndices são textos ou documentos elaborados pelo autor, a fim de complementar sua argumentação, sem prejuízo da unidade nuclear do trabalho.

ANEXO A – Primeiro anexo

Os anexos são textos ou documentos não elaborado pelo autor, que servem de fundamentação, comprovação e ilustração.