



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**

## **Título do trabalho**

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.

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

Mês (por extenso) 2018

Monografia de Graduação sob o título *Título da monografia* apresentada por Nome do aluno 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:

---

Titulação e nome do(a) orientador(a)

Orientador(a)  
Departamento  
Universidade

---

Titulação e nome do membro da banca examinadora

Co-orientador(a), se houver  
Departamento  
Universidade

---

Titulação e nome do membro da banca examinadora

Departamento  
Universidade

---

Titulação e nome do membro da banca examinadora

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
2	Teste de uma figura em formato .png . . . . .	p. 20

# Lista de tabelas

1	Tabela sem sentido . . . . .	p. 22
---	------------------------------	-------

# Lista de abreviaturas e siglas

UFRN – Universidade Federal do Rio Grande do Norte

DIMAp – Departamento de Informática e Matemática Aplicada

# Lista de símbolos

$\lambda$  (algum símbolo)

# Sumário

<b>1</b>	<b>Introdução</b>	p. 15
1.1	Organização do trabalho . . . . .	p. 15
<b>2</b>	<b>Capítulo 2</b>	p. 16
2.1	TTS . . . . .	p. 16
2.1.1	Breve história? . . . . .	p. 16
2.1.2	Fonemas e fones . . . . .	p. 16
2.1.3	Abordagens . . . . .	p. 16
2.1.3.1	Difones . . . . .	p. 16
2.1.3.2	Unit selection . . . . .	p. 16
2.1.3.3	HMM . . . . .	p. 16
2.1.3.4	DNN . . . . .	p. 16
2.1.4	TTS em português . . . . .	p. 16
2.2	Prosódia . . . . .	p. 17
2.2.1	Tipos de prosódia . . . . .	p. 17
2.2.1.1	Aumentativa . . . . .	p. 17
2.2.1.2	Suprasegmental . . . . .	p. 17
2.2.1.3	Afetiva . . . . .	p. 17
2.2.2	Elementos . . . . .	p. 17
2.2.3	Prosódia como elemento extra-textual . . . . .	p. 17
2.2.4	Prosódia no português brasileiro . . . . .	p. 17
2.2.5	Modelos de prosódia . . . . .	p. 17

2.2.6	Trabalhos semelhantes? . . . . .	p. 18
2.3	Implementação . . . . .	p. 18
2.3.1	espeak-ng . . . . .	p. 18
2.3.2	MBROLA . . . . .	p. 18
2.3.2.1	Formato . . . . .	p. 18
2.3.3	Arquitetura . . . . .	p. 18
2.3.4	Módulo de prosódia . . . . .	p. 18
2.3.4.1	Sintaxe . . . . .	p. 18
2.3.4.2	Editor gráfico . . . . .	p. 18
2.3.5	Sintaxe prosódica . . . . .	p. 18
2.4	Resultados . . . . .	p. 19
2.4.1	Metodologia . . . . .	p. 19
2.4.2	MOS . . . . .	p. 19
2.5	Trabalhos futuros . . . . .	p. 19
2.6	Seção 1 . . . . .	p. 20
2.7	Seção 2 . . . . .	p. 20
2.8	Seção 3 . . . . .	p. 20
2.9	Seção 4 . . . . .	p. 20
<b>3</b>	<b>Capítulo 3</b>	p. 21
3.1	Seção 1 . . . . .	p. 22
3.2	Seção 2 . . . . .	p. 22
3.2.1	Subseção 2.1 . . . . .	p. 22
3.2.2	Subseção 2.2 . . . . .	p. 23
3.3	Seção 3 . . . . .	p. 23
<b>4</b>	<b>Capítulo 4</b>	p. 24

4.1	Seção 1 . . . . .	p. 24
4.2	Seção 2 . . . . .	p. 24
<b>5</b>	<b>Capítulo 5</b>	p. 25
5.1	Seção 1 . . . . .	p. 25
5.2	Seção 2 . . . . .	p. 25
5.2.1	Subseção 5.1 . . . . .	p. 25
5.2.2	Subseção 5.2 . . . . .	p. 26
5.3	Seção 3 . . . . .	p. 26
<b>6</b>	<b>Considerações finais</b>	p. 27
	<b>Referências</b>	p. 28
	<b>Apêndice A – Primeiro apêndice</b>	p. 29
	<b>Anexo A – Primeiro anexo</b>	p. 30

# 1 Introdução

A introdução é a parte inicial do texto e que possibilita uma visão geral de todo o trabalho, devendo constar a delimitação do assunto tratado, objetivos da pesquisa, motivação para o desenvolvimento da mesma e outros elementos necessários para situar o tema do trabalho.

## 1.1 Organização do trabalho

Nesta seção deve ser apresentado como está organizado o trabalho, sendo descrito, portanto, do que trata cada capítulo.



## 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?

## 2.3 Implementação

### 2.3.1 espeak-ng

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

### 2.3.2 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.

#### 2.3.2.1 Formato

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

### 2.3.3 Arquitetura

Diagrama aqui

### 2.3.4 Módulo de prosódia

#### 2.3.4.1 Sintaxe

O programa foi codificado em Python em sua versão 3.6.

#### 2.3.4.2 Editor gráfico

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

### 2.3.5 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)

## 2.4 Resultados

### 2.4.1 Metodologia

Pesquisa feita online, arquivos de áudio, soluções comerciais, soluções open-source

### 2.4.2 MOS

Comparar MOS (Mean Opinion Score) com outras soluções

## 2.5 Trabalhos futuros

Usar técnicas de Natural Language Understanding para gerar prosódia utilizando notação desenvolvida neste trabalho

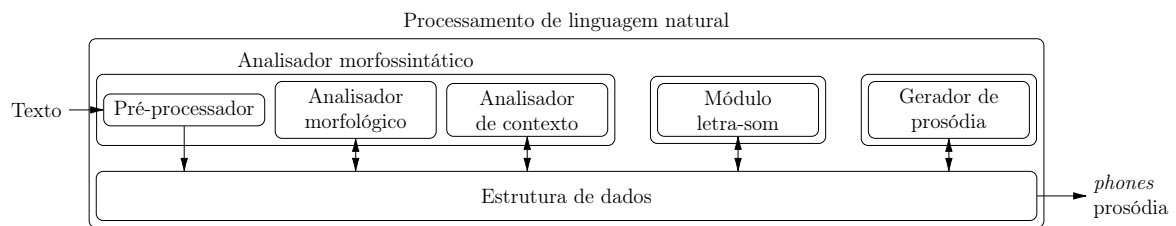


Figura 1: Diagrama

Este é o primeiro capítulo da parte central do trabalho, isto é, o desenvolvimento, a parte mais extensa de todo o trabalho. Geralmente o desenvolvimento é dividido em capítulos, cada um com subseções e subseções, cujo tamanho e número de divisões variam em função da natureza do conteúdo do trabalho.

Em geral, a parte de desenvolvimento é subdividida em quatro subpartes:

- *contextualização ou definição do problema* – consiste em descrever a situação ou o contexto geral referente ao assunto em questão, devem constar informações atualizadas visando a proporcionar maior consistência ao trabalho;
- *referencial ou embasamento teórico* – texto no qual se deve apresentar os aspectos teóricos, isto é, os conceitos utilizados e a definição dos mesmos; nesta parte faz-se a revisão de literatura sobre o assunto, resumindo-se os resultados de estudos feitos por outros autores, cujas obras citadas e consultadas devem constar nas referências;

- *metodologia do trabalho ou procedimentos metodológicos* – deve constar o instrumental, os métodos e as técnicas aplicados para a elaboração do trabalho;
- *resultados* – devem ser apresentados, de forma objetiva, precisa e clara, tanto os resultados positivos quanto os negativos que foram obtidos com o desenvolvimento do trabalho, sendo feita uma discussão que consiste na avaliação circunstanciada, na qual se estabelecem relações, deduções e generalizações.

É recomendável que o número total de páginas referente à parte de desenvolvimento não ultrapasse 60 (sessenta) páginas.

## 2.6 Seção 1

Teste de figura:

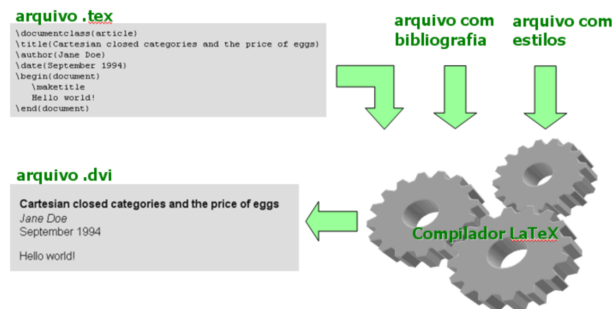


Figura 2: Teste de uma figura em formato .png

## 2.7 Seção 2

Referenciamento da figura inserida na seção anterior: 2.6

## 2.8 Seção 3

Seção 3

## 2.9 Seção 4

Seção 4

## 3 Capítulo 3

Algumas regras devem ser observadas na redação da monografia:

1. ser claro, preciso, direto, objetivo e conciso, utilizando frases curtas e evitando ordens inversas desnecessárias;
2. construir períodos com no máximo duas ou três linhas, bem como parágrafos com cinco linhas cheias, em média, e no máximo oito (ou seja, não construir parágrafos e períodos muito longos, pois isso cansa o(s) leitor(es) e pode fazer com que ele(s) percam a linha de raciocínio desenvolvida);
3. a simplicidade deve ser condição essencial do texto; a simplicidade do texto não implica necessariamente repetição de formas e frases desgastadas, uso exagerado de voz passiva (como *será iniciado*, *será realizado*), pobreza vocabular etc. Com palavras conhecidas de todos, é possível escrever de maneira original e criativa e produzir frases elegantes, variadas, fluentes e bem alinhavadas;
4. adotar como norma a ordem direta, por ser aquela que conduz mais facilmente o leitor à essência do texto, dispensando detalhes irrelevantes e indo diretamente ao que interessa, sem rodeios (verborragias);
5. não começar períodos ou parágrafos seguidos com a mesma palavra, nem usar repetidamente a mesma estrutura de frase;
6. desprezar as longas descrições e relatar o fato no menor número possível de palavras;
7. recorrer aos termos técnicos somente quando absolutamente indispensáveis e nesse caso colocar o seu significado entre parênteses (ou seja, não se deve admitir que todos os que lerão o trabalho já dispõem de algum conhecimento desenvolvido no mesmo);
8. dispensar palavras e formas empoladas ou rebuscadas, que tentem transmitir ao leitor mera idéia de erudição;

9. não perder de vista o universo vocabular do leitor, adotando a seguinte regra prática:  
*nunca escrever o que não se diria*;
10. termos coloquiais ou de gíria devem ser usados com *extrema* parcimônia (ou mesmo nem serem utilizados) e apenas em casos muito especiais, para não darem ao leitor a idéia de vulgaridade e descaracterizar o trabalho;
11. ser rigoroso na escolha das palavras do texto, desconfiando dos sinônimos perfeitos ou de termos que sirvam para todas as ocasiões; em geral, há uma palavra para definir uma situação;
12. encadear o assunto de maneira suave e harmoniosa, evitando a criação de um texto onde os parágrafos se sucedem uns aos outros como compartimentos estanques, sem nenhuma fluência entre si;
13. ter um extremo cuidado durante a redação do texto, principalmente com relação às regras gramaticais e ortográficas da língua; geralmente todo o texto é escrito na forma impessoal do verbo, não se utilizando, portanto, de termos em primeira pessoa, seja do plural ou do singular.

## 3.1 Seção 1

Teste de uma tabela:

Tabela 1: Tabela sem sentido

Titulo Coluna 1	Título Coluna 2
X	Y
X	W

## 3.2 Seção 2

Seção 2

### 3.2.1 Subseção 2.1

Referência à tabela definida no início: 3.1

### **3.2.2 Subseção 2.2**

Subsection 2.2

## **3.3 Seção 3**

Seção 3



## 4 Capítulo 4

### 4.1 Seção 1

Teste para símbolo

$\lambda$

### 4.2 Seção 2

Teste para abreviatura

UFRN

DIMAp

## 5 Capítulo 5

### 5.1 Seção 1

Seção 1

### 5.2 Seção 2

Alguns exemplos de citação:

(APPLE, 2017) é um programa.

Na tese de Doutorado de Paquete (PAQUETE, 2005), discute-se sobre algoritmos de busca local estocásticos aplicados a problemas de Otimização Combinatória considerando múltiplos objetivos. Por sua vez, o trabalho de (KNOWLES; CORNE; FLEISCHER, 2003), publicado nos anais do IEEE CEC de 2003, mostra uma técnica de arquivamento também empregada no desenvolvimento de algoritmos evolucionários multi-objetivo, trabalho esse posteriormente estendido para um capítulo de livro dos mesmos autores (KNOWLES; CORNE, 2004). Por fim, no relatório técnico de Jaskiewicz (1998), fala-se sobre um algoritmo genético híbrido para problemas multi-critério, enquanto no artigo de jornal de Lopez *et al.* (LÓPEZ-IBÁÑEZ; PAQUETE; STÜTZLE, 2006) trata-se do *trade-off* entre algoritmos genéticos e metodologias de busca local, também aplicados no contexto multi-critério e relacionado de alguma forma ao trabalho de Jaskiewicz (1998).

Outros exemplos relacionados encontram-se em (SILBERSCHATZ; KORTH; SUDARSHAN, 2002) (livro), (TURAU, 2001) (referência da Web) e (AGRA, 2004) (dissertação de Mestrado).

#### 5.2.1 Subseção 5.1

Subseção 5.1

### 5.2.2 Subseção 5.2

Subsection 5.2

## 5.3 Seção 3

Seção 3

## 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

- AGRA, A. Projeto de Diplomação, *Implementação de uma Proposta para Atualização de Bancos de Dados através de Visões*. Porto Alegre, RS, Brasil: [s.n.], jul. 2004.
- APPLE. *Siri*. 2017. Acesso em 2017-09-25. Disponível em: <<https://www.apple.com/ios/siri/>>.
- JASZKIEWICZ, A. *Genetic local search for multiple objective combinatorial optimization*. [S.l.], 1998.
- KNOWLES, J.; CORNE, D. Metaheuristics for multiobjective optimisation. In: \_\_\_\_\_. [S.l.]: Springer, 2004. (Lecture Notes in Economics and Mathematical Systems, v. 535), cap. Bounded Pareto Archiving: Theory and Practice, p. 39–64.
- KNOWLES, J. D.; CORNE, D. W.; FLEISCHER, M. Bounded archiving using the Lebesgue measure. In: *Proceedings of the IEEE Congress on Evolutionary Computation*. [S.l.]: IEEE Press, 2003. p. 2490–2497.
- LÓPEZ-IBÁÑEZ, M.; PAQUETE, L.; STÜTZLE, T. Hybrid population-based algorithms for the bi-objective quadratic assignment problem. *Journal of Mathematical Modelling and Algorithms*, v. 5, n. 1, p. 111–137, 2006.
- PAQUETE, L. *Stochastic Local Search Algorithms for Multiobjective Combinatorial Optimization Problems: Methods and Analysis*. Tese (Doutorado) — Technische Universität Darmstadt, 2005.
- SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. *Database system concepts*. 4th. ed. Boston: McGraw-Hill, 2002.
- TURAU, V. *DB2XML 1.4: Transforming relational databases into XML documents*. out. 2001. Out., 2001. Disponível em: <<http://www.informatik.fh-wiesbaden.de/~turau/DB2XML/index.html>>. Acesso em Abril 9, 2004.

## 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.