



ALGORITMOS DE PROCESSAMENTO DA LINGUAGEM E SÍNTESE DE
VOZ COM EMOÇÕES APLICADOS A UM CONVERSOR TEXTO-FALA
BASEADO EM HMM

Denilson da Cruz da Silva

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Fernando Gil Vianna Resende Jr.

Rio de Janeiro
Março de 2011

ALGORITMOS DE PROCESSAMENTO DA LINGUAGEM E SÍNTESE DE
VOZ COM EMOÇÕES APLICADOS A UM CONVERSOR TEXTO-FALA
BASEADO EM HMM

Denilson da Cruz da Silva

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

Prof. Fernando Gil Vianna Resende Jr., Ph.D.

Prof. Luis Pereira Calôba, Dr.Ing.

Prof^ª. Mariane Rembold Petraglia, Ph.D.

Prof. Abraham Alcaim, Ph.D.

Prof. Amaro Azevedo de Lima, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2011

Silva, Denilson da Cruz da

Algoritmos de processamento da linguagem e síntese de voz com emoções aplicados a um conversor texto-fala baseado em HMM/Denilson da Cruz da Silva. – Rio de Janeiro: UFRJ/COPPE, 2011.

XVI, 163 p.: il.; 29,7cm.

Orientador: Fernando Gil Vianna Resende Jr.

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2011.

Referências Bibliográficas: p. 92 – 102.

1. Síntese de voz com emoções. 2. Processamento da linguagem. 3. Conversão texto-fala. 4. Processamento de voz. I. Resende Jr., Fernando Gil Vianna. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*“O cavalo se prepara para a
batalha, mas do Senhor vem a
vitória”
Provérbios 21.31*

Agradecimentos

Eu gostaria de agradecer à minha família, colegas e amigos que sempre estiveram ao meu lado, durante o desenvolvimento deste trabalho. À minha esposa Silvia, e filhas, Silvia Valéria, Denise Vitória e Desirêe, pela compreensão dispensada e pelo apoio irrestrito.

Quero agradecer ao meu orientador, Prof. Fernando Gil, por ter confiado a mim este projeto, tão importante para a comunidade científica no escopo das tecnologias de voz.

Gostaria de agradecer aos Professores: Mariane Rembold Petraglia, Luis Pereira Calôba, Amaro Azevedo de Lima e Abraham Alcaim, pela participação na banca de defesa de tese e pelas importantes contribuições.

Agradeço aos meus colegas do LPS, que de alguma forma contribuíram para a realização deste trabalho e pelos bons momentos passados juntos. A todos os funcionários, alunos e professores do PEE e da COPPE.

Agradeço, também ao Wagner Pedretti, que por longos meses trabalhou conosco no projeto da nova base de dados de voz, compartilhando seu conhecimento e nos ajudando de forma muito significativa.

Agradeço a todos os locutores que participaram da gravação da nova base, bem como ao órgão financiador, CNPq, e a todos os validadores, que contribuíram *online* com este projeto.

Agradeço também a todos os companheiros do Comando da Aeronáutica que contribuíram para a realização deste trabalho.

E, finalmente, quero expressar o meu principal agradecimento, que não poderia ser direcionado a outro, senão ao meu grandioso Deus, que me deu saúde, força e fé para prosseguir sempre, principalmente nos momentos de maior dificuldade.

“Porque eu conheço que o Senhor é grande e que o nosso Senhor está acima de todos os deuses.” Salmos 135.5

Obrigado a todos!

Denilson da Cruz da Silva

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

ALGORITMOS DE PROCESSAMENTO DA LINGUAGEM E SÍNTESE DE VOZ COM EMOÇÕES APLICADOS A UM CONVERSOR TEXTO-FALA BASEADO EM HMM

Denilson da Cruz da Silva

Março/2011

Orientador: Fernando Gil Vianna Resende Jr.

Programa: Engenharia Elétrica

Esta tese apresenta um conjunto de algoritmos para o processamento da linguagem e síntese de voz com emoções para o idioma português falado no Brasil.

É proposto um conjunto de algoritmos baseados em regras linguísticas para a desambiguação de homógrafos heterófonos. O método proposto é composto de uma análise morfossintática para homógrafos pertencentes a diferentes classes gramaticais, e uma análise semântica para homógrafos de mesma classe gramatical. O conjunto de algoritmos foi desenvolvido para a desambiguação de 111 pares de homógrafos organizados em 23 tipos. Os algoritmos foram implementados e testados com textos de três naturezas, apresentando taxa global de acerto de 99,00%.

Um algoritmo que determina a sílaba tônica, baseado em 19 regras linguísticas, também é apresentado. O algoritmo proposto foi implementado e testado obtendo uma taxa de acerto de 98,90%.

A separação das sílabas é realizada através de um algoritmo baseado em 25 regras linguísticas. O resultado de experimentos computacionais foi uma taxa de acerto de 99,29%.

O algoritmo proposto para realizar a conversão grafema-fone (G2P - *Grapheme-Phone*) foi implementado a partir de 132 regras linguísticas e o resultado foi uma taxa de acerto de 98,57%.

Uma nova base de dados de voz para síntese e reconhecimento foi gravada. Nesta base foram incluídas frases lidas com as seguintes emoções: alegria, tristeza, raiva e neutra. Foi realizado um teste para selecionar as frases a serem utilizadas no treinamento do sistema de síntese com emoções, comparando a emoção pretendida pelo locutor com a emoção percebida por um conjunto de avaliadores. Em 81,8% dos casos a emoção pretendida foi igual a percebida. A implementação do sistema de síntese com emoções foi realizada a partir das emoções percebidas pelos avaliadores.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

ALGORITHMS FOR LANGUAGE PROCESSING AND EMOTIONAL SPEECH SYNTHESIS APPLIED IN A HMM-BASED TEXT-TO-SPEECH CONVERTER

Denilson da Cruz da Silva

March/2011

Advisor: Fernando Gil Vianna Resende Jr.

Department: Electrical Engineering

This thesis presents an algorithm set for language processing and emotional speech synthesis for Brazilian Portuguese.

It's proposed an algorithm set based on linguistic rules for homograph disambiguation. The proposed method consists of a morphosyntactic analysis, for homographs that belong to different part-of-speech (POS), and a semantic analysis, for homographs that belong to the same POS. The algorithms were developed for disambiguation of 111 homograph pairs organized into 23 types. The algorithms were implemented and tested with three types of texts, reaching a overall correctness rate of 99.00%.

An algorithm that determines the stress vowel, based on 19 linguistic rules, is also presented. The proposed algorithm was implemented and tested reaching a correctness rate of 98.90%.

The syllabification is accomplished through an algorithm based on 25 linguistic rules. The result of computer experiments was a correctness rate of 99.29%.

The proposed algorithm to perform a grapheme-phone conversion was implemented from 132 linguistic rules and the result of tests was a correctness rate of 98.57%.

A new speech database for synthesis and recognition was recorded. In this database, it was included a set of phrases read with the following emotions: happiness, sadness, anger and neutral. It was performed a test to select utterances to be used in the training of the emotional speech synthesis system, comparing the intended emotion by the speaker with the felt emotion by an evaluator group. In 81.8% of cases, the intended emotion was equal to the perceived emotion. The implementation of the emotional speech synthesis system was undertaken with the emotions perceived by evaluators.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xiii
Lista de Acrônimos	xiv
1 Introdução	1
1.1 Sistemas de Conversão Texto-fala	4
1.1.1 Processamento da Linguagem	5
1.2 Objetivos	6
1.3 Resumo dos resultados obtidos	6
1.4 Estrutura do trabalho	8
2 Algoritmo para desambiguação de homógrafos	9
2.1 Introdução	9
2.2 Consequências da ambiguidade	11
2.3 Metodologia aplicada na desambiguação	12
2.4 Bibliotecas empregadas na análise do texto	15
2.5 Conjunto de regras para desambiguação de homógrafos	18
2.5.1 Os novos homógrafos derivados do Acordo Ortográfico de 2008.	18
2.6 Testes realizados com o conjunto de regras	20
2.6.1 Base de dados utilizada	20
2.6.2 Resultados obtidos	23
2.7 Conclusões e trabalhos futuros	28
3 Algoritmo para determinação de tonicidade	30
3.1 Introdução	30
3.2 Conjunto de regras para determinação de tonicidade	32
3.3 Resultados obtidos	35
3.4 Conclusões e trabalhos futuros	36

4	Algoritmo para separação das sílabas	37
4.1	Introdução	37
4.2	Papel da sílaba na entonação	38
4.3	Conjunto de regras para separação das sílabas	38
4.4	Resultados obtidos	39
4.5	Conclusões e trabalhos futuros	42
5	Algoritmo para transcrição fonética	45
5.1	Introdução	45
5.2	Unidades acústicas utilizadas	45
5.3	Conjunto de regras para transcrição fonética	46
5.4	Resultados obtidos	48
5.5	Conclusões e trabalhos futuros	49
6	Síntese de voz e as emoções da fala	54
6.1	Introdução	54
6.2	Base de dados de voz para Síntese com emoções	56
6.3	O projeto da base de dados de voz com emoções para o PB	57
6.4	A validação cruzada das emoções	60
6.5	Conclusões e trabalhos futuros	63
7	Síntese de voz com emoções baseada em HMM no PB	65
7.1	Introdução	65
7.2	Modelos escondidos de Markov	66
7.2.1	Cálculo da probabilidade do HMM	68
7.3	Síntese de voz com emoções baseada em HMM	70
7.4	Testes realizados e resultados obtidos	74
7.4.1	Distância Mel-log espectral	75
7.4.2	Espectros gerados	76
7.5	Conclusões e trabalhos futuros	76
8	Conclusões da tese e trabalhos futuros	88
8.1	Conclusões da tese	88
8.2	Trabalhos futuros	90
	Referências Bibliográficas	92
A	Publicações	103
A.1	Artigo em Revista	103
A.2	Artigos em Conferências	103

B	Algoritmos para Treinamento do HBPTTS	105
B.1	Ferramentas necessárias	105
B.2	Preparação do módulo de treinamento	106
B.3	Segmentação automática da base de dados	109
B.3.1	Preparação dos dados de voz	112
B.3.2	Criação do Dicionário	115
B.3.3	Criação da transcrição fonética	116
B.3.4	Treinamento dos modelos acústicos	119
C	Uma nova base de dados de voz para reconhecimento e síntese.	123
C.1	1000 frases foneticamente balanceadas.	123
C.2	161 frases nas emoções: Alegria, Tristeza e Raiva.	150
C.3	85 frases interrogativas.	154
C.4	29 frases exclamativas.	157
C.5	23 Frases com onomatopéias diversas	158
C.6	Aprenda a dizer te amo	159
C.7	Sobre a gloriosa noite em que eu apitei o Fla-Flu	161

Lista de Figuras

1.1	Exemplo de um TTS baseado em HMM.	3
1.2	Composição básica de um TTS.	4
1.3	Espelho da interface gráfica do HBPTTS.	7
2.1	Localização do desambiguador de homógrafos no TTS.	11
2.2	Estrutura do desambiguador de homógrafos.	13
3.1	Determinação da tonicidade na frase “Síntese de voz no português brasileiro”.	32
5.1	Rótulos relativos a sentença “Síntese de voz com emoções para o PB”.	49
6.1	Site utilizado para validação “online”.	61
6.2	Percentual de acerto cruzado na matriz de confusão para ALEGRIA.	62
6.3	Percentual de acerto cruzado na matriz de confusão para RAIVA.	62
6.4	Percentual de acerto cruzado na matriz de confusão para TRISTEZA.	63
6.5	Percentual de acerto cruzado na matriz de confusão para NEUTRA	63
7.1	HMM ergódico ou sem restrição.	67
7.2	HMM esquerda-direita.	67
7.3	Função densidade de probabilidade gaussiana.	68
7.4	Função densidade de probabilidade Multi-Mistura.	69
7.5	Função densidade de probabilidade Multi-Fluxo.	70
7.6	Sistema de Síntese de Voz baseado em HMM.	71
7.7	Composição do vetor de observação para treino dos HMMs.	72
7.8	Transformação linear no espaço das emoções.	73
7.9	Evolução da MLSD, fixando o locutor e variando dado de adaptação	77
7.10	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (EDU)”	79
7.11	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (FMA)”	80

7.12	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MIR)”	81
7.13	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MAN)”	82
7.14	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (FER)”	83
7.15	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (EDU)”	84
7.16	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MRO)”	85
7.17	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (EDU)”	86
7.18	Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MRO)”	87
B.1	Exemplo de arquivo .uttinfo.	107

Lista de Tabelas

2.1	Conjunto de homógrafos separados por tipo.	14
2.2	Exemplos de homógrafos com diferentes classes gramaticais.	17
2.3	Exemplos de homógrafos com mesma classe gramatical.	17
2.4	Simbologia usada nos algoritmos de desambiguação.	19
2.5	Testes com o algoritmo proposto - CETEN-Folha.	27
2.6	Testes com o algoritmo proposto - Bíblia Cristã.	28
2.7	Testes com o algoritmo proposto - Dom Casmurro.	29
3.1	Regras para determinação da tonicidade (da 1 até a 8)	33
3.2	Regras para determinação da tonicidade (da 9 até a 15)	34
3.3	Regras para determinação da tonicidade (da 16 até a 19).	35
3.4	Tabela com resultados na determinação de tonicidade.	36
4.1	Símbolos usados nas regras de separação das sílabas.	39
4.2	Casos e ações tomados pelo algoritmo.	40
4.3	Regras utilizadas na separação das sílabas (de 1 a 5).	40
4.4	Regras utilizadas na separação das sílabas (de 6 a 11).	41
4.5	Regras utilizadas na separação das sílabas (de 12 a 15).	42
4.6	Regras utilizadas na separação das sílabas (de 16 a 21).	43
4.7	Regras utilizadas na separação das sílabas (de 22 a 25).	44
4.8	Tabela com resultados na separação das sílabas.	44
5.1	Conjunto de fones representados pelo SAMPA, com alguns exemplos.	47
5.2	Símbolos usados nas regras de transcrição fonética.	48
5.3	Regras para os grafemas <a, b, c, d, e, f, g, h, i>.	50
5.4	Regras para os grafemas <j, k, l, m, n, o, p, q, r>.	51
5.5	Regras para os grafemas <s, t, u, v, w, x, y, z>.	52
5.6	Exceções para o grafema <x>.	52
5.7	Tabela com resultados do mapeamento dos grafemas.	53
6.1	Matriz de confusão no teste de validação “online”.	61

7.1	Testes realizados na Síntese com emoções.	75
-----	---	----

Lista de Acrônimos

ASR - Reconhecimento Automático da Fala (*Automatic Speech Recognition*)

ATR - *Advanced Telecommunications Research Institute International*

CMLLR - Regressão Linear forçada no sentido da Máxima Verossimilhança (*Constrained Maximum-Likelihood Linear Regression*)

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico

EM - Maximização da Expectativa (*Expectation-Maximization*)

G2P - Grafema-Fone (*Grapheme-Phone*)

GV - Variância Global (*Global Variance*)

HBPTTS - Conversor texto-fala para o Português Brasileiro baseado em HMM (*HMM-based Brazilian Portuguese text-to-speech*)

HMM - Modelo Escondido de Markov (*Hidden Markov Model*)

HMMs - Modelos Escondidos de Markov (*Hidden Markov Models*)

HTK - *Hidden Markov Model Toolkit*

HTS - *HMM-based Speech Synthesis System*

IBM - *International Business Machines*

MFCC - Coeficiente mel-cepestral ou coeficiente cepestral da mel-frequência (*Mel-frequency cepstral coefficient*)

MLF - *Master Label File*

MLLR - Regressão Linear no sentido da Máxima Verossimilhança (*Maximum-Likelihood Linear Regression*)

MLSA - Aproximação Mel Log Espectro (*Mel Log Spectrum Approximation*)

MLSD - Distância Mel Log Espectral (*Mel Log Spectral Distance*)

MSD - Distribuição de Probabilidade Multi-espaco (*Multi-Space Probability Distribution*)

NLP - Processamento da Linguagem Natural (*Natural Language Processing*)

PB - Português Brasileiro

POS - Classe gramatical (*Part-of-Speech*)

PSOLA - *Pitch-Synchronous Overlap and Add*

SAMPA - *Speech Assessment Methods Phonetic Alphabet*

SPTK - *Speech Signal Processing Toolkit*

TTS - Conversão texto-fala (*Text-to-Speech*)

Capítulo 1

Introdução

Os primeiros estudos acerca das expressões emocionais em seres humanos e animais são datados do século XIX [1]. A partir de então, pesquisadores tem acumulado crescentes conhecimentos na área e procurado abordar de forma profunda os limites das emoções humanas através da fala [2]. Após os primeiros trabalhos na relação do homem com máquinas, principalmente entre as tecnologias de voz, aqueles pesquisadores perceberam que poderiam avançar ainda mais nas suas pesquisas relativas ao processo de interação homem-máquina, ao inserir as expressões emocionais nesta relação [3].

Das tecnologias de voz pesquisadas atualmente, a síntese de voz tem sido uma das que mais tem encontrado campo para avançar, talvez pelo fato de que, na maioria dos sintetizadores de voz existentes, a qualidade da voz sintetizada ainda permitamos buscar melhorias para que o ouvinte tenha minimizada a percepção de que é uma máquina falando e não um humano. Sendo assim, para desenvolver uma máquina que “fale” da mesma forma que uma pessoa, é fundamental que os sintetizadores sejam capazes de reproduzir a voz com características, estilos e emoções variados [4].

Criar sistemas que possam transmitir emoções através da fala é o novo horizonte para os pesquisadores das tecnologias de voz e a idéia inicial, aplicada nesta expectativa, tem mais de um século, quando Darwin [1] critica o trabalho de outros pesquisadores e posiciona uma base evolutiva para as emoções do ser humano. Darwin relaciona as emoções da fala a antigas canções, considerando que quando a voz é usada sob fortes emoções, ela tende a assumir uma característica musical. Desde então, outros pesquisadores tem proposto teorias que explicam e classificam emoções e expressões emocionais sob várias perspectivas. Passando pelos anos 30, com o trabalho de Grant Fairbanks [2], até os anos 80, pesquisadores de linguística computacional e desenvolvedores de sistemas de conversão texto-fala (TTS - *text-to-speech*) começaram a entrelaçar seus trabalhos. Uma consequência disto foi o trabalho de Janet Cahn [3], em 1988. Atualmente, trabalhos sobre as emoções e a

voz têm sido um atrativo internacional, envolvendo os esforços de pesquisadores e empresas em todo o mundo (IBM, ATR, Microsoft, etc).

As características mais importantes de um sintetizador são a naturalidade e a inteligibilidade [5, 6]. A naturalidade descreve quanto o som reproduzido é semelhante ao de um humano, enquanto a inteligibilidade expressa quão fácil o som reproduzido é compreendido. Um sintetizador próximo do ideal é aquele que possui uma reprodução natural e inteligível. Desde de que os primeiros trabalhos concretos surgiram [7], essa tem sido a meta das pesquisas sobre síntese de voz, buscando otimizar, simultaneamente, naturalidade e inteligibilidade.

Os principais sintetizadores desenvolvidos recentemente são aqueles baseados na concatenação de formas de onda, onde unidades acústicas são apropriadamente selecionadas de uma ampla base de dados, sejam oriundas de palavras, sílabas, demi-sílabas¹, fones, difones, etc, apresentando uma alta qualidade na síntese em relação àqueles baseados na síntese articulatória [8, 9] e síntese de formantes [6].

Utilizar as palavras como elemento de concatenação na síntese tem a vantagem de que toda a coarticulação dentro da mesma é capturada nas próprias unidades gravadas. O problema reside no fato de que concatenar formas de onda, que representam palavras, resulta num prejuízo grande para a inteligibilidade, devido às descontinuidades de pitch e formantes nas fronteiras das palavras, além do mais, palavras gravadas isoladamente possuem uma duração menor do que aquelas pronunciadas no meio das frases, devido ao desgaste respiratório durante a pronúncia. No caso de optarmos por utilizar sílabas como unidades de concatenação, nos deparamos com a vantagem de termos as coarticulações intra-unidades preservadas, mas as coarticulações inter-unidades não são mais fracas do que as coarticulações internas, tornando mais difícil uma transição entre elas com qualidade.

O uso de demi-sílabas [10] tem a vantagem de que são necessário apenas algumas unidades na gravação, além de preservarem grupos altamente coarticulados, por se tratarem de unidades menores, entretanto, a questão da coarticulação inter-unidades também está presente. Por fim, utilizar difones na concatenação parece ser uma boa opção, já que as transições inter-unidades vão ocorrer no meio dos fones, apresentando descontinuidades mínimas, pelo fato de que esta região possui boa estabilidade espectral. Entretanto, as atuais pesquisas buscam sintetizadores com características, estilos e emoções variados [11] e a gravação para a concatenação de quaisquer destas unidades vão gerar uma forma de onda, cuja a voz sintetizada torna-se similar a voz do locutor que a gravou. Sem contar que seria necessário gravar enormes bases de dados, com todas estas variações, para alcançar a síntese que desejamos, sem degradar a qualidade. O problema é que gravações de longas bases de dados é um procedimento difícil e oneroso [4]. PSOLA (*Pitch Synchronous*

¹Metades iniciais e finais das sílabas

Overlap and Add) [12, 13, 14, 15] é um bom exemplo, pois apesar de apresentarem uma alta qualidade na síntese de voz e de serem utilizados em muitas aplicações, possuem estas desvantagens.

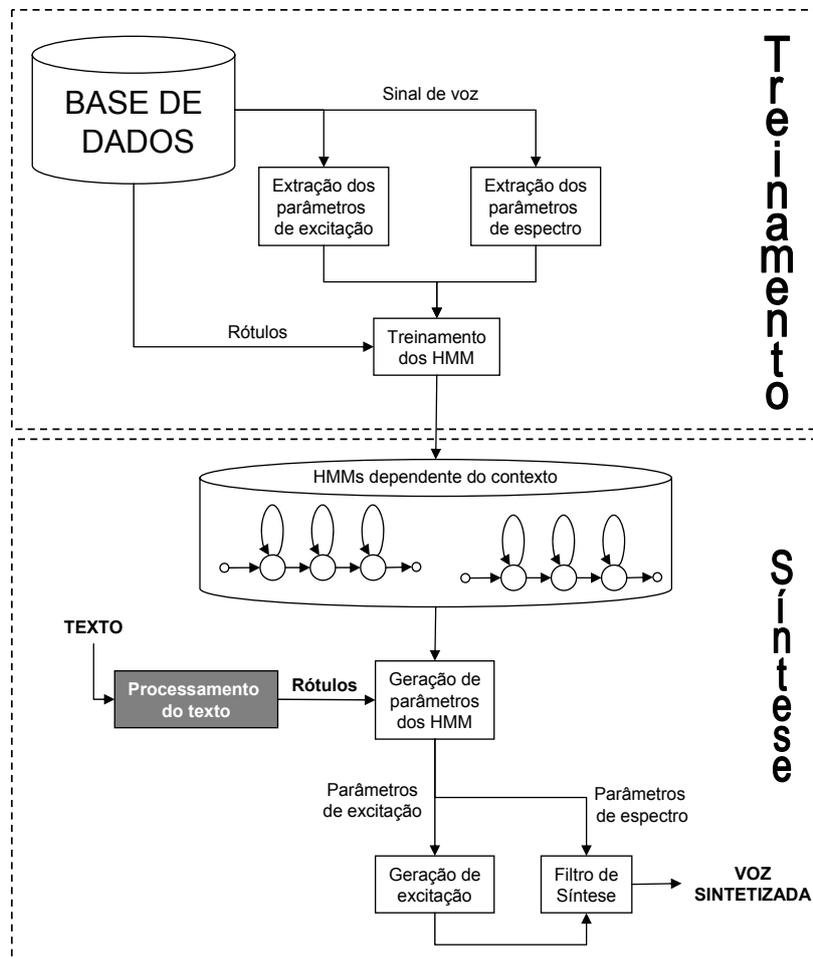


Figura 1.1: Exemplo de um TTS baseado em HMM.

Nos últimos anos, sintetizadores de voz baseados em Modelos Escondidos de Markov (HMMs - *Hidden Markov Models*) têm crescido em popularidade [16, 17, 18, 19, 20, 21, 22, 23]. Nestes sistemas, HMMs dependentes do contexto são treinados a partir da base de dados com fala natural e as formas de onda são geradas a partir destes HMMs. Estes sistemas tem a habilidade de modelar diferentes estilos sem requerer a gravação de uma base de dados muito grande. A Figura 1.1 mostra a estrutura básica de um TTS baseado em (HMM - *Hidden Markov Model*), destacando o bloco de análise e processamento do texto, que também será um dos pontos discutidos ao longo desta tese.

Com os avanços alcançados atualmente no escopo das tecnologias de voz, uma linha de pesquisa tem atraído os especialistas de inteligência artificial, processamento de voz e linguístas: A síntese de voz com emoções. Depois de terem acumulado con-

hecimento suficiente e desenvolvido sistemas capazes de reproduzir a voz humana com destacável qualidade, pesquisadores buscam integrar a esta reprodução, elementos de expressão e emoção da fala humana [24, 25, 26, 27]. Entretanto há uma carência muito grande nesta área quando tratamos a respeito destes sistemas para o PB (Português Brasileiro). Esta tese foi motivada, entre outros fatores, pela carência de dados de voz gravados oficialmente para o nosso idioma, vislumbrando desenvolver o processo de síntese, não apenas sintetizando voz e entregando uma resposta, mas inserindo características emocionais na interação, dando-nos a sensação de que temos voz sintetizada impregnada de alguma emoção.

1.1 Sistemas de Conversão Texto-fala

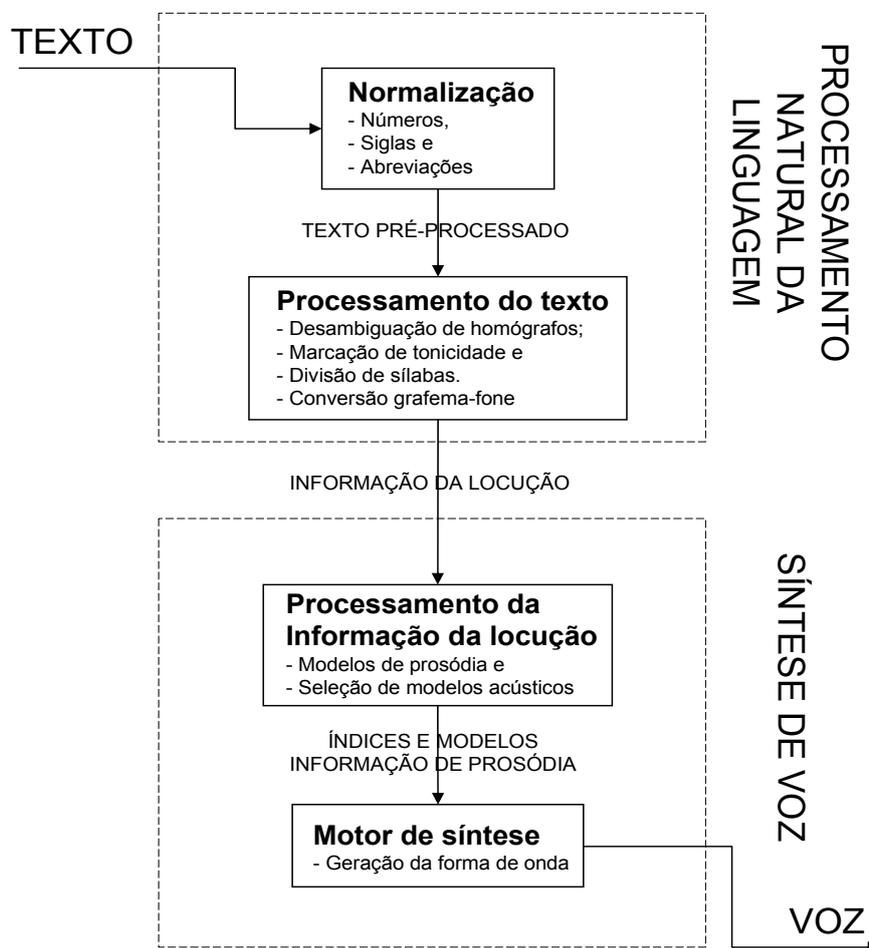


Figura 1.2: Composição básica de um TTS.

Sistemas TTS são aqueles capazes de converter a linguagem de texto em fala audível, reproduzida pelo sistema. Um sistema TTS é, basicamente, composto de duas partes: o processamento da linguagem (NLP - *Natural Language Processing*),

que é geralmente dependente da linguagem, e a síntese de voz, como pode ser visto na Figura 1.2. O NLP tem como objetivo obter um conjunto de representações linguísticas do texto, ou seja, unidades acústicas básicas, que representem o texto de entrada. Esse objetivo é alcançado através de duas tarefas: Primeiro, converter o texto da entrada, contendo símbolos (números, abreviações, siglas, etc), em uma forma equivalente com o texto por extenso. Este processo é frequentemente chamado de normalização do texto, pré-processamento ou tokenização. Depois, o NLP atribui uma transcrição fonética para cada palavra, além de dividir o texto em unidades prosódicas. O processo de associação das palavras à transcrição fonética é chamado de conversão texto-fonema ou G2P (*Grapheme-Phone*). A transcrição fonética, juntamente com as informações prosódicas compõem a representação linguística na saída do NLP. A Síntese de Voz converte a representação linguística em som [28, 29].

Como a síntese de voz propriamente dita é realizada a partir de um texto pré-processado, ou seja, submetido a um processamento prévio da linguagem, a eficiência do sintetizador está diretamente relacionada a um processamento da linguagem adequado. Estudos no contexto do processamento da linguagem têm sido realizados para tornar a conversão e, conseqüentemente, a reprodução, a melhor possível [30, 31, 32]. Muitos métodos tem sido utilizados, contudo, aqueles baseados em regras linguísticas tendem a se apresentar como as melhores opções [33, 34, 35], especialmente no caso do PB, que é um idioma fonologicamente regular, além de gerar economia computacional e possibilitar a adição de novas entradas com facilidade, aumentando a capacidade do processamento [36, 37].

1.1.1 Processamento da Linguagem

Uma correta e adequada conversão G2P contribui, com um grau bem elevado, numa síntese eficiente. Porém, esta conversão depende da eficiência de outros fatores antecedentes, tais como: a normalização, a desambiguação de homógrafos, marcação de tonicidade, a separação das sílabas, etc.

Uma data, na forma “11-jan-71”, deve ser normalizada como “onze de janeiro de setenta e um” e não como “onze traço jan traço setenta e um”. Isto soa de forma muito desagradável para qualquer usuário de Sistemas TTS. Da mesma forma, são também desagradáveis, equívocos na ambiguidade de palavras, como “sede”, “cerca”, “forma”, que dependem do contexto para ter a definição de sua vogal tônica ser aberta ou fechada. Esta tese analisa estes e outros casos visando uma melhoria do processamento do texto através de regras linguísticas.

1.2 Objetivos

A proposta desta tese foi desenvolver um conjunto de algoritmos capazes de realizar a conversão de texto em fala com o nosso TTS para o PB baseado em HMM (HBPTTS - *HMM-based Brazilian Portuguese text-to-speech*) [28, 29], reproduzindo a fala com emoções variadas, além de boa naturalidade e inteligibilidade. Para tal fim, algumas metas foram traçadas:

- O desenvolvimento de algoritmos para realizar a análise do texto (desambiguação de homógrafos, marcação de tonicidade, separação de sílabas e conversão grafema-fone), proporcionando uma transcrição fonética adequada e, conseqüentemente, um boa reprodução;
- O projeto e desenvolvimento de uma nova base de dados de voz para reconhecimento e síntese com emoções, desde a montagem do estúdio, seleção de textos e características, seleção de locutores / dubladores, gravação, edição e validação, a fim de que o TTS pudesse ser treinado na nova vertente de sintetização;
- Desenvolver algoritmos para Síntese com estados emocionais variados.

1.3 Resumo dos resultados obtidos

- Um algoritmo baseado em regras linguísticas para a desambiguação de homógrafos foi desenvolvido, capaz de tratar a ambigüidade de 111 pares de homógrafos (mas não apenas estes) encontrados no PB, separados por tipo, mediante uma análise semântica e/ou morfossintática. O algoritmo foi testado com textos de naturezas variadas (jornalística, formal-religiosa e histórico-literária), atingindo uma taxa de acerto de aproximadamente 99,00%.
- Um algoritmo baseado em regras linguísticas para a determinação da sílaba tônica foi desenvolvido, capaz de identificar e marcar a vogal tônica de cada palavra. O conjunto é composto de 19 regras e foi testado com texto de natureza jornalística, contendo cerca de 16.000 palavras, atingindo uma taxa de acerto em torno de 98,90%.
- Um algoritmo baseado em regras linguísticas para a separação de sílabas foi desenvolvido, capaz de realizar a separação das sílabas do texto de entrada. O conjunto é composto de 25 regras e foi testado com texto de natureza jornalística, contendo cerca de 18.000 sílabas, atingindo uma taxa de acerto em torno de 99,29%.



Figura 1.3: Espelho da interface gráfica do HBPTTS.

- Um algoritmo baseado em regras linguísticas para conversão grafema-fone foi desenvolvido, capaz de realizar a transcrição dos grafemas de entrada nos fones descritos em [38]. O algoritmo foi testado com texto de natureza jornalística, contendo cerca de 40.000 fones, atingindo uma taxa de acerto de 98,57%, superando resultados anteriores que eram de 97,44%. [30].
- Gravação de uma nova base de dados de voz no PB, a ser aplicada em reconhecimento e síntese de voz com emoções, gravada em estúdio de alta qualidade, com 20 locutores/dubladores profissionais, distribuídos em mais de 300 horas de gravação, contendo uma variedade de estilos gramaticais e emocionais. Um site para validação on line da nova base de dados foi criado, onde 20 avaliadores realizaram o teste subjetivo. Uma taxa de percepção emotiva de 81,8% foi alcançada.
- Um algoritmo para segmentação automática dos dados gravados em estúdio, nos moldes necessários ao treinamento dos HMMs.
- Um algoritmo para treinamento dos HMMs nos moldes necessários ao HBPTTS.
- Uma interface gráfica também foi desenvolvida para tornar mais amigável a interação dos usuários com o TTS (Figura 1.3), mantendo a opção de Síntese

via linha de comando, até então utilizada.

1.4 Estrutura do trabalho

No Capítulo 2, apresentamos as nossas discussões sobre desambiguação de homógrafos, bem como a metodologia utilizada para minimizar este problema, experimentos com textos de distintas naturezas e os resultados obtidos.

O Capítulo 3 trata a questão da tonicidade e sua importância no contexto da conversão grafema-fone. As regras linguísticas para esta marcação são apresentadas, além de testes e resultados obtidos relacionados ao assunto.

Tratamos, no Capítulo 4, a questão da separação do texto processado em sílabas. O conjunto de regras utilizadas é apresentado, além de discutirmos a importância da sílaba no ritmo fala.

No Capítulo 5, o método utilizado para realizar a transcrição fonética é tratado, bem como o impacto de módulos anteriores ao G2P na transcrição. Testes realizados e resultados obtidos também são mostrados.

No Capítulo 6, tratamos de forma sucinta da questão das emoções no escopo da síntese, bem como os detalhes inerentes ao desenvolvimento de uma nova base de dados de voz para síntese e reconhecimento. Apresentamos também os resultados inerentes a validação dos dados através de teste perceptual “online”.

No Capítulo 7, tratamos da síntese de voz com emoções propriamente dita. O assunto HMM é introduzido e suas considerações são analisadas sucintamente, como recurso para a síntese com emoções. A teoria é apresentada, bem como o sistema completo de adaptação das emoções através da técnica CMLLR. Testes são realizados e seus resultados avaliados.

No Capítulo 8, apresentamos as conclusões da tese, bem como os próximos passos a serem tomados a respeito do que já foi realizado.

Por fim, nos Apêndice A, B e C, temos, respectivamente, uma lista com publicações do autor desta tese, um tutorial completo com os algoritmos para treinamento do HBPTTS, até então inexistente, e uma lista com as frases gravadas na base de dados utilizada nesta tese.

Capítulo 2

Algoritmo para desambiguação de homógrafos

Neste capítulo, apresentamos os detalhes inerentes ao desambiguador de homógrafos, nas suas 111 regras desenvolvidas para este fim. Na Seção 2.1, fazemos uma introdução do assunto, mostrando sua importância, bem como algumas pesquisas realizadas e a motivação para o desenvolvimento deste conjunto de regras. Na Seção 2.2, temos uma análise sobre as consequências da ambiguidade de palavras em sistemas TTS. A Seção 2.3 contem os pares de homógrafos, separados por tipo. Na Seção 2.4, apresentamos as bibliotecas utilizadas no processamento linguístico, bem como a alternância vocálica e a oposição gramatical relativa a cada par de homógrafos. A Seção 2.5 contem o conjunto de regras linguísticas utilizado na desambiguação de homógrafos, além de algumas observações oriundas do Acordo Ortográfico da Língua Portuguesa de 2008, que impactaram diretamente no estudo abordado pelo conjunto de regras mencionado neste capítulo. A Seção 2.6 segue com a apresentação da base de dados utilizada, os testes realizados, bem como os resultados obtidos. Finalmente, na Seção 2.7, apresentamos nossas conclusões e trabalhos futuros.

2.1 Introdução

A distinção entre homógrafos é um grande problema, com solução extremamente complexa na transcrição G2P, porque a saída da transcrição fonética não é única para cada homógrafo, ou seja, o algoritmo que faz a transcrição precisa decidir entre duas situações possíveis: ou a vogal tônica é aberta ([E]/[O]) ou ela é fechada ([e]/[o]). Palavras como <seca> (substantivo, “a s[e]ca”, e verbo, “ele s[E]ca”) e <sede> (substantivo, “a s[e]de”, e também substantivo, “a s[E]de”), possuem a mesma grafia, porém significados diferente. Por este motivo, elas são chamadas de homógrafos

heterófonos, que a partir de agora serão mencionadas apenas como homógrafos, e se não forem analisadas corretamente, podem comprometer a transcrição fonética e, conseqüentemente, a qualidade da síntese. Mesmo que o número de homógrafos existente em um segmento de texto represente um percentual bem pequeno em relação ao texto analisado (cerca de 1%, como verificado nesta tese), no contexto da síntese de voz, a transcrição fonética equivocada e a conseqüente reprodução da voz, não soam de forma agradável a qualquer ouvinte, atraindo muito a atenção para o erro corrente, mesmo que ocorram em número reduzido. Diminuir os erros entre homógrafos, melhora significativamente a qualidade do sintetizador, bem como sua naturalidade e inteligibilidade.

Em geral, podemos ter homógrafos que pertencem a diferentes classes gramaticais (POS - *part-of-speech*) (<seca>) e homógrafos pertencentes a mesma classe gramatical (<sede>), conforme exemplificado anteriormente. Desta forma, para uma Conversão G2P adequada, o uso de regras linguísticas para coleta de informações morfossintáticas, no caso de homógrafos de diferentes classes gramaticais, e de informações semânticas, para o caso de homógrafos com mesma classe, são um bom recurso para decidir qual será a tonicidade da vogal diferencial do homógrafo, ou seja, se a palavra possui vogal tônica aberta ou fechada. O problema reside no fato de que desconsiderar estas informações linguísticas, fatalmente irá conduzir a transcrição fonética a falhas que resultarão no prejuízo da inteligibilidade, de forma bastante desagradável ao ouvinte ou usuário do sistema TTS, já que teremos apenas um fone de saída para cada grafema de entrada.

Homógrafos é um assunto bastante explorado pela comunidade científica dada a sua importância em síntese de voz em diversos idiomas. Em [39] temos uma tipologia de pares de homógrafos na língua inglesa e algumas técnicas para desambiguação tradicionalmente utilizadas, tais como N-gram, taggers, classificadores bayesianos e árvore de decisão, bem como a proposta de um sistema híbrido, combinando as três melhores técnicas descritas. Em [40], o assunto é tratado em idiomas como tailandês, chinês e japonês, onde as palavras não possuem delimitação de fronteira. Uma técnica de Reconhecimento de Padrões, denominada “winnow”, é utilizada para fazer a segmentação das palavras e resolver o problema de ambigüidade dos homógrafos. [41] apresenta um estudo da relação entre caracteres chineses e sua pronúncia, bem como propõe uma solução para a disambiguação de caracteres polifônicos. Em relação a desambiguação de homógrafos em sistemas TTS no português europeu, [42, 43] analisam as melhorias que podem ocorrer no desempenho da conversão texto-fala, através da influência de informações morfossintáticas na desambiguação de homógrafos. Para o PB, em [44, 45] um analisador morfossintático é apresentado para solucionar o problema de alternâncias vocálicas entre substantivos e verbos, sem, no entanto, abordar a desambiguação de homógrafos semanticamente. Já em

[46, 47, 48], temos ambas as abordagens, morfossintática e semântica, mas o método foi testado apenas com um único exemplo.

Nesta tese, apresentamos um método de desambiguação de homógrafos heterófonos aplicado a um sistema de TTS para o PB [28], que soluciona a ambiguidade de um número amplo de pares de homógrafos, tanto morfossintaticamente como semanticamente. O algoritmo proposto foi implementado e testado com um segmento de texto extraído aleatoriamente da base de dados de texto com naturezas distintas, a saber: CETEN-Folha (natureza jornalística) [49], Bíblia Cristã (natureza formal e religiosa) [50] e da obra literária “Dom Casmurro”, de Machado de Assis (natureza histórico-literária) [51]. Os resultados alcançaram uma taxa de acerto, respectivamente, de 99,00%, 99,00% e 99,02%.

2.2 Consequências da ambiguidade

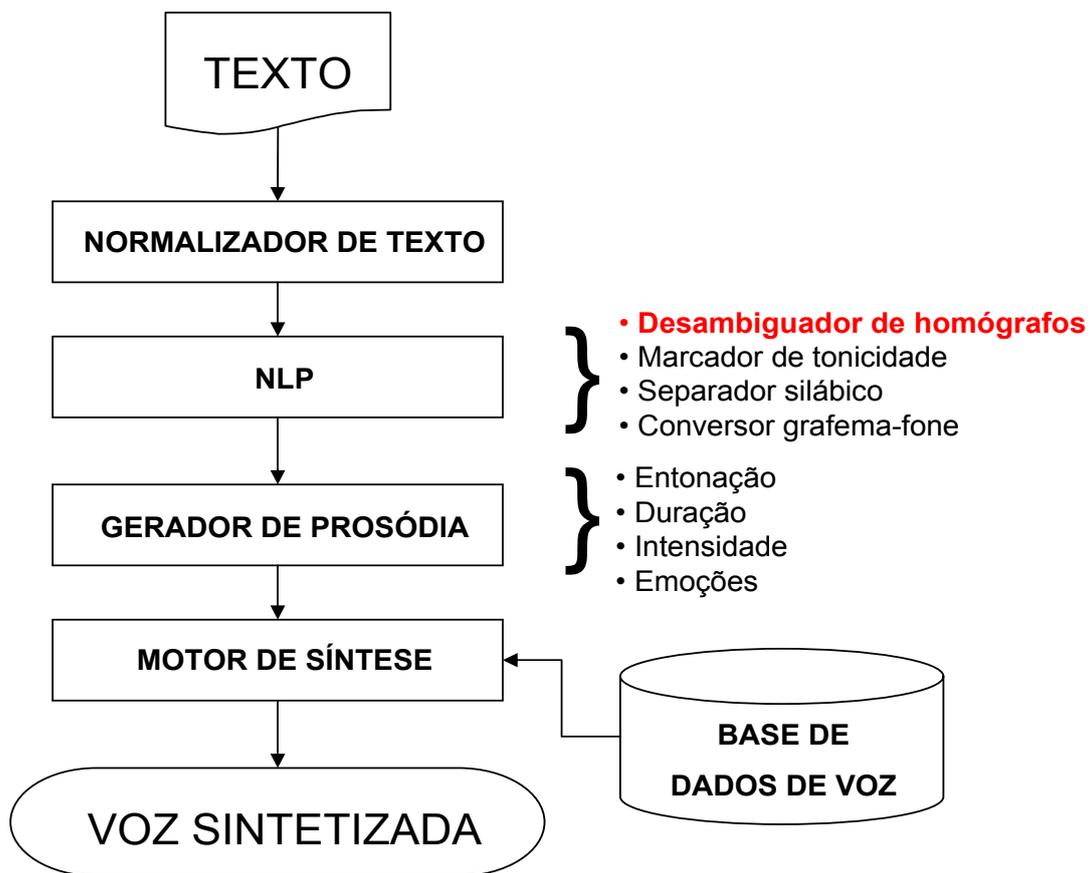


Figura 2.1: Localização do desambiguador de homógrafos no TTS.

Como mencionado anteriormente, a ambiguidade de palavras pode ser um problema sério na conversão texto-fala, principalmente vislumbrando a possibilidade de

que, num futuro bem próximo, esta atividade será bastante usual no nosso dia-a-dia.

Observe o trecho a seguir:

- “Tenho uma cerca na minha propriedade. Ela cerca toda a área e tem cerca de 8.000 metros”.

Podemos perceber no mesmo trecho, as três situações possíveis de ocorrência do homógrafo “cerca”: SUBSTANTIVO, VERBO E PREPOSIÇÃO, respectivamente.

A correta representação vocálica seria: “Tenho uma c[e]rca na minha propriedade. Ela c[E]rca toda a área e tem c[e]rca de 8.000 metros”. Caso o assunto homógrafos não fosse abordado no escopo da síntese de voz, esta frase certamente apresentaria algum equívoco na tonicidade, seja com todas as vogais abertas ou todas fechadas.

Este é um caso em que uma análise dupla é necessária (morfofossintática e semântica), já que temos palavras de classes gramaticais diferentes com a mesma transcrição fonética (“c[e]rca”, substantivo e preposição e “c[E]rca”, verbo). Este caso, particularmente, é analisado através do algoritmo 14.

A localização do desambiguador de homógrafos proposto no TTS pode ser vista na Figura 2.1. Podemos observar que a desambiguação de homógrafos faz parte de um conjunto de algoritmos importantes para a transcrição fonética, já que aquele algoritmo faz uma prévia marcação de tonicidade em cerca de 1% do texto analisado, como verificado empiricamente nesta tese.

2.3 Metodologia aplicada na desambiguação

O algoritmo proposto para desambiguação de homógrafos é apresentado na Figura 2.2. Este algoritmo é, em essência, um conjunto de regras linguísticas, separadas por tipo, baseado em um método composto de duas partes: A análise morfofossintática, que soluciona problemas com homógrafos de classes gramaticais diferentes e a análise semântica, que soluciona problemas com homógrafos de mesma classe.

A metodologia aplicada neste trabalho partiu da coleta de pares de homógrafos existentes no português falado no Brasil, resultando num conjunto de 111 pares de homógrafos, que pode ainda ser atualizado, já que temos a versatilidade de acrescentar quantos homógrafos forem necessários na biblioteca de homógrafos existente. O conjunto de homógrafos está organizado baseado em regras por tipos, de acordo com a oposição gramatical existente e a alternância fonética no par de homógrafos. Para cada tipo de homógrafos, temos um algoritmo com várias perguntas relativas ao contexto morfofossintático do homógrafo, que vão contribuir para a correta transcrição fonética da vogal tônica. Os homógrafos puderam ser agrupados no mesmo tipo, quando o conjunto de perguntas do algoritmo era o mesmo, ou seja, com mesma classe gramatical.

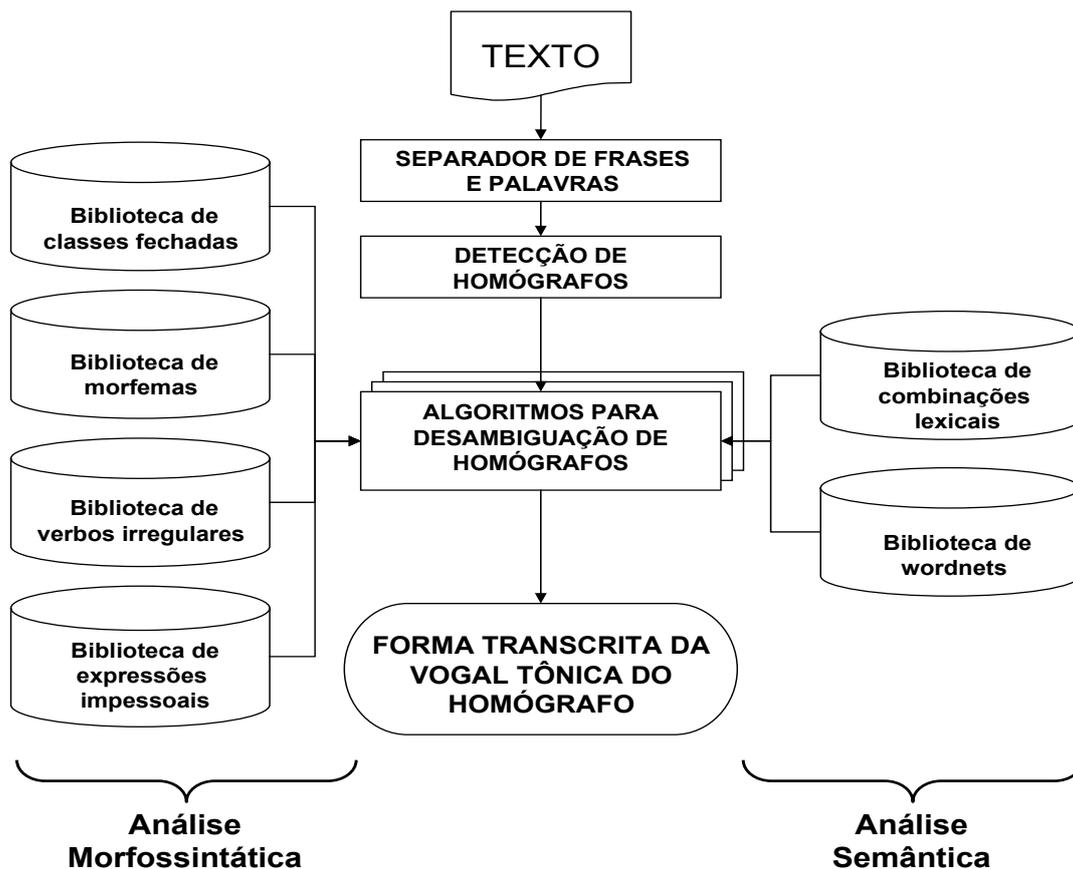


Figura 2.2: Estrutura do desambiguador de homógrafos.

Para a realização das perguntas nos algoritmos e a determinação das classes gramaticais das palavras vizinhas ao homógrafo, várias bibliotecas foram coletadas, formando um analisador morfológico.

Os algoritmos de desambiguação foram testados com CETEN-Folha (corpus contendo linguagem de jornal), além de textos com outras naturezas. Na Tabela 2.1,

Tabela 2.1: Conjunto de homógrafos separados por tipo.

Tipo	Homógrafo
1	acerto, apelo, aperto, apreço, começo, concerto, conserto, desemprego, desespero, emprego, enredo, erro, esmero, espeto, flagelo, gelo, governo, interesse, interesses, modelo, pego, peso, rego, selo, testo ¹ e zelo.
2	aborto, acordo, adorno, aforro, almoço, apoio, arrojo, arroteio, choco, choro, conforto, consolo, contorno, controle, coro, desgosto, despojo, destroço, encosto, endosso, esforço, estorvo ² , folgo ³ , gosto, jogo, logro ⁴ , namoro, olho, piloto, reforço, rodo, rogo, rolo, sopro, suborno, sufoco, toco, toldo, topo, torno, troco e troço ⁵ .
3	rola e rolha ⁶ .
4	colher e meta.
5	desses, deste e destes.
6	fora.
7	seco, seca e secas.
8	boto.
9	este.
10	leste.
11	sobre.
12	rota, rotas, tola e tolas.
13	corte, cortes, forma, formas, molho e soco.
14	cerca.
15	pega e pegas ⁷ .
16	pelo, pela e pelas.
17	besta e bestas.
18	sede e sedes.
19	medo e medos.
20	termos.
21	cor.
22	lobo e lobos.
23	bola e bolas ⁸ .

temos o conjunto de homógrafos usados nesta tese, separados por tipo.

¹adj (de testa) 1 Enérgico, firme, resoluto, sério. 2 Que não gosta de brincadeiras. sm Reg (Ceará) Murro. (ê) sm (lat testu) 1 Tampa de barro para vasilha da mesma substância. 2 Tampa de ferro para tacho ou panela do mesmo metal. 3 Vaso de barro onde se põe a cal para caiar. 4 Testico. 5 O casco da cabeça do boi. 6 Camada endurecida de fuligem no fundo das panelas. 7 Camada de barro para filtrar a água dos pães de açúcar. sm pl pop Cabeça, mioleira.

²(ô) sm (der regressiva de estorvar) 1 Ato ou efeito de estorvar. 2 Embaraço. 3 Impedimento, obstáculo. 4 Dificuldade. 5 Oposição. Var: estorvamento.

³(ô) sm (corr de fôlego, com síncope) 1 V fôlego. 2 ant Folgança.

⁴(ô) sm (lat lucru) 1 Ato ou efeito de lograr. 2 Engano propositado contra alguém. 3 Burla. 4 Engano jocoso; partida, peça. 5 Desfrute, posse. 6 Preço excessivo.

⁵adj (de troço) gír Diz-se de pessoa importante, influente: Fulano é troço na política. sm pop O mesmo que objeto. (ô) sm (fr arc tros) 1 Pedaco de pau roliço e tosco. 2 O mesmo que trocho. 3 Peça em que se fixam os degraus das escadas dos navios. 4 Artilh Cada uma das partes do molde do canhão composto de aduelas ligadas entre si por anilhas de ferro. 5 Mil Parte de uma armada. 6 Mil Parte de um corpo de tropa. 7 Náut Obra de marinheiro feita de amarra ou de cabos velhos. 8 Pedaco de estrada etc. 9 Porção de gente; rancho, grupo. sm pl 1 Mec Pedaco de arame recurvado, que se mete num orifício de uma haste roscada, para impedir que uma porca se afrouxe ou se desenrosque. 2 Cacaréus, trastes velhos e quebrados. 3 Coisas móveis. T. de lança: pedaco de haste quebrada de lança. T. descendente, Bot: parte mais grossa da raiz, da qual nascem várias radículas, por cujo meio ela chupa a substância que alimenta a planta. Troços grossos: maquinismo

2.4 Bibliotecas empregadas na análise do texto

Várias bibliotecas foram usadas na análise morfossintática e na análise semântica:

- Biblioteca de homógrafos, contendo 111 pares de homógrafos agrupados em 23 tipos, que podem ser vistos na Tabela 2.1.
- Biblioteca de classes fechadas, contendo as classes gramaticais que tem um número fixo de componentes (pronomes, preposições, conjunções, contrações, artigos, numerais, interjeições).
- Biblioteca de morfemas, contendo sufixos nominais, verbais, adjetivais e adverbais, prefixos e radicais gregos e latinos.
- Biblioteca de lemas, contendo o dicionário em português Jspell com aproximadamente 34000 palavras morfológicamente anotadas [52].
- Biblioteca de verbos irregulares, contendo a flexão dos principais verbos irregulares existentes no PB.

para a primeira torcedura, nas fábricas de fiação. A troços: com interrupção ou intervalos. vtd e vti (lat *tortiare* - de troçar) Escarnecer, fazer escárnio de; ridicularizar; zombar: Troçar alguém (ou com alguém, ou de alguém).

⁶(ô) sf (lat *rotula*) 1 Peça oblonga e quase sempre cilíndrica, feita de cortiça ou de outra substância, para tapar a boca ou gargalo das garrafas, frascos etc. 2 pop Pessoa manhosa, patife, tratante. 3 Imposição de silêncio. 4 Partida em que os perdedores não fizeram a metade sequer dos pontos do ganhador; capote. 5 Censura, repreensão. R.-do-maluvo: arbusto africano, com cujas folhas os indígenas tapam as cabaças do maluvo. V (de *rolha*+*ar2*) arrolhar.

⁷sf (der regressiva de pegar) 1 Ação de pegar. 2 Ato de segurar com as mãos a rês nas touradas ou em serviço. 3 Ponto por onde se pega, como a asa, o cabo etc. 4 Recrutamento forçado. 5 Azo. 6 Solidificação do cimento, gesso, estuque etc. sm 1 Discussão acalorada. 2 Desavença, desordem. 3 Luta. 4 Alarma, clamor público, perseguição. 5 Braga de ferro com que se prendiam os pés dos escravos fugitivos. 6 Reg (Paraíba) Tragada em cigarro de maconha; tapa. (ê) sf (lat *pica*) 1 Ornit Pássaro grande da família dos Corvídeos (*Corvus pica*) que ocorre na Europa. 2 Náut Peça de madeira que cobre a cabeça dos mastros. 3 pop Mulher feia e vestida sem gosto. 4 Mulher que fala muito. 5 pop Qualquer mulher. 6 Zootécni Animal que tem o pêlo de duas cores apresentando porções de branco e de qualquer outra cor geralmente simples. 7 Ornit Denominação dada no Brasil a alguns pássaros da família dos Icterídeos, também chamados encontro. adj Diz-se do cavalo malhado, também chamado pampa. P.-real, Náut: peça que mantém verticalmente o mastaréu da gávea.

⁸sf (lat *bullā*) 1 Corpo redondo em toda a sua superfície; esfera. 2 Objeto esférico ou arredondado. 3 Objeto esférico ou ovóide de qualquer espécie e material diverso, maciço ou inflado de ar comprimido, para ser jogado, batido ou chutado em jogos ou esportes. 4 Peça de carne envenenada com que se matam cães. 5 fam Cabeça, juízo. 6 Reg (Nordeste) Bala, rebuçado. 7 Pessoa baixa e gorda. 8 Chiste, piada. sf pl 1 O mesmo que boleadeiras. 2 Pó de carvão amassado com greda, em forma de disco, para conservar o calor nos fogareiros. B.-de-milho, Reg (Ceará e Alagoas): peteca feita de palha de milho. B.-de-neve: a) Bot: arbusto das Caprifoliáceas, de flores brancas dispostas em corimbo esférico, e também chamado sabugueiro-dos-pântanos (*Viburnum opulus*); b) progressividade, numa sucessão de fatos. B. de sabão: a) bolha que se forma na extremidade de um tubo, molhado em água de sabão, quando se sopra pela outra extremidade; b) coisa inconstante e passageira; c) coisa fútil. B. elástica: a que é feita de goma-elástica. Dar bola: dar confiança, dar atenção. Não ser certo da bola, sofrer da bola: não ter bom juízo, ser meio maluco. Trocar as bolas: enganar-se; dizer ou fazer uma coisa por outra. Humoristicamente, e com o mesmo sentido, é usada a expressão bolar as trocas.

- Biblioteca de expressões impessoais, contendo expressões com o verbo ser na terceira pessoa + adjetivo.
- Biblioteca de combinações lexicais restritas, contendo expressões idiomáticas, provérbios ou expressões fixas de uma ou mais palavras. Esta biblioteca é usada na análise semântica.
- Biblioteca de wordnets, desenvolvidas sob o conceito de Wordnet [53], contendo palavras que são semanticamente e cognitivamente relacionadas com o homógrafo analisado. Esta biblioteca também é usada na análise semântica.

Na Figura 2.2, a estrutura do sistema é mostrada. O texto é inicialmente dividido em palavras e frases. Em seguida, o sistema realiza a busca por homógrafos, através da biblioteca existente. Uma vez identificado um homógrafo, este é conduzido ao algoritmo correspondente ao seu tipo. As bibliotecas necessárias a realização da análise morfossintática podem ser vistas a esquerda, da mesma forma que as bibliotecas usadas na análise semântica, podem ser vistas a direita.

Na Tabela 2.2 e na Tabela 2.3, podemos ver homógrafos pertencentes a diferentes classes gramaticais, bem como pertencentes a mesma classe gramatical, respectivamente. A ambigüidade dos homógrafos da Tabela 2.2 é solucionada com uma análise morfossintática, porém informações semânticas são necessárias para solucionar o problema da ambigüidade nos homógrafos da Tabela 2.3. Como as Tabelas mostram, as oposições gramaticais que mais ocorrem são entre Substantivo e Verbo, do ponto de vista morfológico, e entre [e]/[E] e [o]/[O], do ponto de vista fonético. Uma evidência sistemática é que em Substantivos, a vogal tônica é tipicamente fechada, embora nas formas verbais a vogal tônica seja aberta. Homógrafos do Tipo 1 e 2 representam 61,3% do total de homógrafos da biblioteca. Homógrafos do Tipo 13, 14, 15 e 20 necessitam tanto da análise morfológica como da análise semântica.

Tabela 2.2: Exemplos de homógrafos com diferentes classes gramaticais.

Tipo	Alternância vocálica e Oposição gramatical	Exemplo
1	[e] Substantivo / [E] Verbo	Nosso <u>er</u> ro foi muito grande. Eu <u>er</u> ro bastante.
2	[o] Substantivo / [O] Verbo	Ele fechou o <u>o</u> lho esquerdo. Eu <u>o</u> lho para cima.
3	[o] Substantivo / [O] Verbo	Eu vi uma <u>ro</u> la branca. Ele deita e <u>ro</u> la.
4	[e] Substantivo / [E] Verbo	É época de <u>co</u> lher o tomate. Essa é a nossa <u>me</u> ta.
5	[e] Contração / [E] Verbo	Ele ganhou dois <u>de</u> sses prêmios. Era bom que nunca <u>de</u> sses a notícia.
6	[o] Verbo / [O] Advérbio	Ele <u>fo</u> ra uma pessoa honesta. Eu estou <u>fo</u> ra do jogo.
7	[e] Adjetivo ou Substantivo / [E] Verbo	O rio estava muito <u>se</u> co. Eu <u>se</u> co os pés na entrada.
8	[o] Adjetivo ou Substantivo / [O] Verbo	Ele viu um <u>bo</u> to na praia. Eu <u>bo</u> to azeite na salada.
9	[e] Demonstrativo / [E] Adjetivo ou Substantivo	<u>Es</u> te armário é meu. Norte, sul, <u>es</u> te, oeste.
10	[e] Verbo / [E] Adjetivo ou Substantivo	<u>Le</u> ste a notícia? Seguiu para o <u>le</u> ste.
11	[o] Preposição / [O] Verbo	Comentou <u>so</u> bre o fato. É bom que <u>so</u> bre uma garrafa.
12	[o] Adjetivo ou Verbo [O] Substantivo /	Ela andava toda <u>ro</u> ta. Nós seguimos a <u>ro</u> ta.
13	[o] Substantivo / [O] Verbo / Substantivo	Ela comprou pão de <u>fo</u> rma. De qualquer <u>fo</u> rma iremos ao passeio.
14	[e] Preposição / Substantivo / [E] Verbo	Eles andaram <u>ce</u> rca de dez quilômetros. Ele <u>ce</u> rca seu terreno com arame farpado.
15	[e] Substantivo / [E] Verbo / Substantivo	Aquela ave parece uma <u>pe</u> ga. Olha que essa moda ainda <u>pe</u> ga.
16	[e] Contração / Substantivo [E] Verbo / Substantivo	Nós passamos <u>pe</u> la rua. Ela <u>pe</u> la todo seu corpo.

Tabela 2.3: Exemplos de homógrafos com mesma classe gramatical.

Tipo	Alternância vocálica e Oposição gramatical	Exemplo
17	[e] Substantivo / [E] Substantivo	Ele é metido a <u>be</u> sta. Ele conseguia disparar a <u>be</u> sta.
18	[e] Substantivo / [E] Substantivo	Ele estava com uma <u>se</u> de insuportável. A <u>se</u> de da empresa fica em Paris.
19	[e] Substantivo / [E] Substantivo	Ela estava com <u>me</u> do de morrer. Eles venceram todo o Império <u>Me</u> do-Persa.
20	[e] Substantivo / Verbo [E] Substantivo	Estes são os nossos <u>te</u> rmos. A <u>te</u> rmos tinha café quente.
21	[o] Substantivo / [O] Substantivo	O vestido era <u>co</u> r de rosa. Sabia tudo de <u>co</u> r e salteado.
22	[o] Substantivo / [O] Substantivo	Na estória não tinha <u>lo</u> bo mau. Ele feriu o <u>lo</u> bo temporal.
23	[o] Substantivo / [O] Substantivo	Só amassei a <u>bo</u> la de carne. Eu não tenho <u>bo</u> la de cristal.

2.5 Conjunto de regras para desambiguação de homógrafos

Depois das sentenças serem divididas em palavras e frases, o sistema busca candidatas a homógrafos e as relaciona com sua biblioteca de homógrafos. Se o sistema identifica uma palavra como sendo homógrafo, ele a relaciona com um tipo. Cada tipo possui um algoritmo correspondente com perguntas sobre o contexto de um determinado homógrafo. As bibliotecas existentes são utilizadas para esta tarefa. Nesta Seção, os algoritmos de desambiguação, bem como as regras linguísticas aplicadas, são mostrados para cada um dos tipos apresentados nesta tese. Os símbolos usados nos algoritmos podem ser vistos na Tabela 2.4. Os 23 algoritmos são apresentados na sequência, ao longo do texto, do Algoritmo 1 até o Algoritmo 23. Na maioria dos algoritmos, o primeiro conjunto de perguntas conduz a saída mais provável, segundo a gramática vigente. Se a resposta for negativa, o sistema realiza outra seção de perguntas e, caso a resposta seja positiva, a saída é a ocorrência estatisticamente menos provável. Se ainda for negativa, o sistema conduz ao caso padrão (default). Em outros algoritmos temos mais do que dois conjuntos de perguntas, proporcionando várias possibilidades de saída, conforme a necessidade da análise morfológica.

2.5.1 Os novos homógrafos derivados do Acordo Ortográfico de 2008.

Nesta Subseção, apresentamos as alterações na língua portuguesa que originaram novos pares de homógrafos e, conseqüentemente, repercutiram neste trabalho. O Acordo é apenas ortográfico, restringindo-se apenas à linguagem escrita, não tendo efeito algum sobre a linguagem falada.

Os ditongos abertos ([E]i, [O]i) perderam o acento em palavras paroxítonas acentuadas. Então, palavras como <apóio> (verbo, “Eu apoio a idéia.”), que perdeu o acento grave, tornou-se homógrafo de <apoio> (substantivo, “Ele tem meu apoio no partido.”). Dada a alternância vocálica e oposição gramatical (ap[o]io - substantivo; ap[O]io - conjugação do verbo “apoiar”), é caracterizado como homógrafo pertencente ao tipo 2.

Da mesma forma, o acento diferencial foi extinto. Sendo assim, palavras como as formas verbais <pêla>, <pêlas> (“Minha filha pêla todo o corpo.”) e <pêlo> (“Eu pêlo o meu rosto.”), bem como as formas nominais <pêla>, <pêlas> (“Ele sempre joga pêla.”) e <pêlo> (“O pêlo do animal é muito bonito.”), tornaram-se homógrafos com as contrações <pelo>, <pelos>, <pela> e <pelas>. A alternância vocálica e a oposição gramatical destas palavras são caracterizadas como: p[e]la, p[e]las (con-

Tabela 2.4: Simbologia usada nos algoritmos de desambiguação.

Símbolo	Significado
P-1, P-2	Última e penúltima palavra, respectivamente.
P+1	próxima palavra.
F0, F-1, F+1	frase atual, última frase e próxima frase, respectivamente.
P_DEM	pronome demonstrativo.
P_IND	pronome indefinido.
P_INT	pronome interrogativo.
P_POSS	pronome possessivo.
A_IND	pronome indefinido.
P_RELA	pronome relativo.
PREPO	preposição.
CONTR	contração.
P_PESS_SU	pronome pessoal reto.
P_PESS_O	pronome pessoal oblíquo (<me>,<mim>,<te>,<ti>,<se>,<si>,<nos>,<vos>,<lhe(s)>,<no-lo(s)>,<no-la(s)>,<vo-lo(s)>,<vo-la(s)>,<lho(s)> ou <lha(s)>).
CS	conjunção subordinativa.
CC	conjunção coordenativa.
HN	“a”, “o”, “as” ou “os” (Homonímia - pronome ou artigo definido).
nc	substantivo.
adv	advérbio.
ad	adjetivo.
NUM	numeral.
DESV	sufixos verbais.
EXPIMP	expressões impessoais.
PART	particípio.
PART_IRR	particípio irregular.
VERB_IRR	verbos irregulares.
VERB	verbos regulares.
BC	combinações lexicais restritas.
WN	wordnet.
V	vogal tônica do homógrafo.

tração da preposição “por” + artigo definido “a”, “as”) / p[E]la, p[E]las (conjugação do verbo “pelar” ou Substantivo “péla”), p[e]lo (Substantivo ou contração da preposição “por” + artigo definido “o”) / p[E]lo (conjugação do verbo “pelar”). Para este caso, foi necessário criar um novo algoritmo para a desambiguação (Algoritmo 16), pois estes pares de homógrafos não foram classificados em nenhum algoritmo previsto em [54]. Nesta tese, estes pares de homógrafos foram classificados no algoritmo do tipo 16, como o último algoritmo que necessita apenas de análise morfossintática, apesar deste tipo possuir similaridade gramatical (Substantivo, “pêlo” / Substantivo, “péla”), esta oposição não foi sustentada com o arranjo dos pares:

- p[e]lo (contração ou Substantivo) / p[E]lo (verbo);
- p[e]la(s) (contração) / p[E]la(s) (verbos ou Substantivos).

Do algoritmo 17 até o algoritmo 23 são requeridas análises semânticas, diferentemente da designação apresentada em [54].

Algoritmo 1

```
1: if (A palavra é homógrafo do tipo 1) then
2:   if (P-1 = P_DEM, P_IND, P_INT ou P_POSS) ou (P-1, P-2 ou P-3 = A_IND) ou (P-1 ou
   P-2 = HN, CONTR ou PREPO) ou (P+1 = <que> ou P_RELA) then
3:     V = [e]
4:   else if (P-1 = P_PESSO_SU, P_PESS_O_1 ou CS) ou (P+1 = PREPO, CONTR, P_
   PESS_O_1 ou HN) ou (P+1 = A_IND e P+2 = nc) ou (P-1 ou P-2 = <não> ou <nunca>)
   then
5:     V = [E]
6:   else
7:     V = [e]
8:   end if
9: else
10:   Vá para o algoritmo 2
11: end if
```

Algoritmo 2

```
1: if (A palavra é homógrafo do tipo 2) then
2:   if (P-1 = P_DEM, P_IND, P_INT ou P_POSS) ou (P-1, P-2 ou P-3 = A_IND) ou (P-1 ou
   P-2 = HN, CONTR ou PREPO) ou (P+1 = <que> ou P_RELA) then
3:     V = [o]
4:   else if (P-1 = P_PESSO_SU, P_PESS_O_1 ou CS) ou (P+1 = PREPO, CONTR, P_
   PESS_O_1 ou HN) ou (P+1 = A_IND e P+2 = nc) ou (P-1 ou P-2 = <não>, <nunca> ou
   <ainda>) then
5:     V = [O]
6:   else
7:     V = [o]
8:   end if
9: else
10:   Vá para o algoritmo 3
11: end if
```

Algoritmo 3

```
1: if (A palavra é homógrafo do tipo 3) then
2:   if (P+1 = <pelo>, ad ou adv) ou (P-2 ou P-3 = A_IND ou HN) ou (P-1 = <que>, <ele>, <ela>,
   <se>, <não>, <já>, <as>, nc, CC ou CS) ou (P-1 ou P-2 = P_DEM, P_IND, P_INT ou P_
   POSS) ou (P+1 = <e> e P+2 = <rebola>) then
3:     V = [O]
4:   else if (P-1 = <à> e P-2 = <tiro> ou <caça>) ou (P-1 = <uma> ou <a>) ou (P-1 ou P-2 = CONTR
   ou PREPO) ou (P+1 = <brava>) then
5:     V = [o]
6:   else
7:     V = [O]
8:   end if
9: else
10:   Vá para o algoritmo 4
11: end if
```

2.6 Testes realizados com o conjunto de regras

2.6.1 Base de dados utilizada

O sistema foi testado com extratos de texto provenientes de três bases com naturezas distintas (jornalística, formal-religiosa e histórico-literária). Estas vertentes foram selecionadas visando explorar o rigor gramatical, além de ter a possibilidade de encontrar alguns homógrafos e testar o respectivo algoritmo, já que seria necessário uma quantidade de palavras extremamente grande para poder encontrar todos num mesmo texto. A seguir são definidas as origens de cada fonte de texto:

Algoritmo 4

```
1: if (A palavra é homógrafo do tipo 4) then
2:   if Homógrafo = <colher> then
3:     if (P+1 = <de>) ou (P-2 começa com <met-> e P-1 = <a>) ou (P-1 = <ã>, HN ou A_IND)
4:       then
5:         V = [E]
6:       else if (P+1 = HN ou A_IND e P+2 = nc) ou (P+1 = P_DEM, P_POSS ou P_IND e P+2
7:         = nc) ou (P+1 = <em>, <no>, <na>, <nos> ou <nas>) ou (P+1 termina com <-os>, <-as>,
8:         <-ões>, <-ães>, <-ãs>, <-res> ou <-es>) ou (P-1 ou P-2 = <não>) then
9:           V = [e]
10:        else
11:          V = [E]
12:        end if
13:      else if Homógrafo = <meta> then
14:        if (P-1 = P_DEM, P_IND, P_INT ou P_POSS) ou (P-1, P-2 ou P-3 = A_IND) ou (P-1
15:        ou P-2 = HN, CONTR ou PREPO) ou (P+1 = <que> ou P_RELA) then
16:          V = [E]
17:        else if (P-1 = P_PESSO_SU, P_PESS_O_1 ou CS) ou (P+1 = PREPO, CONTR, P_
18:        PESS_O_1 ou HN) ou (P+1 = A_IND e P+2 = nc) ou (P-1 ou P-2 = <não>, <nunca> ou
19:        <que>) then
20:          V = [e]
21:        else
22:          V = [E]
23:        end if
24:      end if
25:    else
26:      Vá para o algoritmo 5
27:    end if
```

Algoritmo 5

```
1: if (A palavra é homógrafo do tipo 5) then
2:   if (P-1 = NUM) then
3:     V = [e]
4:   else if (P-1, P-2 ou P-3 = <tu>, <vós>, <ontem>, <se>, <talvez>, <oxalá> ou CS) ou (P+1 = P_
5:   PESS_O_1 ou P_PESS_O_2) ou (P-1 ou P-2 = <não> ou <nunca>) ou (P-1 = P_PESS_O_1)
6:   then
7:     V = [E]
8:   else
9:     V = [e]
10:  end if
11: else
12:  Vá para o algoritmo 6
13: end if
```

- CETEN-Folha - Esta base de dados é um corpus contendo cerca de 24 milhões de palavras do PB, criado pelo projeto Processamento Computacional do Português com base nos textos do jornal Folha de São Paulo e que fazem parte do corpus NILC/São Carlos, compilado pelo Núcleo Interinstitucional de Linguística Computacional (NILC) [49]. O teste realizado com esta base foi composto de um segmento de texto contendo 1.564.591 palavras, onde foram identificadas 20.308 homógrafos, ou seja, 1,30% do texto são homógrafos.
- Bíblia Cristã - Esta base de dados é uma versão em formato texto da Bíblia Sagrada para o PB [50], composta de mais de 750.000 palavras, apresentando um estilo linguístico mais elaborado do que a base do CETEN-Folha, motivo pelo qual foi incluída no teste dos algoritmos. Além disso, pelo estilo existente, alguns homógrafos puderam ser encontrados em maior número e, portanto, ter

Algoritmo 6

```
1: if (A palavra é homógrafo do tipo 6) then
2:   if (P+1 ou P+2 termina com <-ndo>, <-ado>, <-ada>, <-ido> ou <-ida>) ou (P+1 ou P+2 =
   PART IRR) ou (P+1 = <apenas>, A IND ou HN) ou (P-1 = <eu>, <ele>, <ela>, <você>,
   <onde>, <como>, <quando> ou <quem>) then
3:     V = [o]
4:   else if (P+1 = <de>, <do>, <da>, <dos>, <das> ou CONTR) ou (P-1 ou P-2 = <lá>, <cá> ou
   <ai>) ou (P-1 ou P-2 termina com <-mente>) ou (P-1 ou P-2 begins with <deit->, <deix->, <atir->,
   <empat->, <consider->, <fic->, <est-> ou <jog->) ou (P-1 = <borda>, <jantar>, <comer>, <noite>,
   <mundo>, <dia>, <tarde>, <por>, <de> ou <para>) ou (P-1 ends by <-ar>, <-er> ou <-ir>) then
5:     V = [O]
6:   else
7:     V = [o]
8:   end if
9: else
10:  Vá para o algoritmo 7
11: end if
```

Algoritmo 7

```
1: if (A palavra é homógrafo do tipo 7) then
2:   if (P-1 = P PESS SU, P PESS O 1 ou CS) ou (P+1 = P PESS O 1, CONTR ou HN)
   ou (P-1 ou P-2 = <não>, <nunca>, <ainda> ou <já>) then
3:     V = [E]
4:   else
5:     V = [e]
6:   end if
7: else
8:  Vá para o algoritmo 8
9: end if
```

Algoritmo 8

```
1: if (A palavra é homógrafo do tipo 8) then
2:   if (P-1 = <eu>, P PESS O 1 ou CS) ou (P-1 ou P-2 = <não> ou <nunca>) ou (P+1 = <fora>,
   P PESS O 1, CONTR ou HN) then
3:     V = [O]
4:   else
5:     V = [o]
6:   end if
7: else
8:  Vá para o algoritmo 9
9: end if
```

Algoritmo 9

```
1: if (A palavra é homógrafo do tipo 9) then
2:   if (P+1, P+2 ou P+3 = <oeste>) ou (P-1 = <vento>) then
3:     V = [E]
4:   else
5:     V = [e]
6:   end if
7: else
8:  Vá para o algoritmo 10
9: end if
```

os seus algoritmos mais solicitados. O teste, neste caso, foi realizado com toda a base e 7.904 homógrafos foram identificados (1,05% do texto processado).

- Dom Casmurro - Esta obra literária é um romance do escritor brasileiro Machado de Assis e serviu como base de dados de texto contendo quase 70.000 palavras. Foi publicado em 1899 e a história se passa no Rio de Janeiro

Algoritmo 10

```
1: if (A palavra é homógrafo do tipo 10) then
2:   if (P-1 = <não> ou <já>) ou (P-1 ou P-2 = <ainda> ou <nunca>) ou (P-1, P-2 ou P-3 = <tu>) ou
   (P+1 = HN, A_IND ou P_PESS_O_1) then
3:     V = [e]
4:   else
5:     V = [E]
6:   end if
7: else
8:   Vá para o algoritmo 11
9: end if
```

Algoritmo 11

```
1: if (A palavra é homógrafo do tipo 11) then
2:   if (P+1 = <ti>, <mim> ou <si>, HN, P_PESS_SU ou P_PESS_O_1) ou (P-1 = P_PESS_-
   SU ou P_PESS_O_1 e P+1 = A_IND) ou (P-1, P-2 ou P-3 = VERB ou VERB_IRR) ou (P-1
   = nc ou P_PESS_SU e P+1 ou P+2 = nc) then
3:     V = [o]
4:   else if (P-1 = P_PESS_SU, P_PESS_O_1 ou CS) ou (P-1 ou P-2 = <não> ou <nunca>) ou
   ((P-1 ou P-2 = <que> ou <ainda>) e (P+1 = A_IND)) ou (P+1 = PREPO, CONTR ou P_-
   PESS_O_1) then
5:     V = [O]
6:   else
7:     V = [o]
8:   end if
9: else
10:  Vá para o algoritmo 12
11: end if
```

Algoritmo 12

```
1: if (A palavra é homógrafo do tipo 12) then
2:   if (P-1 = <da>, <das>, <na>, <nas>, <pela>, <pelas> ou <em>) ou (P-1 ou P-2 = <a>, <uma>,
   <mesma>, <ortra>, <de>, <por>, P_DEM, P_POSS ou CONTR) ou (P+1 = CONTR) then
3:     V = [O]
4:   else if (P-1 = <toda>) ou (P-1 ends by <-mente>) ou (P-1 ou P-2 = nc) then
5:     V = [o]
6:   else
7:     V = [O]
8:   end if
9: else
10:  Vá para o algoritmo 13
11: end if
```

do Segundo Império, e conta a trajetória de Bentinho e Capitu [51]. É um romance psicológico, narrado em primeira pessoa por Bentinho, permitindo assim ter alguns dos homógrafos dos tipos 1 e 2, que são verbos, surgirem e testarem os algoritmos. O fato do texto ser histórico e literário nos leva a ter um conteúdo com uma linguística bem elaborada, servindo como um bom teste para os algoritmos de desambiguação. Por fim, neste teste foi identificado um total de 510 homógrafos, ou seja, 0,73% do texto total da base.

2.6.2 Resultados obtidos

De acordo com os textos apresentados na Subseção 2.6.1, foram realizados testes com as regras de desambiguação de homógrafos. Os resultados podem ser verificados

Algoritmo 13

```
1: if (A palavra é homógrafo do tipo 13) then
2:   if (O homógrafo pertence a BC_forma_o) ou (WN_forma_o existe em F0) ou (P-1 = <uma> e a
   palavra é <corte>) ou (P-1 = <um> e a palavra é <molho> ou <soco>) then
3:     V = [o]
4:   else if (P-1 ou P-2 = <a>, <uma>, <esta>, <qualquer>, P_IND, P_DEM, P_POSS, CONTR
   ou PREPO) ou (P+1 ou P+2 = ad) ou (O homógrafo pertence a BC_forma_O) then
5:     V = [O]
6:   else
7:     V = [O]
8:   end if
9: else
10:  Vá para o algoritmo 14
11: end if
```

Algoritmo 14

```
1: if (A palavra é homógrafo do tipo 14) then
2:   if (O homógrafo pertence a BC_cerca_e) ou (WN_cerca_e existe em F0) ou (P+2 ou P+3 =
   NUM) then
3:     V = [e]
4:   else if (P-1 = <uma>, <a>, CONTR ou PREPO) ou (P+2 = <madeira>, <arame>, <espinhos>)
   ou (<saltar> ou <levantar> existe em F0) ou (P+1 = ad) then
5:     V = [e]
6:   else if (P-1 ou P-2 = <que>, <não>, <ainda>, <já> ou <também>) ou (P-1 = <ele>, <ela> ou
   P_PESS_O_1) then
7:     V = [E]
8:   else
9:     V = [e]
10:  end if
11: else
12:  Vá para o algoritmo 15
13: end if
```

Algoritmo 15

```
1: if (A palavra é homógrafo do tipo 15) then
2:   if (P+1 = <em>, <no>, <na>, <nos>, <nas> ou <fogo>) ou (P-1 = <nunca>, <não>, <ainda>,
   <já>, <também>, <moda>, <se> ou CS) ou (P+1 = <ao> e P+2 = <colo>) then
3:     V = [E]
4:   else if (P-1 = <na>) ou (P+1 = <a>, <uma>, <outra>, <mesma>, P_DEM ou P_POSS) e (O
   homógrafo pertence a BC_pega_E) ou (WN_pega_E existe em F-1, F0 ou F+1) then
5:     V = [E]
6:   else if (P+1 = <a>, <uma>, <outra>, <mesma>, P_DEM ou P_POSS) e (O homógrafo pertence
   a BC_pega_e) ou (WN_pega_e existe em F-1, F0 ou F+1) then
7:     V = [e]
8:   else
9:     V = [E]
10:  end if
11: else
12:  Vá para o algoritmo 16
13: end if
```

nas Tabelas 2.5, 2.6 e 2.7. Podemos verificar que, apesar da variação no tamanho dos textos submetidos ao teste e do número de homógrafos encontrados, os resultados mostraram uma ligeira regularidade no percentual de acerto global do algoritmo proposto. É verificado também que apenas no teste realizado com o conteúdo de natureza jornalística, foram encontrados homógrafos pertencentes a todos os tipos listados neste trabalho.

Considerando os resultados obtidos, podemos notar que o resultado geral apresenta uma taxa de acerto com valor consideravelmente elevado (cerca de 99,00%

Algoritmo 16

```
1: if (A palavra é homógrafo do tipo 16) then
2:   if (P+1 = <senhor>, <que>, <qual>, <tua>, <teu>, <minha>, <meu>, <sua>, <seu>) ou (P-1 ou
   P-2 = nc) ou (P-1, P-2 ou P-3 = VERB ou VERB_IRR) then
3:     V = [e]
4:   else if (P-1 ou P-2 = <o(s)>, <um>, <uns>, <esse(s)>, <este(s)>, <aquele(s)>, <nesse(s)>,
   <desse(s)>, <deste(s)>, <daquele(s)>) then
5:     V = [e]
6:   else if (P-1 ou P-2 = <a(s)>, <uma(s)>, <essa(s)>, <esta(s)>, <aquela(s)>, <nessa(s)>, <dessa(s)>,
   <desta(s)>, <daquela(s)>) then
7:     V = [E]
8:   else if (P-1 ou P-2 = <eu>, <tu>, <ele>, <ela>) ou (P+1 = HN ou A_IND) then
9:     V = [E]
10:   else
11:     V = [e]
12:   end if
13: else
14:   Vá para o algoritmo 17
15: end if
```

Algoritmo 17

```
1: if (A palavra é homógrafo do tipo 17) then
2:   if (WN_besta_E existe em F-1, F0 ou F+1) then
3:     V = [E]
4:   else if (WN_besta_e existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_besta_e) then
5:     V = [e]
6:   else
7:     V = [E]
8:   end if
9: else
10:   Vá para o algoritmo 18
11: end if
```

Algoritmo 18

```
1: if (A palavra é homógrafo do tipo 18) then
2:   if (WN_sede_e existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_sede_e) then
3:     V = [e]
4:   else if (WN_sede_E existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_sede_E) then
5:     V = [E]
6:   else
7:     V = [E]
8:   end if
9: else
10:   Vá para o algoritmo 19
11: end if
```

em média), porém não podemos deixar de lado o fato de que ainda existem alguns homógrafos com poucas ocorrências ao longo do texto, que acabam ponderando suas taxas de acertos individuais para valores muito baixos, como verificado nas Tabelas 2.5, 2.6 e 2.7. Isto é um fator que demonstra a complexidade de solução do problema dos homógrafos.

Os erros ocorridos nos algoritmos dos tipos 1 e 2 apresentaram um índice elevado de erros quando seguidos, principalmente, por preposições ou contrações, antecedidos por formas verbais flexionadas.

Nos demais casos, podemos notar que os índices de acerto foram muito elevados ou máximos. Alguns erros que ainda assim ocorreram, podem ser justificados

Algoritmo 19

```
1: if (A palavra é homógrafo do tipo 19) then
2:   if (WN_medo_e existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_medo_e) then
3:     V = [e]
4:   else if (WN_medo_E existe em F-1, F0 ou F+1) then
5:     V = [E]
6:   else
7:     V = [e]
8:   end if
9: else
10:  Vá para o algoritmo 20
11: end if
```

Algoritmo 20

```
1: if (A palavra é homógrafo do tipo 20) then
2:   if (P-1 = <a> ou <as>) then
3:     V = [E]
4:   else if (P-1 = <os>, <aos>, <nos>, <em>, <desses>, <destes>, <nesses>, <daqueles>, <daqueles>,
<teus>, <seus>, <dos>, <cujos>, <meus>, <nestes>, <vossos>, <nossos>, <mesmos> ou <esses>)
then
5:     V = [e]
6:   else if (P-1 = <de> ou EXPIMP) ou (P+1 = HN, A_IND, P_POSS, P_DEM ou P_IND)
then
7:     V = [e]
8:   else
9:     V = [e]
10:  end if
11: else
12:  Vá para o algoritmo 21
13: end if
```

Algoritmo 21

```
1: if (A palavra é homógrafo do tipo 21) then
2:   if (WN_cor_o existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_cor_o) then
3:     V = [o]
4:   else if (O homógrafo pertence a BC_cor_O) then
5:     V = [O]
6:   else
7:     V = [O]
8:   end if
9: else
10:  Vá para o algoritmo 22
11: end if
```

por algumas falhas linguísticas existentes na base de dados ou pelo fato das frases estarem muito soltas ao longo do texto, requerendo uma análise mais elaborada, já que os algoritmos são baseados em regras linguísticas do PB. Boa parte do sucesso obtido ocorre devido ao fato de que se todas as respostas falharem, o sistema sai com a transcrição padrão. Estes resultados são bastante animadores, considerando que os homógrafos checados representam mais ou menos 1,0% das bases de dados utilizadas, o que significa que os homógrafos têm muito pouca expressão na base de dados, mas para o fim que se propõe, que é a síntese de voz, basta apenas que uma palavra seja transcrita de forma equivocada, que não soará agradavelmente a qualquer ouvinte. As naturezas das bases de dados também explicam porque alguns dos homógrafos não foram encontrados, como é o caso de b[o]to / b[O]to (algoritmo 8),

Algoritmo 22

```
1: if (A palavra é homógrafo do tipo 22) then
2:   if (WN_lobo_o existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_lobo_o) then
3:     V = [o]
4:   else if (WN_lobo_O existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_lobo_O) then
5:     V = [O]
6:   else
7:     V = [o]
8:   end if
9: else
10:  Vá para o algoritmo 23
11: end if
```

Algoritmo 23

```
1: if (A palavra é homógrafo do tipo 23) then
2:   if (WN_bola_O existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_bola_O) then
3:     V = [O]
4:   else if (WN_bola_o existe em F-1, F0 ou F+1) ou (O homógrafo pertence a BC_bola_o) then
5:     V = [o]
6:   else
7:     V = [O]
8:   end if
9: else
10:  Sair
11: end if
```

Tabela 2.5: Testes com o algoritmo proposto - CETEN-Folha.

Tipo	Ocorrência	Número de acertos	% de acerto
1	3409	3365	98,71%
2	3046	2965	97,34%
3	11	10	90,91%
4	95	90	94,74%
5	637	636	99,84%
6	482	471	97,72%
7	90	80	88,89%
8	5	5	100,00%
9	825	825	100,00%
10	169	169	100,00%
11	2335	2321	99,40%
12	47	45	95,74%
13	826	813	98,43%
14	866	863	99,65%
15	43	43	100,00%
16	6656	6652	99,94%
17	11	10	90,91%
18	148	141	95,27%
19	130	130	100,00%
20	108	101	93,52%
21	68	68	100,00%
22	39	39	100,00%
23	262	262	100,00%
TOTAL	20308	20105	99,00%

comum no PB, porém muito informal ou específico.

O resultado global é obtido como segue:

Tabela 2.6: Testes com o algoritmo proposto - Bíblia Cristã.

Tipo	Ocorrência	Número de acertos	% de acerto
1	209	205	98,09%
2	322	311	96,58%
3	5	4	80,00%
4	27	25	92,59%
5	333	321	96,40%
6	428	422	98,60%
7	61	56	91,80%
8	0	—	—
9	984	984	100,00%
10	5	4	80,00%
11	2740	2726	99,49%
12	11	10	90,91%
13	65	61	93,85%
14	51	49	96,08%
15	5	5	100,00%
16	2345	2344	99,96%
17	46	45	97,83%
18	107	97	90,65%
19	82	81	98,78%
20	60	58	96,67%
21	3	2	66,67%
22	14	14	100,00%
23	1	1	100,00%
TOTAL	7904	7825	99,00%

$$\begin{aligned}
\text{Resultado global} &= \frac{20105 + 7825 + 505}{20308 + 7904 + 510} * 100\% & (2.1) \\
&= 99,00\%
\end{aligned}$$

2.7 Conclusões e trabalhos futuros

Neste trabalho, apresentamos um conjunto de algoritmos baseados em regras linguísticas, separados por tipo, para desambiguação de homógrafos heterófonos aplicados a conversão texto-fala no PB. O método proposto é capaz de realizar a distinção entre 111 pares de homógrafos existentes no PB, organizados em 23 tipos de algoritmos de desambiguação, com a possibilidade de ter sua performance aumentada, a medida que novos homógrafos são inseridos na base de dados de homógrafos. O conjunto de algoritmos proposto foi implementado e testado com extratos de textos com naturezas distintas, atingindo uma taxa de acerto de 99,00% para o de natureza jornalística, 99,00% para o de natureza formal-religiosa e 99,02% no conteúdo com natureza histórico-literária. Estes resultados são bastante animadores, considerando a diversificação da natureza dos texto utilizados nos testes. Os algoritmos propos-

Tabela 2.7: Testes com o algoritmo proposto - Dom Casmurro.

Tipo	Ocorrência	Número de acertos	% de acerto
1	36	36	100,00%
2	73	72	98,63%
3	0	—	—
4	3	3	100,00%
5	30	30	100,00%
6	52	50	96,15%
7	6	6	100,00%
8	0	—	—
9	86	86	100,00%
10	0	—	—
11	35	35	100,00%
12	2	1	50,00%
13	5	5	100,00%
14	2	2	100,00%
15	1	1	100,00%
16	123	123	100,00%
17	7	7	100,00%
18	1	1	100,00%
19	22	22	100,00%
20	5	4	80,00%
21	17	17	100,00%
22	0	—	—
23	4	4	100,00%
TOTAL	510	505	99,02%

tos podem resolver a ambiguidade de vários pares de homógrafos existentes no PB. Como trabalhos futuros, desejamos conduzir os testes com extratos de textos ainda maiores e de diferentes naturezas, bem como avaliar a necessidade da inclusão ou exclusão de homógrafos, além de realizar comparações com métodos estatísticos.

Capítulo 3

Algoritmo para determinação de tonicidade

Este capítulo contém o detalhamento do marcador de tonicidade do nosso TTS, através das suas 19 regras linguísticas criadas com este propósito. Na Seção 3.1 apresentamos uma introdução do assunto, mostrando algumas pesquisas realizadas e a motivação em torno deste assunto. Na Seção 3.2 temos as regras utilizadas na determinação de tonicidade, bem como a metodologia aplicada. A Seção 3.3 contém os testes e resultados obtidos, bem como as suas análises e, finalmente, na Seção 3.4 temos nossas conclusões e trabalhos futuros.

3.1 Introdução

Tratando-se de sistemas artificiais, que vão transformar texto em fala audível, uma boa naturalidade dificilmente será alcançada se as características prosódicas não forem levadas em consideração. A prosódia tem como peculiaridade analisar o ritmo, a entonação e as demais características relacionadas à fala, descrevendo um conjunto de propriedades acústicas da fala que não tem como ser previstas pela ortografia tradicional. Neste sentido, a relação da tonicidade com a prosódia é muito forte.

A tonicidade vai representar a sílaba com pronúncia mais forte nas palavras (sílabas tônicas). A tonicidade da sílaba, eventualmente, pode ser sinalizada com acentuação gráfica ou não, conforme o idioma tratado. Não existe marcação no francês, ou alemão, dada a regularidade na pronúncia. No inglês, a tonicidade ocorre de forma inesperada, mas também não apresenta marcação gráfica. Já em idiomas como o espanhol e o português encontramos acentuação, porém, de acordo com as regras gramaticais, nem toda tonicidade é marcada com acentuação gráfica.

Visto que o assunto é complexo, tratando-se da determinação automática da

tonicidade em sistemas TTS, pesquisas têm sido desenvolvidas para que a qualidade da síntese seja melhorada, em diversos idiomas, através de uma determinação da tonicidade adequada. Em [55], temos um estudo onde a sílaba tônica é determinada em sinais acústicos de palavras isoladas, oriundos de uma base de reconhecimento de voz com amplo vocabulário. A determinação da sílaba tônica foi testada com 1600 palavras faladas por 11 locutores e uma taxa de erro de 2% foi alcançada. Também concernente ao inglês, em [10], um algoritmo para determinar o padrão de tonicidade em palavras polissílabas, foi implementado. Em [56] temos um estudo de tonicidade para promover a desambiguação existente em pares de palavras no inglês. No trabalho registrado em [57], um método para determinação da tonicidade no inglês falado na Nova Zelândia foi implementado. O método utiliza árvore de decisão e máquina de vetor suporte na classificação das vogais em tônicas ou átonas. Foram testadas locuções de 60 adultos do sexo feminino, com cerca de 700 vogais e uma precisão de quase 85% foi atingida. Em [58] foi desenvolvido um método para determinar a posição exata da sílaba tônica no romeno, já que este idioma, semelhante ao inglês, não possui uma posição tradicional da tonicidade. O método visa melhorar a qualidade da síntese e o algoritmo implementado fez uso de parâmetros morfológicos, fonéticos e léxicos, alcançando uma taxa de erro na detecção automática de cerca de 6%.

Considerando estudos recentemente realizados no PB, em [59] é analisado o uso de sílabas átonas como marcadores na verificação de tonicidade, através da frequência fundamental (F0) e da duração. Um estudo sobre o “stress” secundário no PB é apresentado em [60], bem como a verificação de “clicks” em posições tônicas da palavra. “Stress” secundário é o segundo mais forte na pronúncia de uma palavra, quando o idioma possui os dois graus de “stress”, sendo o mais forte o “stress” primário.

Considerando que podemos encontrar, no máximo, uma sílaba tônica em cada palavra encontrada no PB, a tarefa de marcar esta sílaba é de suma importância no uso dos conversores G2P, visto que, algumas regras para G2P utilizam a informação de vogal tônica [30]. Além do mais, analisando prosodicamente, a tonicidade está relacionada ao aumento da frequência fundamental, intensidade e duração da vogal. O ritmo da fala também é determinado pela transição entre tônicas e átonas [61, 62]. A Figura 3.1 destaca a tonicidade na frase: “Síntese de voz no português brasileiro”. Cada palavra apresenta sua sílaba tônica, dando ritmo e cadência à frase (S[í]ntese de v[o]z no Portugu[ê]s Brasil[e]iro). Qualquer deslocamento da marcação tônica diferente daquele determinado na frase, certamente vai prejudicar a qualidade da fala reproduzida.

Dada a importância deste assunto para a síntese de voz e, também pelo fato de não dispormos de trabalhos suficientes para o PB [30, 44, 45], Propomos um

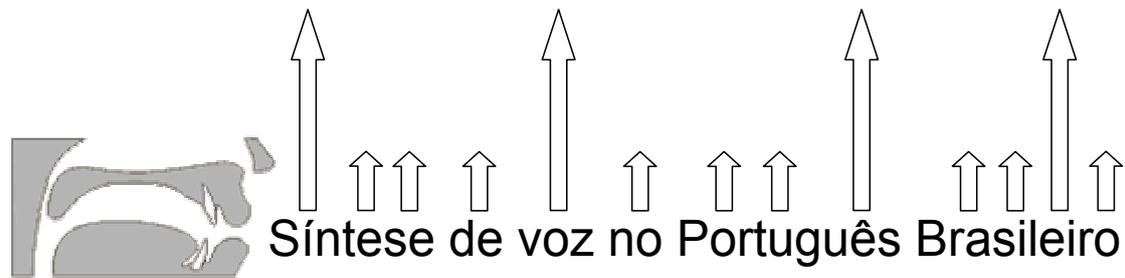


Figura 3.1: Determinação da tonicidade na frase “Síntese de voz no português brasileiro”.

algoritmo baseado em 19 (dezenove) regras, para determinação da sílaba tônica em palavras do português falado no Brasil. O algoritmo foi testado com um extrato do CETEN-Folha [49] composto de 15.974 palavras, atingindo uma taxa de acerto de 98,90%.

3.2 Conjunto de regras para determinação de tonicidade

Inicialmente, apresentamos o significado dos símbolos utilizados nas regras para determinação de tonicidade:

- O símbolo “ \wedge ” é um ponteiro para os caracteres das palavras, a partir do final. Por exemplo, $\wedge(0)$ é o último caractere de uma palavra, $\wedge(1)$ é o penúltimo caractere da mesma.
- O símbolo “T” indica a posição da vogal tônica em uma determinada palavra. Por exemplo, $T = 2$ significa que o antepenúltimo caractere é tônico.

As regras para a determinação da tonicidade são mostradas nas Tabelas 3.1, 3.2 e 3.3. O conjunto de regras apresentado foi inicialmente proposto em [30] com uma taxa de acerto de 98,53%, aqui obtivemos uma sensível melhora ao tratarmos algumas regras.

O algoritmo funciona mediante uma análise das sequências grafêmicas, partindo do último caractere, e das regras de acentuação gráfica da língua portuguesa [63]. As palavras são submetidas à sequência de regras, desde a Regra 1 até a Regra 19. Caso a palavra esteja enquadrada em uma das regras, ela tem a sua sílaba tônica marcada, passando, então, para a palavra seguinte.

A primeira regra já faz uma varredura na palavra buscando marcação gráfica prévia, obedecendo a questão da precedência. Algumas palavras, como homógrafos,

Tabela 3.1: Regras para determinação da tonicidade (da 1 até a 8)

Regra 1
Se existe acento, a vogal marcada é tônica. O acento agudo e o circunflexo tem precedência sobre o til. Exemplos: <i>órfão</i> , <i>órfã</i> <i>bênção</i> , <i>bênções</i> .
Regra 2
if $\hat{^}(0) = \{r, l, z, x, n\}$ then $T = 1$ end if Exemplos: <i>propor</i> , <i>carrossel</i> , <i>rapaz</i> , <i>triplex</i> , <i>durex</i> .
Regra 3
if $\hat{^}(0) = \{m\}$ and $\hat{^}(1) = \{i, o, u\}$ then $T = 1$ end if Exemplos: <i>pudim</i> , <i>bombom</i> , <i>comum</i> .
Regra 4
if $\hat{^}(0) = \{s\}$ and $\hat{^}(1) = \{n\}$ and $\hat{^}(2) = \{i, o, u\}$ then $T = 2$ end if Exemplos: <i>pudins</i> , <i>bombons</i> , <i>comuns</i> .
Regra 5
if $\hat{^}(0) = \{i\}$ and $\hat{^}(1) = \{u, ü\}$ and $\hat{^}(2) = \{q, g\}$ then $T = 0$ end if Exemplos: <i>caqui</i> , <i>aqui</i> , <i>sagüi</i> .
Regra 6
if $\hat{^}(0) = \{s\}$ and $\hat{^}(1) = \{i\}$ and $\hat{^}(2) = \{u, ü\}$ and $\hat{^}(3) = \{q, g\}$ then $T = 1$ end if Exemplos: <i>caquis</i> , <i>sagüis</i> .
Regra 7
if $\hat{^}(0) = \{i, u\}$ then if $\hat{^}(1)$ é vogal $\neq \{u\}$ $T = 1$ else $T = 0$ end if end if Exemplos: <i>caiu</i> , <i>grau</i> , <i>pneu</i> , <i>caju</i> , <i>javali</i> .
Regra 8
if $\hat{^}(0) = \{s\}$ and $\hat{^}(1) = \{i, u\}$ and $\hat{^}(2)$ é vogal then $T = 2$ end if Exemplos: <i>andais</i> , <i>pauis</i> , <i>graus</i> .

já chegam a este ponto com a definição da sua tonicidade. Isso antecipa a marcação tônica em até 1,0% das palavras do texto, se a desambiguação for realizada com precisão.

As palavras átonas são tratadas como exceção no contexto das regras por não possuírem tonicidade, tais como algumas preposições, contrações, pronomes, conjunções, etc.

As dez primeiras regras buscam identificar marcação de tonicidade em palavras oxítonas¹. Da Regra 11 até a Regra 17, são tratadas as marcações de palavras paroxítonas². A Regra 10 trata de uma exceção da Regra 11, já que “porque” é oxítona no PB. A Regra 18 também é uma exceção da Regra 19, para que o <u>

¹A última sílaba é a tônica

²A penúltima sílaba é a tônica

Tabela 3.2: Regras para determinação da tonicidade (da 9 até a 15)

<p>Regra 9</p> <p>if $\hat{(0)} = \{s\}$ and $\hat{(1)} = \{i,u\}$ and $\hat{(2)}$ não é vogal then $T = 1$ end if</p> <p>Exemplos: <i>cajus, javalis</i>.</p>
<p>Regra 10</p> <p>if $\hat{(0)} = \{e\}$ and $\hat{(1)} = \{u\}$ and $\hat{(2)} = \{q\}$ and $\hat{(3)} = \{r\}$ and $\hat{(4)} = \{o\}$ and $\hat{(5)} = \{p\}$ then $T = 0$ end if</p> <p>Exemplos: <i>porque</i>.</p>
<p>Regra 11</p> <p>if $\hat{(0)} = \{e\}$ and $\hat{(1)} = \{u\}$ and $\hat{(2)} = \{q,g\}$ then if $\hat{(3)}$ é vogal $T = 3$ else $T = 4$ end if</p> <p>Exemplos: <i>alambique, Henrique, destaque, bosque</i>.</p>
<p>Regra 12</p> <p>if $\hat{(0)} = \{s\}$ and $\hat{(1)} = \{e\}$ and $\hat{(2)} = \{u\}$ and $\hat{(3)} = \{q,g\}$ then if $\hat{(4)}$ é vogal $T = 4$ else $T = 5$ end if</p> <p>Exemplos: <i>alambiques, Henriques, destaques</i>.</p>
<p>Regra 13</p> <p>if $\hat{(0)}$ é vogal and $\hat{(1)} = \{i,u\}$ and $\hat{(2)}$ é vogal then $T = 2$ end if</p> <p>Exemplos: <i>meia, seio</i>.</p>
<p>Regra 14</p> <p>if $\hat{(0)}$ and $\hat{(3)}$ são vogais and $\hat{(2)} = \{i,u\}$ and $\hat{(1)}$ não é vogal and $\hat{(4)} \neq \{q,g\}$ then $T = 3$ end if</p> <p>Exemplos: <i>cadeira, queima, louco</i>.</p>
<p>Regra 15</p> <p>if $\hat{(0)} = \{s\}$ and $\hat{(1)}$ and $\hat{(4)}$ são vogais and $\hat{(3)} = \{i,u\}$ and $\hat{(2)}$ não são vogais and $\hat{(5)} \neq \{q,g\}$ then $T = 4$ end if</p> <p>Exemplos: <i>cadeiras, queimas, loucos</i>.</p>

não seja equivocadamente marcado no lugar do <e>, na palavra “quem”.

As Regras 4, 6, 9, 12 e 15 vão determinar a tonicidade de palavras no plural, até então, associadas a regras anteriores, quando no singular.

O conjunto de regras apresentado foi inicialmente proposto em [30] com uma taxa de acerto de 98,53%. Nesta tese obtivemos uma sensível melhora ao tratarmos algumas daquelas regras. Naquela ocasião, o algoritmo proposto apresentava um conjunto de 20 (vinte) regras. As melhorias implementadas para este trabalho estão relacionadas com a fusão das Regras 7 e 8 de [30] (aqui, Regra 7); um ajuste nas Regras 12 e 13 (aqui, Regras 11 e 12), com uma expansão para o caso de termos $\hat{(2)}$ ou $\hat{(3)} = \{g\}$. Outra providência foi ajustar a Regra 16 (aqui, Regra 15), introduzindo a exceção para $\hat{(5)} \neq \{q,g\}$. Estas melhorias proporcionaram um aumento na taxa de acerto considerável (de 98,53% para 98,90%), com um extrato de texto três vezes maior do que o testado em [30].

Tabela 3.3: Regras para determinação da tonicidade (da 16 até a 19).

<p>Regra 16</p> <p>if $\hat{^}(0) = \{a,e,o\}$ and $\hat{^}(1)$ é consoante and $\hat{^}(2) = \{n\}$ and $\hat{^}(3) = \{i,u\}$ and $\hat{^}(4)$ é vogal then</p> <p style="padding-left: 2em;">$T = 3$</p> <p>end if</p> <p>Exemplos: <i>ainda, caíndo, fluindo, incluindo.</i></p>
<p>Regra 17</p> <p>if $\hat{^}(k)$ = penúltima vogal and $\hat{^}(k) = \{i,u\}$ and $\hat{^}(k+1)$ é vogal and $\hat{^}(k-1)$ não é vogal and $\hat{^}(k+2)$ não é $\{q,g\}$ then</p> <p style="padding-left: 2em;">$T = k+1$</p> <p>end if</p> <p>Exemplos: <i>outro, claustró.</i></p>
<p>Regra 18</p> <p>if $\hat{^}(0) = \{m\}$ and $\hat{^}(1) = \{e\}$ and $\hat{^}(2) = \{u\}$ and $\hat{^}(3) = \{q\}$ then</p> <p style="padding-left: 2em;">$T = 1$</p> <p>end if</p> <p>Exemplos: <i>quem.</i></p>
<p>Regra 19</p> <p>$T =$ penúltima vogal da palavra</p> <p>Exemplos: <i>casa, homem, guerra.</i></p>

3.3 Resultados obtidos

As regras para determinação da sílaba tônica foram testadas com um extrato de texto da base de dados CETEN-Folha [49], composto de 15.974 palavras. Os resultados obtidos podem ser verificados na Tabela 3.4, onde pôde ser alcançada uma taxa de acerto global de 98,90%.

Diante destes resultados, podemos constatar a boa performance do marcador de tonicidade proposto. Além disso, dada a natureza do texto, podemos encontrar uma variedade muito grande de palavras, que tornam os algoritmos bastante funcionais.

Os erros ocorridos, relacionados a nomes próprios, justificam-se pelo fato de que muitas vezes esses nomes não estão de acordo com as regras vigentes no PB, como por exemplo “Sônia”, que muitas vezes é grafado equivocadamente como “Sonia”, determinando uma falsa informação de tonicidade (Son|i)a). Além do mais, verificamos que a ocorrência deste tipo de erro, em número acentuado, foi devido ao fato de que um conteúdo jornalístico apresenta diversas informações e notícias relacionadas às pessoas e de forma repetitiva.

Erros de estrangeirismos são justificáveis por si só, já que dificilmente estarão enquadrados nas regras linguísticas.

Por fim, os acrônimos, que apesar de apresentarem ocorrência reduzida, muitas vezes, mesmo sendo palavras silabadas³, também não ficam enquadradas no contexto das regras. Podemos citar como exemplos, para este caso, as Siglas “ABRAVEST” e “CRECI”, que fonologicamente possuem marcação tônica como “ABRAV[E]ST” (Associação Brasileira do Vestuário) e “CR[E]CI” (Conselho Regional de Corretores de Imóveis), pela ortografia das regras propostas, tem a sua tonicidade definida com

³Condição em que a palavra pode ser pronunciada como uma sequência de sílabas.

Tabela 3.4: Tabela com resultados na determinação de tonicidade.

Tipo de erro	Erros encontrados	% de acerto
Nomes próprios	86	99,46% (parcial)
Estrangeirismos	53	99,67% (parcial)
Acrônimos	20	99,87% (parcial)
Falha na marcação	16	99,90% (parcial)
TOTAL	175	98,90% (global)

“ABR[A]VEST” e “CREC[I]”.

As falhas de marcação tiveram um índice baixíssimo, mostrando que os algoritmos tem uma performance boa, considerando o tamanho do teste, com quase 16.000 palavras.

Diante destes resultados, o conjunto de regras proposto pode ser considerado uma boa escolha para a marcação de tonicidade em sistemas TTS no PB. Além disso, dada a natureza do texto, podemos encontrar uma variedade ampla de palavras, que tornam os algoritmos bastante funcionais.

3.4 Conclusões e trabalhos futuros

Neste trabalho, apresentamos um conjunto de algoritmos baseados em regras linguísticas, num total de 19 (dezenove) regras, capazes de realizar a identificação da vogal tônica de palavras do português brasileiro. O algoritmo proposto foi implementado e testado com um extrato de texto da base de dados CETEN-Folha, composto de 15.974 palavras e uma taxa de acerto de 98,90% foi atingida, superando os resultados apresentados em [30], ao serem trabalhadas as Regras 7, 8, 12, 13 e 16, que proporcionaram um aumento na taxa de acerto com um texto três vezes maior do que o testado anteriormente.

Estes resultados são bastante satisfatórios, considerando que o erro real do conjunto de regras implementado pode ser reduzido a 0,1%, se forem considerados apenas as falhas na marcação. Como trabalhos futuros, desejamos isolar os erros de marcação ocorridos e ajustar ou criar novas regras, de forma que tenhamos uma eficiência ainda maior do marcador de tonicidade, aprimorando indiretamente o G2P, além do mais pretendemos realizar novos testes com textos de natureza não-jornalística, a fim de que possamos evitar o excesso de nomes próprios e acrônimos.

Capítulo 4

Algoritmo para separação das sílabas

Neste capítulo são apresentadas as características do algoritmo de separação das sílabas nas suas 25 regras linguísticas desenvolvidas. Na Seção 4.1, apresentamos uma introdução do assunto e sua importância no âmbito das tecnologias de voz. A Seção 4.2 contém alguns comentários relativos ao papel da sílaba na entonação, especificando o PB. As regras linguísticas para separação das sílabas são apresentadas na Seção 4.3, bem como a metodologia utilizada na implementação do algoritmo. A Seção 4.4 segue com a apresentação dos testes realizados e resultados obtidos. Por fim, na Seção 4.5, apresentamos as conclusões e trabalhos futuros relativos à separação silábica.

4.1 Introdução

Tem sido reportado através de várias pesquisas, seja no âmbito do português europeu [62, 64, 65, 66], ou brasileiro [59, 67, 68], bem como em outros idiomas [33, 69], que há uma importância significativa no mapeamento das sílabas quando tratamos de TTS, seja para transcrição fonética, como também para a geração de prosódia, já que numa sílaba tônica há um aumento na sua duração e intensidade. O mapeamento das sílabas, juntamente com a definição da sílaba tônica, são importantes para a verificação da alternância entre as sílabas tônicas e átonas nas palavras do português. Essa alternância vai ditar o ritmo da fala em sistemas TTS. Outro ponto importante é que a identificação da sílaba tônica pode solucionar, fonologicamente, a ambiguidade de palavras como “para”, “pára” e “Pará”, situação esta que traz ainda consequências morfosintáticas sobre a palavra, já que haverá uma mudança na classe gramatical [30, 37].

4.2 Papel da sílaba na entonação

Desde a sua infância e de forma imutável até a fase adulta, o ser humano desenvolve o ritmo da sua fala como uma das principais características do seu idioma.

Concernete ao ritmo aplicado em cada idioma, eles podem ser classificados como *syllable-timed* e *stress-timed* [70, 71].

Nos idiomas classificados como *syllable-timed*, a sílaba é a unidade responsável por dar o ritmo da fala e cada sílaba tem a sua pronúncia realizada num intervalo de tempo com durações semelhantes. Dessa forma, fica claro que a duração da pronúncia de uma frase está diretamente relacionada com o número de sílabas, sendo muito raro o surgimento de contrações (japonês, francês e o PB), onde cada sílaba tem uma pronúncia clara e distinta, determinando o fluxo de sons.

Já nos idiomas do tipo *stress-timed*, como o inglês e o russo, as sílabas tônicas de determinadas palavras marcam o ritmo da fala. Essas sílabas tônicas ocorrem em intervalos aparentemente irregulares quando atentamos para a frase escrita. Os trechos de algumas sub-palavras atônicas intermediárias ficam comprimidas e aglutinadas, chegando até ao quase desaparecimento de algumas delas. Diferentemente do tipo anterior, nos idiomas classificados como *stress-timed*, o tempo de duração para pronunciar uma frase em inglês ou russo vai depender do número de sílabas que recebem marcação tônica e não do número total de sílabas.

Justificamos o uso de regras linguísticas para divisão das palavras em unidades silábicas, aproveitando a condição de que há uma regularidade no ritmo da fala existente no Português falado no Brasil. Ritmo este cadenciado pelas sílabas, que são a base para a formação das palavras no nosso idioma, ou seja, toda palavra vai possuir pelo menos uma sílaba e toda sílaba possui uma vogal como base [63]. Além de já ter sido comprovado que o uso de regras linguísticas é uma boa escolha para a conversão texto-fala no português [37], já que uma economia computacional é gerada pela hierarquia existente, bem como pela facilidade de podermos agregar novas entradas, rapidamente, melhorando o processamento.

4.3 Conjunto de regras para separação das sílabas

Nesta Seção, apresentamos o conjunto, composto por 25 regras, utilizado para realizar a separação em sílabas das palavras processadas no texto. Para uma fácil compreensão das regras, estamos dispendo na Tabela 4.1, a simbologia inserida na descrição das regras de separação.

As regras estão baseadas na busca das vogais existentes na palavra, seguida de análise dos caracteres existentes à esquerda e à direita, para então decidir qual a ação a ser tomada. Como as vogais são a base da sílaba, a decisão ou ação tomada

vai ser dada no sentido de unir ou não a referida vogal ao conjunto de grafemas já separados. Uma nova sílaba é formada caso os grafemas já separados possuam uma vogal. As seis primeiras regras tratam de vogais abrindo sílabas, os demais casos vão analisar sílabas com vogais posicionadas internamente ou como terminais. O conjunto de ações está listado na Tabela 4.2. As regras linguísticas são apresentadas nos Algoritmos das Tabelas 4.3, 4.4, 4.5 e 4.6. Os algoritmos foram desenvolvidos baseados na ortografia, mas prevalecendo a fonologia das palavras, na tentativa de conciliar as teorias fonológicas da língua, com as necessidades técnicas dos sistemas TTS, como é o caso das Regras 2, 5, e 17, onde temos situações em que a fonologia prevalece em palavras como “obstruir” (*o-bs-tru-ir*), “arredondar” (*a-rre-don-dar*), “assado” (*a-ssa-do*), “sócio” (*só-cio*), etc.

As regras são dispostas hierarquicamente desde a primeira até a Regra 25. Este procedimento faz com que a carga computacional não seja muito grande, fator importante em sistemas que requerem uma resposta rápida, uma vez que, tendo sido a vogal analisada, enquadrada em uma das regras, passa-se imediatamente para a vogal seguinte. Além do mais, ter a versatilidade de manipular regras é também importante, pois elas podem ser excluídas, incluídas ou modificadas rapidamente, de acordo com as necessidades técnicas e linguísticas.

Tabela 4.1: Símbolos usados nas regras de separação das sílabas.

Símbolo	Significado
V	Vogal (a, e, o, á, é, ó, í, ú, ã, õ, â, ê, ô, à, ü)
V_t	Vogal tônica
G	Semivogal (i, u)
C	Uma consoante qualquer (<lh>, <nh>, CO , CF , CL , CN)
CO	Oclusivas (p; t; c+a, o, u; qu+e, i; b; d; g+a, o, u; gu+e, i)
CF	Fricativas (f; v; s; c+e, i; ç; z; ss; ch; j; g+e, i; x)
CL	Líquida (l; r; rr exceto <lh>)
CN	Nasal (m; n)
SP	Espaço
F	Final de linha ou final do arquivo
p⁰	Início da sílaba
^(0)	Posição do grafema analisado
^(+1) = C	Primeiro grafema à direita da vogal é uma consoante qualquer
^(+2) = G	Segundo grafema à direita da vogal é semivogal
^(+3) = V	Terceiro grafema à direita da vogal é vogal
^(-1) ≠ CN	Primeiro grafema à esquerda da vogal não é consoante nasal
^(-1) = CO	Primeiro grafema à esquerda da vogal é oclusiva

4.4 Resultados obtidos

As regras para separação das palavras em sílabas foram testadas com um extrato de texto da base de dados CETEN-Folha, composto de 18.037 sílabas. Os resultados

Tabela 4.2: Casos e ações tomados pelo algoritmo.

Caso	Ação
1	V é separada do grafema seguinte
2	V é unida ao próximo grafema da direita e é separada do subsequentes.
3	V é unida ao grafema anterior e é separada dos seguintes.
4	V é unida ao grafema anterior e ao seguinte e é separada dos subsequentes.
5	V é unida aos dois grafemas seguintes e é separada do terceiro grafema.
6	V é unida ao grafema anterior e a todos os grafemas seguintes até o final da palavra.
7	V é unida ao grafema anterior e aos dois grafemas seguintes e separa-se dos subsequentes.
8	V é unida aos dois grafemas anteriores e separa-se do seguinte.

Tabela 4.3: Regras utilizadas na separação das sílabas (de 1 a 5).

<p>Regra 1</p> <p>if $V = p^0$ and $\wedge(0) \neq \langle \tilde{a} \rangle, \langle \tilde{o} \rangle$ and $\wedge(+1) = V$ and $\wedge(+1) \neq G$ then if $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) = SP, F$ then Caso 5 else Caso 1 end if end if Exemplo: <i>a-eronave, a-inda</i></p>
<p>Regra 2</p> <p>if $V = p^0$ and $\wedge(+1) = C$ and $\wedge(+2) = C$ and $\wedge(+3) = CO$ then Caso 5 end if Exemplo: <i>o-bstar, a-dstrito</i></p>
<p>Regra 3</p> <p>if $V = p^0$ and $\wedge(+1) = G, CN, \langle s \rangle, \langle r \rangle, \langle l \rangle, \langle x \rangle$ and $\wedge(+2) = C$ then if $\wedge(+2) \neq \langle s \rangle, \langle h \rangle$ and $\wedge(+1) \neq \wedge(+2)$ then Caso 2 else if $\wedge(+1) = CN$ and $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) \neq V$ then Caso 7 else if $\wedge(+1) = \wedge(+2)$ or $\wedge(+2) = \langle h \rangle$ then Caso 1 else if $\wedge(+2) = \langle s \rangle$ and $(\wedge(+3) = C$ and $\wedge(+3) \neq \langle s \rangle)$ or $\wedge(+3) \neq C, V$ then Caso 7 else Caso 2 end if end if Exemplo: <i>am-bos, an-te, as-pas, al-tura, ar-gúcia, eu-ropa</i></p>
<p>Regra 4</p> <p>if $V = p^0$ and $\wedge(+1) = CO, CF, \langle g \rangle, \langle p \rangle$ and $\wedge(+2) = CO, CF, CN, \langle \zeta \rangle$ and $\wedge(+3) = V$ then if $\wedge(+1) = \wedge(+2)$ then Caso 1 else Caso 2 end if end if Exemplo: <i>as-tral, ex-por, ei-ra, ai-po, o-ptar, a-dvogar, a-gnóstico, a-florar, a-fta</i></p>
<p>Regra 5</p> <p>if $V = p^0$ and $\wedge(+1) = C$ and $\wedge(+2) = V, CL, \langle h \rangle$ then Caso 1 end if Exemplo: <i>a-rrendar, a-tlas, a-lho, a-mor, a-clamado</i></p>

obtidos podem ser verificados na Tabela 4.8, onde pôde ser alcançada uma taxa de acerto de 99,41%.

Testar o separador silábico com um amplo conjunto de sílabas pôde mostrar que o algoritmo apresenta uma boa performance ao realizar a separação em sílabas de palavras sujeitas às regras linguísticas do Português falado no Brasil. Já era

Tabela 4.4: Regras utilizadas na separação das sílabas (de 6 a 11).

Regra 6
if $V \neq p^0$ and $\wedge(+1) = G$ and $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) = CO$ then
Caso 5
end if
Exemplo: <i>Áus-tria</i>
Regra 7
if $V \neq p^0$ and $(\wedge(-1) = C, \{\langle u \rangle, \langle ü \rangle$ seguido de $\langle g \rangle, \langle q \rangle\}$ and $\wedge(+1) = C$ and $\wedge(+2) = V$ then
Caso 3
end if
Exemplo: <i>ca-lha, ca-la, me-ta, ca-choeira</i>
Regra 8
if $V \neq p^0$ and $\wedge(-1) = C$ and $\wedge(+1) = G$ and $\wedge(+2) = \langle r \rangle$ and $\wedge(+3) = C$ then
if $\wedge(0) = V_t$ then
Caso 4
else
Caso 3
end if
end if
Exemplo: <i>ca-irmos</i>
Regra 9
if $V \neq p^0$ and $\wedge(-1) = C$ and $\wedge(+1) = G, CN$ and $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) = CO$ then
Caso 7
end if
Exemplo: <i>cons-tante</i>
Regra 10
if $V \neq p^0$ and $\wedge(-1) = C, G$ and $\wedge(+1) = \langle i \rangle, \langle u \rangle, \langle e \rangle, \langle o \rangle$ and $\wedge(+1) \neq V_t$ and $\wedge(0) \neq \wedge(+1)$ and $\wedge(+2) = C$ seguido de C, V and $\wedge(+2) \neq \langle s \rangle$ then
if $\wedge(0) = V_t$ and $\wedge(+2) \neq \langle n \rangle$ and $\wedge(+3) \neq C$ then
Caso 4
else if $\wedge(0) \neq V_t$ and $\wedge(+1) = \langle i \rangle$ and $\wedge(+2) \neq \langle n \rangle$ then
Caso 2
else if $(\wedge(+1) \neq \langle i \rangle$ and $(\wedge(+2) = CN, \langle r \rangle$ and $\wedge(+3) \neq \langle h \rangle, V_t$ or $\wedge(0) = \langle a \rangle, \langle e \rangle, \langle o \rangle$ and $\wedge(+1) = \langle a \rangle, \langle e \rangle, \langle o \rangle$ or $\wedge(+2) = CN$ and $\wedge(+3) \neq \langle h \rangle, \langle s \rangle$ and $\wedge(+4) = V, C$ then
Caso 1
else
Caso 4
end if
end if
Exemplo: <i>condicio-nador</i>
Regra 11
if $V \neq p^0$ and $\wedge(-1) = C$ and $\wedge(+1) = G$ and $\wedge(+2) = V, SP$ then
Caso 4
end if
Exemplo: <i>prai-a, mei-a</i>

esperado que houvesse um número razoável de erros decorrentes de acrônimos e estrangeirismos, pois um conteúdo jornalístico, normalmente, apresenta muitos casos deste tipo, e por eles não seguirem as regras linguísticas vigentes. Por exemplo, “ABRAVEST” (*Associação Brasileira do Vestuário*), que teve a sua última consoante isolada e formando sílaba (*A-BRA-VES-T*), o que não ocorre no PB, pois toda sílaba, obrigatoriamente deve possuir uma vogal. Fato que não ocorreu com palavras estruturalmente semelhantes, com é o caso de “FAPERJ” (*Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro*), que teve a sua separação como esperada (*FA-PERJ*). As ocorrências relativas aos acrônimos foram pequenas, não impactando na robustez do algoritmo, mas necessitando de análise e ajuste para não prejudicar a síntese de voz. No caso de estrangeirismos, as incorreções foram ainda mais explícitas (*Was-hington, sus-his, im-pe-achment, Dai-hatsu, Re-nau-lt*, etc),

Tabela 4.5: Regras utilizadas na separação das sílabas (de 12 a 15).

<p>Regra 12</p> <p>if $V \neq p^0$ and $\wedge(-1) = C$ and $\wedge(0) = G$ and $\wedge(+1) = V$ and $\wedge(0) \neq \wedge(+1)$ and $\wedge(+2) = C$ and $\wedge(+3) = V$ then</p> <p> if $\wedge(+2) = \langle \zeta \rangle$ and $\wedge(+3) = \langle \tilde{a} \rangle, \langle \tilde{o} \rangle$ or $\wedge(+1) = V_t$ or $\wedge(-1) = \langle r \rangle$ and $\wedge(+3) = V_t$ then</p> <p> Caso 1</p> <p> else</p> <p> Caso 8</p> <p> end if</p> <p>end if</p> <p>Exemplo: <i>correligio-nários</i></p>
<p>Regra 13</p> <p>if $V \neq p^0$ and $\wedge(-1) = C, \{ \langle u \rangle, \langle \tilde{u} \rangle$ antecidos de $\langle q \rangle, \langle g \rangle \}$ and $\wedge(+1) = CL, CN, \langle c \rangle, \langle x \rangle$ and $\wedge(+2) = C$ and $\wedge(+3) = V, \langle h \rangle, \langle l \rangle, \langle r \rangle$ then</p> <p> if $\wedge(+2) = \langle h \rangle, \langle l \rangle, \langle r \rangle$ then</p> <p> if $\wedge(+1) = \wedge(+2)$ or $\wedge(+1) = \langle c \rangle, \langle l \rangle, \{ \langle n \rangle$ seguido de $\langle h \rangle \}$ then</p> <p> Caso 1</p> <p> else</p> <p> Caso 4</p> <p> end if</p> <p> end if</p> <p>end if</p> <p>Exemplo: <i>gue-rra</i></p>
<p>Regra 14</p> <p>if $V \neq p^0$ and $\wedge(-1) = C$ and $\wedge(+1) = CL, CN, \langle i \rangle$ and $\wedge(+2) = \langle s \rangle$ then</p> <p> if $\wedge(+3) \neq C, V$ then</p> <p> Caso 6</p> <p> else if $\wedge(0) = V_t$ or $\wedge(+3) = V$ then</p> <p> Caso 4</p> <p> else</p> <p> Caso 5</p> <p> end if</p> <p>end if</p> <p>Exemplo: <i>va-ral, cais</i></p>
<p>Regra 15</p> <p>if $V \neq p^0$ and $\wedge(+1) = V$ igual or $\wedge(+1) = V$ and $\wedge(+2) = V$ then</p> <p> if $\wedge(0) = V_t$ and $\wedge(+1) = G$ and $\wedge(+3) = C$ then</p> <p> Caso 2</p> <p> else</p> <p> Caso 1</p> <p> end if</p> <p>end if</p> <p>Exemplo: <i>co-operação, bu-eiro</i></p>

fator somente solucionável após a integração de um leitor de estrangeirismos. As falhas na separação ocorreram de forma significativa, apesar de termos alcançado um percentual de acerto parcial considerável (99,69%). Estas falhas mostraram que o algoritmo ainda pode ser mais refinado, apenas tratando algumas regras, como foi verificado em relação a [72].

4.5 Conclusões e trabalhos futuros

Nesta tese, apresentamos um algoritmo baseado em regras linguísticas, num total de 25 regras, capaz de realizar a separação em sílabas de palavras do PB. O algoritmo proposto foi implementado e testado com um extrato de texto da base de dados CETEN-Folha, composto por 18.037 sílabas e uma taxa de acerto global de 99,41% foi atingida. Considerando que o erro real do conjunto de regras implementado foi

Tabela 4.6: Regras utilizadas na separação das sílabas (de 16 a 21).

<p>Regra 16 if $V \neq p^0$ and $\hat{(-1)} = C$ and $\hat{(+1)} = V$ and $\hat{(+2)} = CN$ then Caso 3 end if Exemplo: <i>ci-ência</i></p>
<p>Regra 17 if $V \neq p^0$ and $\hat{(0)} = \langle i \rangle$ and $\hat{(-2)} = \langle \acute{a} \rangle, \langle \acute{e} \rangle, \langle \acute{i} \rangle, \langle \acute{o} \rangle, \langle \acute{u} \rangle$ and $\hat{(-1)} = C$ and $\hat{(+1)} = \langle a \rangle, \langle o \rangle$ then Caso 6 end if Exemplo: <i>palá-cio, proprietá-rio</i></p>
<p>Regra 18 if $V \neq p^0$ and $V = \langle \tilde{a} \rangle, \langle \tilde{o} \rangle$ and $\hat{(-1)} = C$ and $\hat{(+1)} = \langle o \rangle, \langle e \rangle$ then Caso 6 end if Exemplo: <i>gamão, limões</i></p>
<p>Regra 19 if $V \neq p^0$ and $\hat{(-1)} = C$ and $\hat{(+1)} = V_t$ and $\hat{(+1)} \neq \langle i \rangle, \langle u \rangle$ then Caso 3 end if Exemplo: <i>ali-ás, juni-or</i></p>
<p>Regra 20 if $V \neq p^0$ and $\hat{(-1)} = C$ and $\hat{(+1)} = V$ and $\hat{(+2)} = CN$ and $\hat{(+3)} = C$ then Caso 7 end if Exemplo: <i>cinquen-ta, sequên-cia</i></p>
<p>Regra 21 if $V \neq p^0$ and $\hat{(+1)} = CO, \langle f \rangle, \langle v \rangle, \langle g \rangle$ and $\hat{(+2)} = CL, CO$ and $\hat{(+3)} = V$ then if $\hat{(+1)} = \langle f \rangle, \langle p \rangle$ and $\hat{(+2)} = \langle t \rangle, \langle \zeta \rangle$ then Caso 2 else Caso 1 end if end if Exemplo: <i>su-blime, ra-pto</i></p>

menor do que 0,5% (relativos às falhas de separação), podemos avaliar o algoritmo como adequado, para o fim que se propõe, apesar de ainda vislumbrarmos que os erros podem ainda ser reduzidos, através do tratamento de algumas regras, bem como a manipulação de palavras estrangeiras, sendo estas tarefas, parte dos nossos trabalhos futuros.

Tabela 4.7: Regras utilizadas na separação das sílabas (de 22 a 25).

Regra 22

if $V \neq p^0$ and $\wedge(-1) = C, \langle u \rangle$ antecedido de $\langle q \rangle, \langle g \rangle$ and $\wedge(+1) = V$ and $\wedge(+2) = C, SP, F$ **then**
 if $\wedge(0) = \langle i \rangle, \langle u \rangle, \langle i \rangle, \langle ú \rangle, \langle é \rangle, \langle ê \rangle$ and $\wedge(0) = V_t$ and $\wedge(+1) \neq \langle u \rangle$ or $\wedge(0) \neq G$ and $\wedge(+2) = \langle s \rangle$
 and $\wedge(+3) \neq C, V$ **then**
 Caso 3
 else if $\wedge(0) = \langle i \rangle$ and $\wedge(0) = V_t$ and $\wedge(+1) = \langle u \rangle$ and $\wedge(+2) = SP, F$ **then**
 Caso 4
 else if $\wedge(0) = G$ and $\wedge(+1) \neq V_t$ and $\wedge(+2) \neq C, V$ or $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) \neq C, V$ or $\wedge(0) \neq V_t$ and
 $\wedge(+1) \neq V_t$ and $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) = SP, F$ **then**
 Caso 6
 else if $\wedge(0) = \langle i \rangle, \langle u \rangle$ and $\wedge(+1) \neq V_t$ or $\langle ã \rangle$ and $\wedge(0) \neq \langle i \rangle, \langle u \rangle$ and $\wedge(0) \neq V_t$ and $\wedge(+1) \neq \langle i \rangle, \langle u \rangle$
 and $\wedge(+1) \neq V_t$ or $\wedge(-1) = \langle u \rangle$ antecedido de $\langle q \rangle, \langle g \rangle$ **then**
 Caso 1
 else if $\wedge(+1) = \langle o \rangle, \langle u \rangle$ and $\wedge(+1) \neq V_t$ and $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) = SP, F$ **then**
 Caso 7
 else
 Caso 2
 end if
end if
 Exemplo: *polici-a, sóci-o*

Regra 23

if $V \neq p^0$ and $\wedge(-1) = C, \langle u \rangle$ antecedido de $\langle q \rangle$ and $\wedge(+1) = C$ and $\wedge(+2) = C$ **then**
 if $\wedge(+1) = \wedge(+2)$ **then**
 Caso 1
 else if $\wedge(+1) = \langle s \rangle$ and $\wedge(+2) \neq \langle s \rangle$ **then**
 Caso 2
 else if $\wedge(+2) = \langle s \rangle$ and $\wedge(+3) = CO$ **then**
 Caso 5
 end if
end if
 Exemplo: *bu-rro, bes-ta*

Regra 24

if $V \neq p^0$ and $\wedge(+1) = C$ and $\wedge(+2) = C$ **then**
 Caso 2
end if
 Exemplo: *dis-cretizar*

Regra 25

Se nenhum dos casos anteriores for verificado e a palavra terminar, **V** forma sílaba com os grafemas que restarem até o espaço em branco, sinal de pontuação ou hífen.
 Exemplo: *des-cer*

Tabela 4.8: Tabela com resultados na separação das sílabas.

Tipo de erro	Erros encontrados	% de erro
Estrangeirismos	43	99,76% (parcial)
Acrônimos	8	99,96% (parcial)
Falha na separação	56	99,69% (parcial)
TOTAL	107	99,41% (global)

Capítulo 5

Algoritmo para transcrição fonética

Este capítulo contém o detalhamento do conversor G2P, suas características, as regras linguísticas e testes realizados. Na Seção 5.1, temos uma breve introdução do assunto, bem como sua importância no processo de sintetização da fala. A Seção 5.2 contém os comentários sobre as unidades acústicas utilizadas, sua apresentação e origem. As regras linguísticas para realização da transcrição G2P encontram-se na Seção 5.3, bem como o detalhamento da metodologia aplicada na implementação do G2P. A Seção 5.4 segue com a apresentação dos testes realizados e resultados obtidos. Finalmente, na Seção 5.5, apresentamos as conclusões e trabalhos futuros relativos ao conversor G2P.

5.1 Introdução

Dentro do contexto da síntese de voz, o módulo de conversão G2P é o mais importante, pois é a partir dele que os rótulos serão gerados para a concatenação dos modelos. Além do mais, todos os algoritmos até aqui apresentados, buscam ter o melhor resultado possível, visando justamente esta fase do processamento do texto [22, 30, 44, 45, 67, 73].

A questão da transcrição fonética em sistemas TTS tem sido bastante estudada e não tem sido uma questão com assunto esgotado, já que aperfeiçoar o módulo G2P, além de outros casos, significa melhorar as características de inteligibilidade e naturalidade da fala sintetizada [41].

5.2 Unidades acústicas utilizadas

O conjunto de unidades acústicas escolhido para ser empregado pelo conversor G2P é mostrado na Tabela 5.1, além de alguns exemplos do SAMPA (*Speech Assessment Methods Phonetic Alphabet*) [38]. Apesar de alguns sistemas TTS considerarem

ditongos como unidades acústicas independentes, como é o caso do sistema de síntese de voz Festival [74] ¹, devido às suas características peculiares onde os formantes da vogal/semi-vogal inicial são suavemente substituídos pelos formantes da semi-vogal/vogal seguinte [75]. Aqui eles não são considerados desta forma, seguindo o exemplo de outros conjuntos de fones reportados para síntese de voz no português brasileiro [76, 77]. Conseqüentemente, ditongos são modelados como seqüência de difones, como [a w], por exemplo, para o ditongo formado pela seqüência de grafemas <a l> na palavra *animal*.

5.3 Conjunto de regras para transcrição fonética

Nesta Seção, apresentamos a simbologia utilizada no conjunto de regras para a transcrição fonética, como visto na Tabela 5.2, além das regras propriamente ditas, que podem ser vistas da Tabela 5.3 até a Tabela 5.5, onde cada tabela corresponde a um grupo de fonemas. A coluna da esquerda contém a seqüência de grafemas, a do meio apresenta a transcrição fonética, e a da direita apresenta os exemplos. Dentro de cada tabela, as regras estão em seqüência, tal que se uma dada regra é associada, o grafema é convertido no fone correspondente.

Em um texto a ser convertido, o algoritmo associa um determinado grafema ao seu fone correspondente, seguindo a ordem de ocorrência de cada grafema. Em outras palavras, o algoritmo de conversão é acionado para cada letra, concatenando as unidades acústicas resultantes, dentro da seqüência fonética de saída. Contudo, existem casos onde a aplicação de uma dada regra pode resultar num salto do próximo grafema. Por exemplo, considerando o segundo *a* na sentença *eles cantam*, como este caso cai na Regra 2, mostrada na Tabela 5.3, a transcrição da letra *m* é desconsiderada.

Finalmente, existem casos em que o grafema *x* pode ser convertido em uma seqüência de dois fones ao invés de um, conforme as exceções mostradas na Tabela 5.6.

Uma vez realizada a transcrição fonética, as informações da sentença são organizadas e enviadas, no formato mostrado na Figura 5.1, ao módulo de síntese de voz. Podemos observar pela figura mostrada que as informações de fones, sílabas, tonicidade, palavras e classe gramatical, são todas reunidas neste rótulo enviado para a síntese. Essas informações contextuais vão referenciar a forma como os modelos serão concatenados. Detalhes a respeito do impacto da classe gramatical podem ser verificados em [78].

¹Situação verificada para a língua inglesa.

Tabela 5.1: Conjunto de fones representados pelo SAMPA, com alguns exemplos.

Fones	Exemplos
Vogais orais	
[a]	jatobá, capacete, cabeça, lua
[E]	é, pajé, pele, ferro, velho
[e]	capacete, resolver, respeito
[i]	justiça, país, lápis, idiota, ele
[O]	ópio, jogos, sozinho, forte
[o]	jogo, golfinho, corpo
[u]	raul, culpa, baú, cururu, logo
Fricativas vozeadas	
[z]	casa, coisa, quase, exato
[v]	vovó, vamos, avião
[Z]	geladeira, trovejar
Africadas	
[tS]	tia, pacote, constituinte
[dZ]	dia, cidade, disco
Plosivas	
[b]	barba, absinto
[d]	dados, administrar
[t]	pató, constituinte
[k]	casca, quero, quanto
[g]	guerra, gato, agüentar, agnóstico
[p]	papai, psicólogo, apto
Vogais nasais	
[a~]	andar, tampar, canção, cama
[e~]	então, tempo, bem, menos
[i~]	ninho, tinta, latina, importa
[o~]	onda, campeões, somos, homem
[u~]	um, muito, umbigo
Semivogais	
[w]	fácil, voltar, eu, quase
[j]	pai, foi, caracóis, micróbio
[w~]	não, cão
[j~]	muito, bem, parabéns, compõe
Líquidas	
[l]	laranja, leitão
[L]	calhar, colheita, melhor
[R]	carro, rua, rato, carga, germe
[X]	casar, certo, arpa, arco
[r]	garoto, frango, por exemplo
Fricativas não-vozeadas	
[f]	festa, fanfarrão, afta, afluyente
[s]	sapo, caçar, crescer, sessão, lápis, capaz, casca, excesso
[S]	chá, xaveco, cachorro
Consoantes nasais	
[m]	mamãe, emancipar
[n]	nome, atenuar, encanação
[J]	casinha, galinha

Tabela 5.2: Símbolos usados nas regras de transcrição fonética.

Símbolo	Significado
...	Caracter qualquer
V	Vogal
V_ton	Vogal tônica
V_atn	Vogal átona
C	Consoante
C_v	Consoante vozeada
C_uv	Consoante não-vozeada
$\langle x_i \rangle$	Grafema x_i
Pont	Pontuação (Espaço em branco, ponto, hífen, exclamação, etc.)
$[y_i]$	Fone y_i
$[*]$	Nenhum fone transcrito
Ltr	uma letra (a, b, c, etc.)
$\dots \langle x_i \rangle \dots$	Caso <i>Default</i> para grafema x_i
$\langle \mathbf{W_bgn} \ x_i \rangle$	Grafema x_i vem no início da palavra
$\langle x_1, (x_2, x_3) \rangle$	$\langle x_1 \ x_2 \rangle$ ou $\langle x_1 \ x_3 \rangle$
x_atn	x é átona
$\langle \Theta - x_i \rangle$	O conjunto Θ excluindo caractere x_i , onde Θ pode ser: V , V_ton , V_atn , C , C_v , C_uv e Ltr
sp	Espaços em branco
Prn_M	Pronome masculino
Prn_F	Pronome feminino

5.4 Resultados obtidos

O G2P teve as suas regras de transcrição fonética testadas com um extrato de texto da base de dados CETEN-Folha [49], composto de 40.007 fones. Os resultados obtidos podem ser verificados na Tabela 5.7, onde pôde ser atingida uma taxa de acerto de 98,57%.

A grande quantidade de fones submetidos a teste no G2P, confrontado com a quantidade e o tipo de erros encontrados, pôde mostrar que o algoritmo apresenta uma performance boa. Os erros de nasalização foram os mais acentuados e já encontra-se em fase de análise, para que sejam diminuídos. Os erros decorrentes da marcação tônica de grafemas [e] e [o], ocorreram escalonadamente decorrentes de alguns erros nos algoritmos anteriores, homógrafos (Capítulo 2) e determinação de tonicidade (Capítulo 3). O tratamento destes problemas, acarretará numa redução considerável de alguns erros do G2P. Foi verificado, também, que os estrangeirismos constituem uma dificuldade de leitura e que, apesar de não terem um impacto muito grande no contexto geral do texto (0,28%), um estudo sobre a integração fonética de palavras estrangeiras no PB está sendo planejado.

phone	syll	stress	word	class
sil				
s	si~	0	si~tezi	content
i~				
t	te	1		
e				
z	zi	0		
i				
dZ	dZi	0	dZi	function
i				
v	v0s	1	v0s	content
0				
s				
k	ko~	0	ko~	content
o~				
e	e	0	emoso~j~	content
m	mo	0		
o				
s	so~j~	1		
o~				
j~				
p	pa	1	para	function
a				
r	ra	0		
a				
u	u	0	u	function
p	poX	0	poXtugejs	content
o				
X				
t	tu	0		
u				
g	gejs	1		
e				
js				
b	bra	0	brazilejru	content
r				
a				
z	zi	0		
i				
l	lej	1		
ej				
r	ru	0		
u				
sil				
sil				
sil				

Figura 5.1: Rótulos relativos a sentença “Síntese de voz com emoções para o PB”.

5.5 Conclusões e trabalhos futuros

Foi apresentado nesta tese, um algoritmo capaz de realizar a transcrição fonética de palavras do PB, realizando a conversão G2P aplicada em um TTS. O algoritmo proposto foi implementado e testado com um extrato de texto da base de dados CETEN-Folha, composto por 40.007 fones, tendo sido atingida uma taxa de acerto global de 98,57%. Considerando o resultado dos testes realizados, pôde ser constatado que a aplicação de regras linguísticas é uma boa escolha para a conversão G2P do PB. Como trabalhos futuros a serem realizados, temos uma pesquisa, seguida de correções nos erros de tonicidade em [e] e [o], já mencionados no Capítulo 3, bem como uma manipulação adequada de estrangeirismos, ainda não implementado.

Tabela 5.3: Regras para os grafemas <a, b, c, d, e, f, g, h, i>.

Regra	Algoritmo para o Grafema <a>	Fone	Exemplos
1	...<an><Pont>...	[a [~]]	Ivan, Itapoan
2	...<am><Pont>...	[a [~] w [~]]	andam, cresçam
3	...<a(V ton)><m,n>...	[a [~]]	cama, banho
4	...<â (m,n)><C-h>...	[a [~]]	lâmpada, cântico
5	...<ão>...	[a [~] w [~]]	avião
6	...<ã,â>...	[a [~]]	amanhã, câmara, tâmara
7	...<á,à>...	[a]	Antártica, àquela
8	...<a (m,n)><C-h>...	[a [~]]	antropofagia, amperímetro
9	...<a>...	[a]	aracnofobia
Regra	Algoritmo para o Grafema 	Fone	Exemplos
1	[b]	abacate
Regra	Algoritmo para o Grafema <c>	Fone	Exemplos
1	...<c><e,i>...	[s]	aceitar, jacinto
2	...<ç>...	[s]	almoço
3	...<c h>...	[S]	acho
4	...<c>...	[k]	claro
Regra	Algoritmo para o Grafema <d>	Fone	Exemplos
1	...<d><i,[i]>...	[dZ]	dia, tarde
2	...<d><C-r,l>...	[dZ]	advogado
3	...<d><Pont>...	[dZ]	Raid
4	...<d>...	[d]	dote
Regra	Algoritmo para o Grafema <e>	Fone	Exemplos
1	...<en><Pont>...	[e [~]]	hífen, líquen, Carmen, abdômen
2	...<e(V ton)><l><C-h,Pont>...	[E]	papel, relva, mel, selva
3	...<e(V aton)><l><C-h,Pont>...	[e]	sensível, selvagem
4	...<ô,â><e>...	[j [~]]	mães, corações, alemães
5	...<a><e>...	[j]	Caetano
6	...<(Prn M) e (V ton)>...	[e]	ele, este
7	...<(Prn F) e (V ton)>...	[E]	ela, esta
8	...<(W bgn) e><x><V>...	[e]	exato, existe
9	...<(W bgn) e,ê><x,s><C>...	[e]	exceto, excelente, êxtase, estrada, escola
10	...<(W bgn) e><C-(m,n,x)>...	[e]	errado, econômico, equestre, elefante
11	...<e><ne,me><Pont>...	[e]	creme, higiene, solene, gene, leme
12	...<e><la,lo><Pont>...	[E]	vela, bela, beringela, aquarela
13	...<e><sa,ssa,za><Pont>...	[e]	Exceções desta regra para as palavras pelo, pela, chinesa, condessa, turquesa, baronesa, pobreza, riqueza
14	...<ê (m,n)><C-h>...	[e [~]]	ciência, existência
15	...<e (m,n)><C-h>...	[e [~]]	embora, entoação
16	...<e (V ton)><m,n>...	[e [~]]	tema, comemos, cena, pena
17	...<(e, ê, è)m><Pont>...	[e [~] j [~]]	contém, têm, ontem
18	...<ê>...	[e]	bebê
19	...<é>...	[E]	época
20	...<e><s><Pont>...	[i]	frases
21	...<e(V ton)i>...	[e j]	fiquei, areia, feio
22	...<e><Ltr><V ton>...	[e]	semestre, perigo
23	...<Pont><e><Pont>...	[i]	o João e a Ana
24	...<e><Pont>...	[i]	índice
25	...<e>...	[e]	bebê
Regra	Algoritmo para o Grafema <f>	Fone	Exemplos
1	...<f>...	[f]	faca
Regra	Algoritmo para o Grafema <g>	Fone	Exemplos
1	...<g><e,i>...	[Z]	gelo
2	...<g u><e,i>...	[g]	guindaste
3	...<g>...	[g]	garoto
Regra	Algoritmo para o Grafema <h>	Fone	Exemplos
1	...<h>...	[*]	hoje
Regra	Algoritmo para o Grafema <i>	Fone	Exemplos
1	...<u (V ton)><i><t>...	[j [~]]	muito
2	...<i e><Pont>...	[i]	superfície, planície, calvície, cárie
3	...<i,i (m,n)><C-h,Pont>...	[i [~]]	acimba, sinto, símbolo, fim
4	...<i><m,n><V,h>...	[i [~]]	inoperante, imaginar, ninho
5	...<V-i><i>...	[j]	coisa, sai, queixa, paisagem
6	...<i,i>...	[i]	líquido, sai

Tabela 5.4: Regras para os grafemas <j, k, l, m, n, o, p, q, r>.

Regra	Algoritmo para o Grafema <j>	Fone	Exemplos
1	...<j>...	[ʒ]	jambo
Regra	Algoritmo para o Grafema <k>	Fone	Exemplos
1	...<k>...	[k]	kelvin
Regra	Algoritmo para o Grafema <l>	Fone	Exemplos
1	...<l><V>...	[l]	ala
2	...<l h>...	[L]	alho
3	...<l>...	[w]	vogal
Regra	Algoritmo para o Grafema <m>	Fone	Exemplos
1	...<e,é,ê,i><m><sp><V>	[j]	Alguém usou, Quem está?
2	...<m>...		Eles têm amor, Tim está pronto, mameluco
Regra	Algoritmo para o Grafema <n>	Fone	Exemplos
1	...<n h>...	[ɲ]	ganho
2	...<n>...	[n]	nata
Regra	Algoritmo para o Grafema <o>	Fone	Exemplos
1	...<o(V ton)><l><C-h,Pont>...	[O]	sol, girassol, futebol
2	...<o(V aton)><l><C-h,Pont>...	[o]	soldadura, soltar
3	...<ou>...	[ow]	ouvir, couve, estou
4	...<o(V ton)><a><Pont,Ltr>...	[o]	Lisboa, pessoa, perda, canoa
5	...<o><so><Pont,Ltr>...	[o]	saudoso, virtuoso, caprichoso
6	...<o><sa><Pont,Ltr>...	[O]	saudosa, virtuosa, caprichosa
7	...<o,ô (m,n)><C-h,Pont>...	[õ]	compositor, gôndola, som, Kibon
8	...<o(V ton)><m,n>...	[õ]	soma, sono, sonho, ponho, risonho, medonho
9	...<o><r><Pont>...	[o]	compor, dor, exterior, amor, sabor Exceção: maior(es), menor(es), melhor(es), pior(es), suor(es) ([O])
10	...<o><z><Pont>...	[O]	voz, algeoz, atroz Exceção: arroz ([o])
11	...<ô>...	[o]	vovô
12	...<ó>...	[O]	vovó
13	...<ô>...	[õ]	corações
14	...<(W bgn)c><o><Pont>...	[o]	co-produção
15	...<a><o><C,Pont>...	[w]	ao, caos
16	...<o><Ltr><V ton>...	[o]	opor
17	...<o><Pont>...	[u]	músico
18	...<o><s><Pont>...	[u]	carros
19	...<o>...	[o]	escopo
Regra	Algoritmo para o Grafema <p>	Fone	Exemplos
1	...<ph>...	[f]	Philippe
2	...<p>...	[p]	pato
Regra	Algoritmo para o Grafema <q>	Fone	Exemplos
1	...<qu><i,e,o>...	[k]	quito, quente, quota
2	...<q><ü,ua>...	[k]	cinquenta, quase
Regra	Algoritmo para o Grafema <r>	Fone	Exemplos
1	...<rr>...	[R]	carro
2	...<r sp r>...	[R]	Um pomar rodeado de flores.
3	...<(W bgn)r>...	[R]	A rua foi interditada.
4	...<r><V>...	[r]	ratoeira
5	...<r><sp><V,h>...	[r]	Falta acertar apenas uma.
6	...<r><sp><C uv>...	[X]	Pecar pelo meio do caminho.
7	...<r><sp><C v>...	[R]	Injetar grãos de arroz.
8	...<r><C uv>...	[X]	perco
9	...<r><C v>...	[R]	carga
10	...<r>...	[X]	Ela irá se lascar.

Tabela 5.5: Regras para os grafemas <s, t, u, v, w, x, y, z>.

Regra	Algoritmo para o Grafema <s>	Fone	Exemplos
1	...<tr(a,ã)n><s><V>...	[z]	transa <u>ç</u> ão, trãns <u>ç</u> ito
2	...<ob><s><équio>...	[z]	obs <u>ç</u> équio
3	...<ã><s><Pont>...	[j~ s]	fã <u>s</u> , ma <u>ç</u> ãs
4	...<V_ton><s><Pont>...	[j s]	ma <u>s</u> , gã <u>s</u> , atrã <u>s</u> , Goiã <u>s</u>
5	...<sh>...	[S]	sh <u>ç</u> iatsu
6	...<(W_bgn)s>...	[s]	O s <u>ç</u> apato está lustrado.
7	...<V><s><V>...	[z]	as <u>ç</u> a
8	...<s><C_v>...	[z]	transgred <u>ç</u> ir
9	...<ss>...	[s]	ass <u>ç</u> ar
10	...<sc><e,i>...	[s]	cre <u>ç</u> cer
11	...<sc>...	[s]	cre <u>ç</u> cam
12	...<s sp j>...	[Z]	E <u>ç</u> les jogaram bola.
13	...<s><sp><V,C_v,h>...	[z]	O <u>ç</u> aros são cromados.
14	...<s>...	[s]	E <u>ç</u> les reberam o prêmio.

Regra	Algoritmo para o Grafema <t>	Fone	Exemplos
1	...<th><Pont>...	[tS]	Ru <u>ç</u> h
2	...<th>...	[t]	Ar <u>ç</u> thur
3	...<t><C-r,l>...	[tS]	algor <u>ç</u> itmo
4	...<t><i,[i]>...	[tS]	t <u>ç</u> ia, me <u>ç</u> e
5	...<t><Pont>...	[tS]	Aquele set foi dif <u>ç</u> icil.
6	...<t>...	[t]	t <u>ç</u> ato

Regra	Algoritmo para o Grafema <u>	Fone	Exemplos
1	...<ü>...	[w]	lingü <u>ç</u> ística
2	...<u(m,n)>...	[u~]	ab <u>ç</u> undante, ret <u>ç</u> umbante
3	...<u(m,n)><Pont>...	[u~]	Ele come at <u>ç</u> um.
4	...<u><m,n>...	[u~]	u <u>ç</u> ma, u <u>ç</u> nha
5	...<ú,u>...	[u]	ac <u>ç</u> ústica

Regra	Algoritmo para o Grafema <v>	Fone	Exemplos
1	...<v>...	[v]	vo <u>ç</u> ando

Regra	Algoritmo para o Grafema <w>	Fone	Exemplos
1	...<w>...	[w]	w <u>ç</u> att

Regra	Algoritmo para o Grafema <x>	Fone	Exemplos
Verificar a Tabela 5.6			

Regra	Algoritmo para o Grafema <y>	Fone	Exemplos
1	...<y><C>...	[i]	Y <u>ç</u> guaçu
2	...<y>...	[j]	Y <u>ç</u> anomami

Regra	Algoritmo para o Grafema <z>	Fone	Exemplos
1	...<z><sp><C_uv>...	[s]	Ferr <u>ç</u> az furou o ferro.
2	...<z><sp><C_v,V,h>...	[z]	Faz anos que não o vemos; Ferraz gosta de pão; Faz horas que o vejo.
3	...<V_ton><z><Pont>...	[j s]	f <u>ç</u> az
4	...<z><Pont>...	[s]	O que José fez?
5	...<z>...	[z]	z <u>ç</u> umbido

Tabela 5.6: Exceções para o grafema <x>.

Mapeamento G2P	Lista parcial de palavras
<x>	[k s] oxít <u>ç</u> ono, oxít <u>ç</u> ona, oxít <u>ç</u> onos, oxít <u>ç</u> onas, oxidar, oxida <u>ç</u> ão, complex <u>ç</u> o, anexar, oxigê <u>ç</u> nio, oxiú <u>ç</u> ro, oxalato, ú <u>ç</u> xer, uxoricida, axila, axiologia, í <u>ç</u> xia, tá <u>ç</u> xi, sintax <u>ç</u> e
<x>	[k z] ixofagia, ixomielite, ixolite, ixô <u>ç</u> metro, ixora, ixoscopia, ox-acé <u>ç</u> tico

Tabela 5.7: Tabela com resultados do mapeamento dos grafemas.

Tipo de erro	Ocorrências	% de erro
[e] ou [E]	123	99,69% (parcial)
[o] ou [O]	60	99,85% (parcial)
Fones incorretos em palavras estrangeiras	114	99,72% (parcial)
Ditongos	11	99,97% (parcial)
Fones incorretos em acrônimos	10	99,98% (parcial)
Erros de nasalização	256	99,36% (parcial)
TOTAL	574	98,57% (global)

Capítulo 6

Síntese de voz e as emoções da fala

Neste capítulo, apresentamos as idéias que relacionam a síntese de voz com as emoções da fala, além das carências e aspirações para a síntese de voz no PB. Na Seção 6.1, temos um introdução do assunto, apontado alguns estudos já realizados. Na Seção 6.2, discutimos a falta de uma base de dados com emoções para o nosso idioma e como foi idealizado o projeto da nossa base. A Seção 6.3 contem os detalhes da base de dados com emoções variadas e suas particularidades. Na sequência, a Seção 6.4 possui a metodologia aplicada na validação dos dados gravados, bem como seus resultados quanto a percepção emotiva dos ouvintes. Finalmente, na Seção 6.5, apresentamos as conclusões e trabalhos futuros concernentes à nova base de dados para síntese e reconhecimento de voz.

6.1 Introdução

As comunicações através da voz são uma das formas mais tradicionais da comunicação humana. Por este motivo, diversos pesquisadores não têm medido esforços para incorporar a voz na interação homem-máquina, seja em sistemas de reconhecimento e síntese de voz, bem como em sistemas de verificação e identificação de locutor, reconhecimento e síntese de voz com emoções, etc.

Tradicionalmente, sistemas TTS buscam converter texto em forma de onda, e tem sido considerada uma das mais importantes áreas de pesquisa entre as tecnologias de voz, principalmente por proporcionar mais acessibilidade a deficientes físicos nas interações com máquinas inteligentes.

As informações linguísticas são tudo que podem ser carregadas pelo texto, mas elas representam apenas uma pequena parte da mensagem falada. Existe muito mais a ser extraído da fala para que a mensagem possa ser completamente mapeada. Como característica de seres humanos, ao ouvirmos a fala de um locutor, somos sensíveis a um conjunto de informações extra-linguísticas, relacionadas à identidade e ao estado do locutor, além de informações paralinguísticas, relacionadas às intenções

do locutor, pré-existentes nas “entrelinhas” da mensagem transmitida. Informações como estas são extremamente necessárias ao adequado processo de síntese da fala, e até mesmo em sistemas de reconhecimento [79], trazendo notória naturalidade e inteligibilidade para a fala sintética.

Um amplo número de sistemas TTS tem sido desenvolvidos e diversas técnicas propostas em diversas aplicações práticas. Os sistemas TTS baseados em concatenação de unidades acústicas são exemplos destas técnicas, com muito sucesso na produção da fala artificial. Entretanto, apesar destes sintetizadores produzirem fala com alta qualidade, ainda existem algumas dificuldades na produção de voz artificial com características variadas, tais como estilos, emoções, etc. Um destes problemas é que as técnicas baseadas na seleção e concatenação de unidades acústicas, normalmente requerem uma grande quantidade de dados de voz armazenados para produzirem a fala sintética com características diversas. O fator complicador reside no fato de que o processo torna-se muito mais oneroso ao se coletar e armazenar uma grande quantidade de dados de voz com características distintas.

Em muitos níveis da conversação humana, as intenções do locutor, estão sinalizadas pelo modo como ele fala. Como os seres humanos estão acostumados a usar estas informações extra-verbais, certamente esperarão isto quando interagirem com máquinas por meio da voz [80].

Recentemente, muitos trabalhos tem sido reportados a respeito do estudo das emoções através da fala [79, 81, 82, 83]. Para os especialistas em produção da fala com emoções, é sabido que informações de entonação e prosódia são extremamente importantes [84], sem as quais seria praticamente impossível sintetizar com emoções variadas. Para um mapeamento adequado dessas informações, uma boa base de dados torna-se necessária, tanto para o estudo das características dos estados emocionais variados, bem como para o desenvolvimento de algoritmos capazes de inserir estas emoções na reprodução da fala.

Como a fala possui características não-verbais em adição às informações fonéticas, é desejável que sistemas TTS sejam capazes de sintetizar a voz com características variadas. Para evitar o problema da síntese concatenativa de formas de onda em sistemas de síntese de voz com estilos emocionais variados, esta tese é dedicada ao desenvolvimento de um sistema TTS, inédito para o PB, com estilos emocionais variados, utilizando HMMs, que são muito utilizados para modelar a voz em sistemas de reconhecimento automático da fala (ASR - *Automatic Speech Recognition*).

A utilização de HMM na síntese de voz torna possível modificar as características da voz para um locutor alvo através de mudanças dos parâmetros espectrais da fala. Com apenas algumas sentenças gravadas, um sistema baseado em HMM pode gerar a fala sintetizada muito semelhante a de qualquer locutor.

6.2 Base de dados de voz para Síntese com emoções

Apesar das inúmeras pesquisas realizadas nas últimas décadas a respeito de base de dados para síntese e reconhecimento de voz [85, 86, 87], uma pequena parte delas tem incluído as emoções da fala, ou dentro do que tem sido desenvolvido, apenas algumas possuem documentação disponível.

Considerando o português falado no Brasil, existe uma escassez ainda maior de materiais disponíveis relativos a pesquisas realizadas no escopo da síntese ou reconhecimento de voz com emoções [82]. Essa falta de dados para pesquisa tornou-se uma das motivações para a gravação de uma base de dados com estados emocionais variados, voltados para síntese e reconhecimento de voz com emoções.

Uma questão ainda é muito discutida no âmbito da comunidade científica: usar fala real ou atuada (lida)?

Estudos realizados têm mostrado que as bases de dados utilizadas tem sido gravadas, em estúdios, através de atores simulando várias emoções. Entretanto é importante considerar que, enquanto as emoções pretendidas puderem ser identificadas posteriormente pelo ouvinte, elas podem ser consideradas satisfatórias, porém, sendo atuada, é questionável se elas representam fielmente as características da fala pronunciada normalmente por uma pessoa, quando elas naturalmente sentem emoções semelhantes [88].

Existem muitos fatores que os especialistas alegam desfavorecer o uso da fala atuada, mas um dos principais é de que existem muitas emoções que não podem ser conscientemente “imitadas”. Por outro lado, coletar gravações do mundo real tem diversas implicações legais, éticas, além de questões relativas à privacidade, momento ímbar para a perfeita coleta de algumas condições e estados emocionais.

Utilizar a fala natural, torna a investigação impossível se as emoções básicas são o objeto de estudo. A base de dados torna-se imprescindível para o desenvolvimento de qualquer estudo sobre estados emocionais na fala. Considerando as implicações legais, supomos que a base de dados pode ser gravada usando-se a fala atuada, desde que algumas providências sejam tomadas:

- Utilizar um número amplo de locutores, com todas as emoções, a fim de proporcionar generalização do conjunto;
- Todos os locutores devem gravar o mesmo conteúdo verbal para permitir uma comparação cruzada entre emoções e locutores;
- As gravações deverão ser feitas em estúdio, com alta qualidade de áudio, minimizando o ruído de fundo, sem prejuízo das medidas espectrais;

Neste sentido, alguns passos tornaram-se necessários para que o objetivo fosse atingido:

- Definir o local para a gravação das vozes com emoções;
- Escolher as emoções a serem gravadas;
- Escolher os locutores que irão simular as emoções;
- Definir o conteúdo a ser gravado;
- Gravar os dados;
- Validar os dados gravados e reclassificar ou descartar as falas reprovadas;
- Realizar uma etiquetagem dos dados gravados, distintamente, por locutor, sexo, locução e emoção.

Diante da carência de uma base de dados de voz com emoções, tendo em vista não existir nenhuma disponível para o PB, desenvolvemos um projeto para a gravação de uma base de dados para síntese e reconhecimento voz com emoções para o PB. Este projeto foi financiado pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), sob o título “Projeto universal do CNPq para gravacao de uma base para síntese e outra para reconhecimento de voz”, base esta que será descrita na Seção 6.3.

6.3 O projeto da base de dados de voz com emoções para o PB

A base de dados de voz foi gravada em duas partes. A primeira parte, que pode ser aplicada em estudos de síntese e reconhecimento com ou sem emoções, foram realizadas gravações por 20 locutores/dubladores, 10 do sexo masculino e 10 do sexo feminino, com notória experiência na mídia brasileira, tais como telejornais, novelas, dublagem de filmes, comerciais, etc. Esta escolha foi feita buscando-se a máxima aproximação entre a fala atuada e a fala real. Cada um destes locutores realizou a leitura de um conteúdo comum, relacionado no Apêndice C, em cerca de 10 horas de gravação, totalizando 200 horas, a saber:

1. 1000 frases foneticamente balanceadas 20 a 20 (texto com natureza jornalística) [89];
2. 161 frases selecionadas das 1000 frases (Este conjunto de frases possui natureza jornalística, portanto não possui emoção intrínseca, sendo interpretadas em 3 emoções definidas, a saber, alegria, tristeza e raiva, totalizando 483 locuções mais as 161 neutras para comparação);

3. 85 frases interrogativas;
4. 29 frases exclamativas;
5. 23 frases com onomatopéias diversas (Frases com textos não linguísticos que foram repetidos 3 vezes por locutor, para aumentar o número de amostras);
6. Crônica - “Aprenda a dizer te amo” (texto com diversos parágrafos e narrativa contínua em terceira pessoa, com tom emotivo de tristeza);
7. Crônica - “Sobre a gloriosa noite em que eu apitei o Fla-Flu” (texto com diversos parágrafos e narrativa contínua em primeira pessoa, com tom emotivo de alegria, ironia, raiva, etc.);

A segunda parte da gravação foi realizada para aplicação voltada para o estudo de síntese de voz com ou sem emoção. Do grupo de 20 locutores participantes do projeto, foram escolhidos dois locutores para gravar conteúdo textual de diversas obras da literatura brasileira. Nos textos elencados existem diversas situações corriqueiras do nosso dia-a-dia, com estilos e emoções variadas, bem como a fala neutra contínua, totalizando cerca de 50 horas de áudio cada um.

Todo o conteúdo da base de dados está discriminado detalhadamente em [90].

Antes que fossem iniciadas as gravações, algumas considerações foram feitas para o pleno êxito do projeto:

1. As gravações foram feitas em estúdio, com alta qualidade de áudio, minimizando o ruído de fundo, sem prejuízo das medidas espectrais;
2. Foi utilizado um número amplo de locutores a fim de proporcionar generalização do conjunto;
3. Todos os locutores gravaram o mesmo conteúdo verbal para permitir uma comparação cruzada entre emoções, quando for o caso, e entre locutores;
4. Nas gravações com emoções foi evitada a utilização de atores de teatro, exclusivamente, por exagerarem demasiadamente em algumas características, com o propósito de tornar o conteúdo emocional mais claro, resultando em uma locução não-espontânea quanto a emoção inserida;
5. Entre a sala do operador e a sala de gravação, uma janela foi colocada de forma que locutores e operador pudessem ver um ao outro o tempo todo. Dois operadores estiveram presentes o tempo todo durante a gravação para reduzir os efeitos de fadiga. Um operador ficou ouvindo o locutor buscando possíveis erros de locução e o outro ficou gravando e controlando o equipamento;

6. Cada locução emotiva foi interpretada pelos locutores nas emoções sob investigação, a saber: alegria, tristeza, raiva e a fala neutra;
7. Os locutores foram colocados de pé, de frente para o microfone, ou sentados, com seus braços livre e descansados, em geral, mantendo a mesma distância do microfone durante todo o período de gravação. As pausas eram definidas pelo próprio locutor, quando este sentia-se cansado, ou pelo operador da sala, quando este percebia alguma alteração orgânica que pudesse prejudicar o processo de gravação.

Com o início do processo de gravação, uma breve orientação foi dada a cada locutor, sobre o conteúdo a ser lido ou interpretado, a forma como a leitura das locuções deviam ser realizadas, a forma de comunicação entre o locutor e os operadores e qual a atitude tomada quando houvesse algum tipo de erro percebido pelo locutor ou pelo operador.

O conteúdo a ser lido foi entregue ao locutor em folha solta, de forma que facilitasse bastante a sua leitura, principalmente quanto ao conforto e comodidade.

Cada locutor recebeu alguns minutos para reexaminar o texto e fazer os esclarecimentos necessários para sua plena compreensão do conteúdo.

Os locutores poderiam realizar a leitura em pé, procurando manter sempre a mesma distância do microfone, ou sentados, com os braços soltos, respiração controlada e a folha de papel, com o conteúdo a ser lido, sobre o púlpito, diante dele. Nas gravações do texto interpretado, procurou-se evitar excessos de gesticulação, para que não fosse perdida a noção de distância do microfone.

Os locutores poderiam realizar interrupções, quando necessário, com o propósito de não permitir a criação de uma atmosfera de estresse e fadiga, porém essas interrupções devem ser feitas entre uma frase e outra, no caso de leitura de frases, e entre um parágrafo e outro, no caso de leitura de texto contínuo. Se a interrupção for muito prolongada, o texto deveria ser lido novamente.

No caso de interpretações de emoções, os locutores foram convidados a usar suas próprias expressões emocionais do dia a dia, sem exagerar nas suas emoções. Foi recomendado que o locutor gastasse algum tempo para imaginar uma situação específica onde já havia expressado aquela emoção e a considerasse na interpretação de cada locução, mesmo que o conteúdo lido não estivesse relacionado com a emoção interpretada. Isto foi muito importante para a perfeita coleta da fala emotiva com naturalidade.

Todo conteúdo da base foi gravado em plataforma Macintosh, com o editor de áudio Pro Tools 9 [91]. As waves foram geradas com 48Khz, Estéreo (dois canais), 24 bits.

Todos os teste realizados nesta tese foram realizados com parte desta base, composta por 10 locutores, cinco do sexo masculino e cinco do sexo feminino. O conteúdo textual, composto das 1000 frases, além das locuções emotivas, em suas três versões, alegria, tristeza e raiva, totalizando 14830 locuções.

6.4 A validação cruzada das emoções

Apesar de todo cuidado no processo de gravação, algumas discrepâncias ainda ocorreram, requerendo um edição mais precisa, tais como: respirações pré e intra-locuções, clicks labiais, confusão emotiva, etc. Sendo assim, foram programadas três tipos de edições:

1. A primária, onde as waves foram segmentadas simplesmente, sem o cuidado de remover-se as respirações e clicks labiais; desde que estes não tivessem uma concentração de energia no nível dos fonemas iniciais, além da realização da validação da base através da minimização da confusão emotiva;
2. A secundária, onde toda e qualquer anomalia seria removida (respirações, clicks, pausas prolongas, etc);
3. E por fim, a terciária, onde seriam inseridos efeitos de áudio, tais como reverberação.

Nos limitamos a trabalhar nesta tese com a edição primária, não preterindo a validação cruzada, tendo em vista que, após várias horas de gravação, é comum termos um desvio da percepção emotiva.

Para isso, um teste de validação online foi realizado, visando confirmar se as emoções gravadas eram realmente percebidas pelos ouvintes como as propostas no ato da gravação. O espelho do site de validação pode ser visto na Figura 6.1. O teste de validação subjetiva foi composto de 100 locuções aleatoriamente organizadas no site, onde cada validador ouvia as locuções e classificava a emoção segundo sua percepção (avaliação subjetiva). Foram computados 20 validadores e a matriz de confusão pode ser vista na Tabela 6.1. O objetivo é mapear as waves com maior índice de confusão (menor percentual de acertos cruzados), confirmar a sua classificação e dar um destino diferente na base para o treinamento dos HMMs, baseado na análise perceptiva. Por exemplo: se a locução “ptbr_mse_ale_000.wav” (alegria), fosse classificada pela maioria dos validadores como “neu” (neutra), para efeito de treinamento, passava a ser considerada como pertencente a outro conjunto de emoções, no caso “neutra”.

A matriz de confusão pode ser vista na Tabela 6.1. Apesar do número reduzido de locuções avaliadas, podemos observar que existe uma precisão de cerca de 81,8%

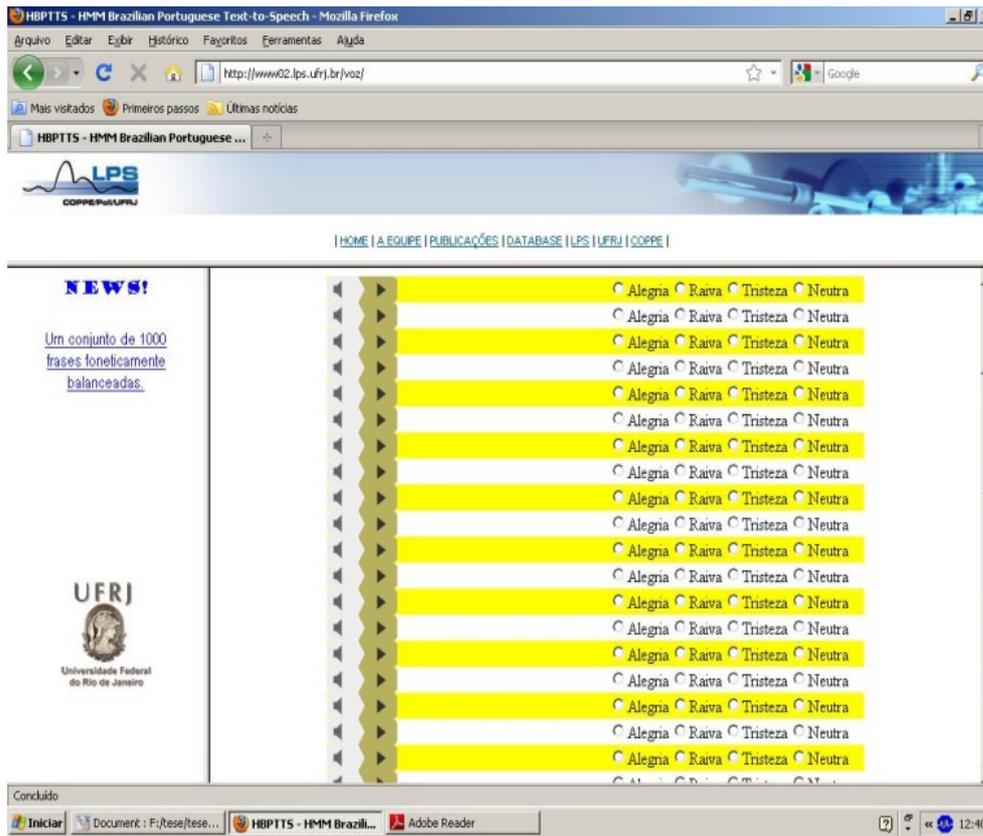


Figura 6.1: Site utilizado para validação “online”.

na percepção das emoções (taxa de percepção emotiva), como verificado pela sobreposição dos valores das Figuras 6.2, 6.3, 6.4 e 6.5 na expressão a seguir:

$$\begin{aligned} \text{Taxa de percepção emotiva} &= \frac{79,1\% + 85,0\% + 79,6\% + 83,5\%}{4} \quad (6.1) \\ &= 81,8\% \end{aligned}$$

Tabela 6.1: Matriz de confusão no teste de validação “online”.

	alegria	raiva	tristeza	neutra
alegria	387	9	16	77
raiva	31	399	7	64
tristeza	1	6	436	70
neutra	22	16	44	415

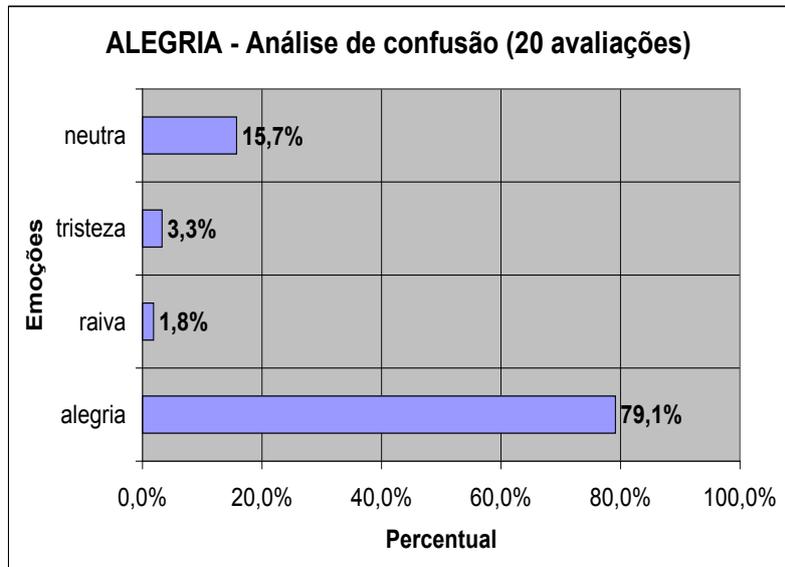


Figura 6.2: Percentual de acerto cruzado na matriz de confusão para ALEGRIA.

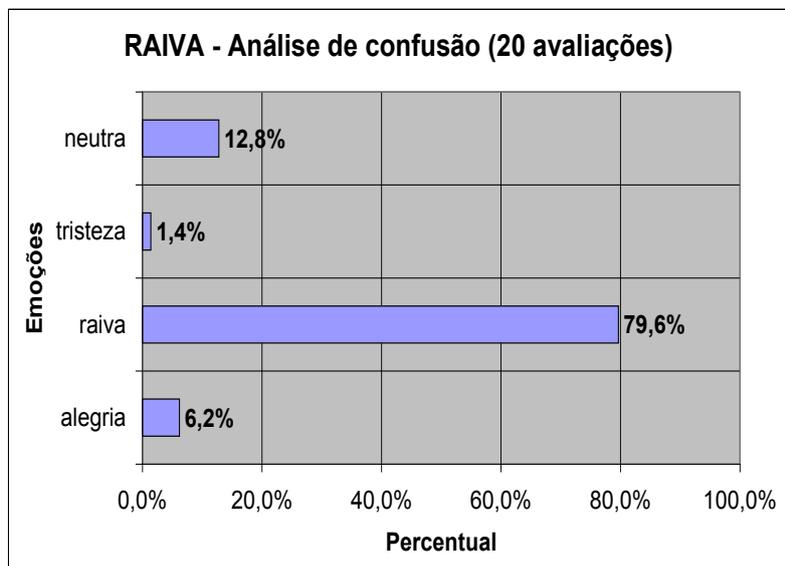


Figura 6.3: Percentual de acerto cruzado na matriz de confusão para RAIVA.

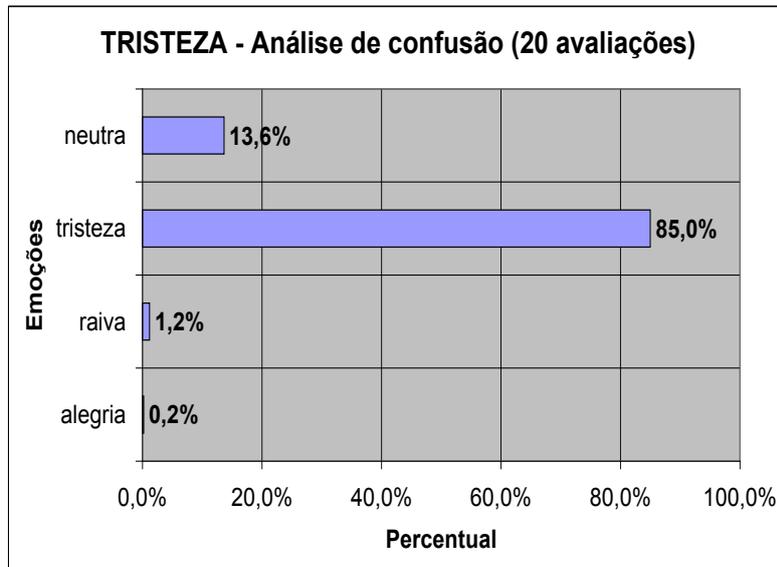


Figura 6.4: Percentual de acerto cruzado na matriz de confusão para TRISTEZA.

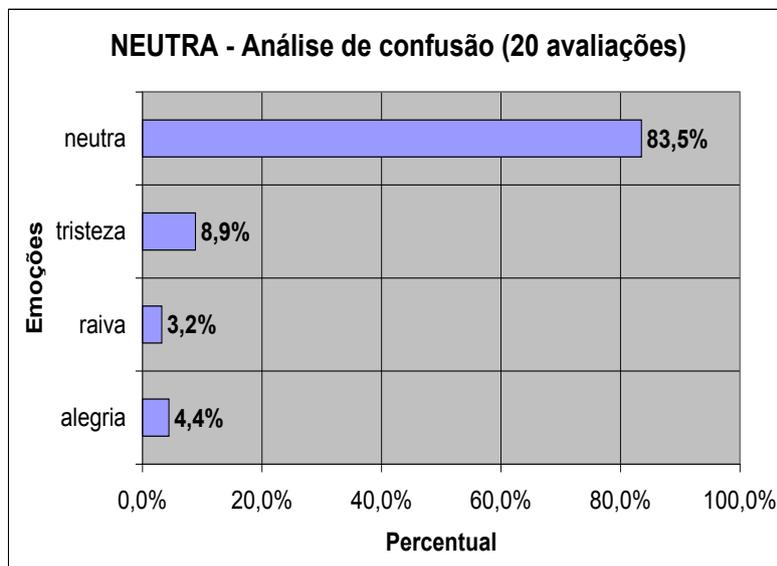


Figura 6.5: Percentual de acerto cruzado na matriz de confusão para NEUTRA

6.5 Conclusões e trabalhos futuros

Uma nova base de dados de voz para Síntese e Reconhecimento com emoções foi projetada, gravada e editada. A base, composta por 20 locutores, 10 de cada

sexo, foi utilizada parcialmente nesta tese, cinco do sexo masculino e cinco do sexo feminino. Um teste de validação foi realizado online e uma taxa de percepção emotiva de 81,8% foi verificada. As maiores confusões ocorreram com a fala neutra sendo confundida com as emoções. Isto pode ser justificado pelo fato do conteúdo textual ser jornalístico e, naturalmente, impregnado com alguma emoção intrínseca, verificada pelo locutor no ato da gravação. A confusão reversa, das emoções com a fala neutra, também pode ser justificada pelo conteúdo textual, pois muitas das emoções não eram contextualizadas e, depois de algumas horas gravando, o locutor poderia perder a sua referência emotiva e interpretativa.

Como trabalhos futuros, pretendemos dar maior amplitude na avaliação, utilizando mais análises perceptivas, além de acrescentar mais duas opções para os avaliadores ao ouvirem as locuções: “não identificada” e “outras”.

Capítulo 7

Síntese de voz com emoções baseada em HMM no PB

Neste capítulo apresentamos as considerações concernetes à síntese de voz com emoções propriamente dita, suas particularidades para o PB, bem como as ferramentas e algoritmos implementados. Iniciamos com uma breve introdução, descrita na Seção 7.1. Na Seção 7.2 apresentamos alguns conceitos relativos aos modelos escondidos de Markov, um resumo da fundamentação teórica e sua aplicação nesta tese. A Seção 7.3 contem o sistema de síntese de voz com emoções propriamente dito, seus estágios e implementações, com principal ênfase à adaptação das emoções. Em seguida, na Seção 7.4, são apresentados os testes realizados e os resultados obtidos na síntese com emoções. Finalmente, na Seção 7.5 temos as conclusões e os trabalhos futuros.

7.1 Introdução

Apesar de vários sistemas de síntese de voz realizarem a reprodução da fala com alta qualidade, existe ainda uma lacuna considerável quanto a reprodução com diversificação de características (estilos, emoções, etc). Obter características variadas em sistemas baseados na concatenação de unidades acústicas, faz com que seja difícil coletar e armazenar os dados, além de tornar o sistema oneroso, pois uma grande quantidade de dados disponíveis seria necessária para que todas as características pudessem ser representadas [92, 93].

Uma alternativa que vem sendo aplicada para síntese com características variadas é a utilização de HMMs [94, 95, 96, 97]. Tem sido observado, no estado da arte, que sistemas de síntese de voz com características variadas podem realizar a síntese baseada nas características de um locutor alvo, realizando uma técnica de adaptação, utilizando uma pequena quantidade de dados disponíveis.

Estes sistemas tem sido pesquisados de forma exaustiva em diversos idiomas, porém para o PB, ainda não existe nenhum sistema que possa reproduzir a fala com características variadas, que tenha documentação disponível. Sendo assim, neste capítulo, apresentamos nosso sistema de síntese de voz com emoções baseado em HMM.

7.2 Modelos escondidos de Markov

HMM (Modelo escondido de Markov - *Hidden Markov Model*) é um modelo probabilístico utilizado para aprendizagem de máquina, mas especificamente como um classificador de sequências [98]. A idéia base dos HMMs é representar uma variável aleatória como um processo markoviano, com a propriedade particular de que tal variável não pode ser medida diretamente, mas sim através de observações possíveis a partir de cada estado da variável. Um processo de Markov é uma máquina de estados finitos cujo estado corrente é alterado regularmente em espaços discretos de tempo, sendo o estado futuro determinado probabilisticamente apenas a partir do estado presente da máquina, condição esta conhecida como propriedade de Markov. HMMs podem ser utilizados para modelar qualquer sequência de observações,

$$O = \{o_1, o_2, \dots, o_T\} \quad (7.1)$$

onde o_t é o vetor de observações no tempo t . Um HMM contém vários estados. HMM ergódico ou sem restrição é aquele em que todas as transições de estados são consideradas. HMM esquerda-direita é aquele em que as transições de estados são limitadas da esquerda para a direita e auto-transições. Ambos são exemplificados nas Figuras 7.1 e 7.2, respectivamente, sendo este último frequentemente utilizado para modelar sequência de parâmetros de voz, visto que podem modelar adequadamente sinais cujas propriedades mudam de maneira sucessiva.

Cada estado i possui distribuição de probabilidade de saída $b_i(o_t)$, que pode ser discreta ou contínua dependendo da observação, distribuição de probabilidade de transição a_{ij} e distribuição de probabilidade inicial π_i . $b_i(o_t)$, a_{ij} e π_i vão representar, respectivamente, a probabilidade de ocorrência de um vetor de observação o_t de tempo t no estado i , a probabilidade discreta de transição do estado i para o estado j e a probabilidade de ocorrência inicial da primeira observação no estado i . Desta forma, um HMM λ de N estados é definido como:

$$\lambda = \{\Pi, A, B\} \quad (7.2)$$

$$\Pi = \{\pi_i\}_{i=1}^N \quad (7.3)$$

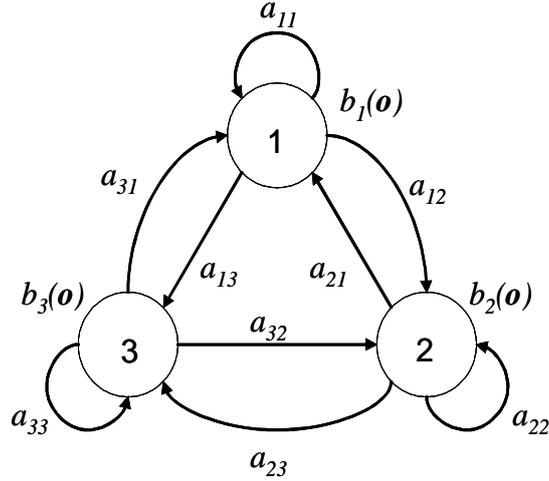


Figura 7.1: HMM ergódico ou sem restrição.

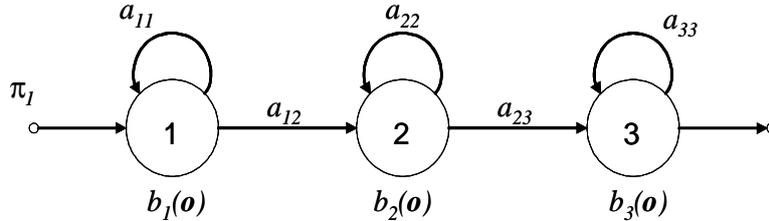


Figura 7.2: HMM esquerda-direita.

$$A = \{a_{ij}\}_{i,j=1}^N \quad (7.4)$$

$$B = \{b_i(o_t)\}_{i=1}^N \quad (7.5)$$

Para o caso de distribuição de probabilidade de saída com densidade contínua (Figura 7.3), $b_i(o_t)$, é escrita como:

$$b_i(o_t) = \mathcal{N}(o_t | \mu_i, \Sigma_i) \quad (7.6)$$

onde $\mathcal{N}(o_t | \mu_i, \Sigma_i)$ é uma distribuição Gaussiana com média μ_i e matriz covariância Σ_i ,

$$\mathcal{N}(o_t | \mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_i|}} e^{-\frac{1}{2}(o_t - \mu_i)' \Sigma_i^{-1} (o_t - \mu_i)}, \quad (7.7)$$

e n é a dimensionalidade de o_t . Para o caso de Misturas de Gaussianas (Figura 7.4), $b_i(o_t)$ é definida pelo somatório de distribuições Gaussianas ponderadas,

$$b_i(o_t) = \sum_{m=1}^M w_{im} \mathcal{N}(o_t | \mu_i, \Sigma_i), \quad (7.8)$$

onde M é o número de misturas e w_{im} , o peso da mistura. Quando o vetor de observações no tempo t é dividido em S fluxos de dados independentes, a distribuição $b_i(o_{st})$ para cada estado é definida pelo produto de Misturas de Gaussianas (Figura 7.5),

$$b_i(o_{st}) = \prod_{s=1}^S \left[\sum_{m=1}^{M_s} w_{ism} \mathcal{N}(o_{st} | \mu_{ism}, \Sigma_{ism}) \right]^{\Gamma_s}, \quad (7.9)$$

onde o_{st} é o vetor observação de fluxo s no tempo t e Γ_s é o peso de fluxo para o fluxo s .

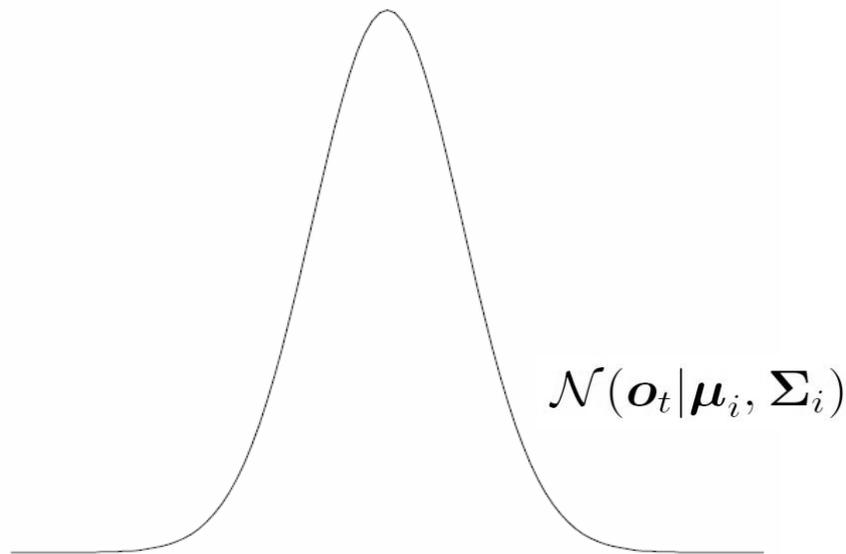


Figura 7.3: Função densidade de probabilidade gaussiana.

7.2.1 Cálculo da probabilidade do HMM

Ao definirmos uma sequência de estados como $Q = \{q_1, q_2, \dots, q_T\}$ a probabilidade de geração da sequência $O = \{o_1, o_2, \dots, o_T\}$ é calculada pelo produto das probabilidades de transição entre estados e a probabilidade de saída ao longo da sequência de estados Q :

$$P(O, Q | \lambda) = \prod_{t=1}^T a_{q_{t-1}q_t} b_{q_t}(o_t) \quad (7.10)$$

A probabilidade total de geração da sequência O no HMM λ é o somatório de

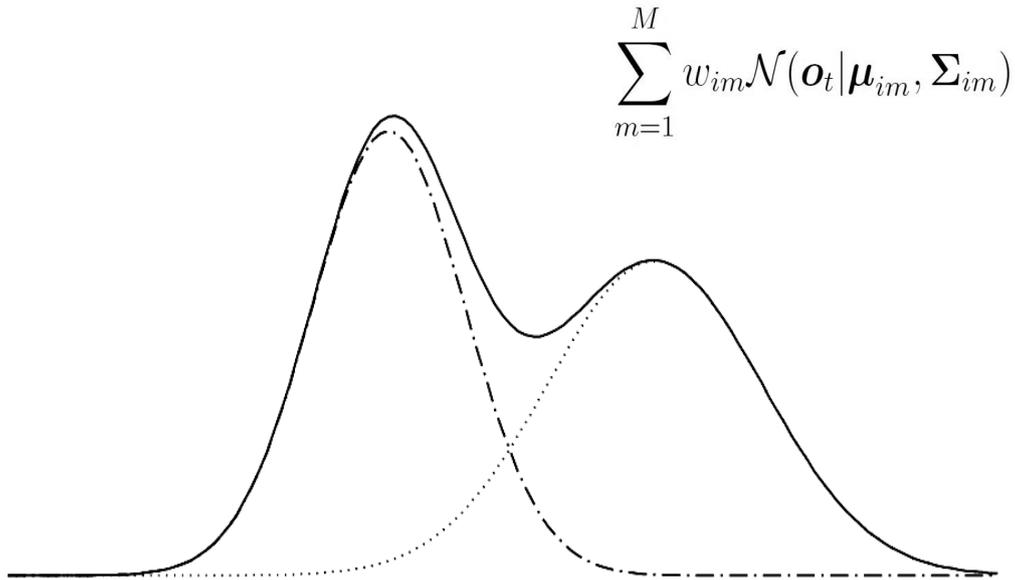


Figura 7.4: Função densidade de probabilidade Multi-Mistura.

$P(O, Q|\lambda)$ para todas sequências de estados possíveis,

$$P(O|\lambda) = \sum_{\text{all } Q} \prod_{t=1}^T a_{q_{t-1}q_t} b_{q_t}(o_t) \quad (7.11)$$

$$P(O|\lambda) = \sum_{q_1, q_2, \dots, q_T} \pi_{q_1} b_{q_2}(o_1) a_{q_1q_2} b_{q_2}(o_2) \dots a_{q_{T-1}q_T} b_{q_T}(o_T) \quad (7.12)$$

onde Q refere-se a sequência possível e a_{0j} indica π_j . A probabilidade da Equação 7.11 é calculada pelo Algoritmo Progressivo-Regressivo, cujas variáveis são $\alpha_t(i) = P(o_1, o_2, \dots, o_t | q_t = i, \lambda)$ e $\beta_t(i) = P(o_{t+1}, o_{t+2}, \dots, o_T | q_t = i, \lambda)$, e está bem descrito na literatura [75, 99, 100].

Para obter a melhor sequência de estados $Q^* = \{q_1, q_2, \dots, q_T\}$, que resulta na maior probabilidade $P(O, Q^*|\lambda)$, para uma dada observação $O = (o_1, o_2, \dots, o_T)$, utiliza-se o Algoritmo de Viterbi [100]. Ao final da execução do algoritmo, a melhor sequência de estados q_t^* é obtida para uma dada sequência de observação O e HMM λ .

O treinamento dos parâmetros para o HMM λ consiste em estimar o parâmetro λ

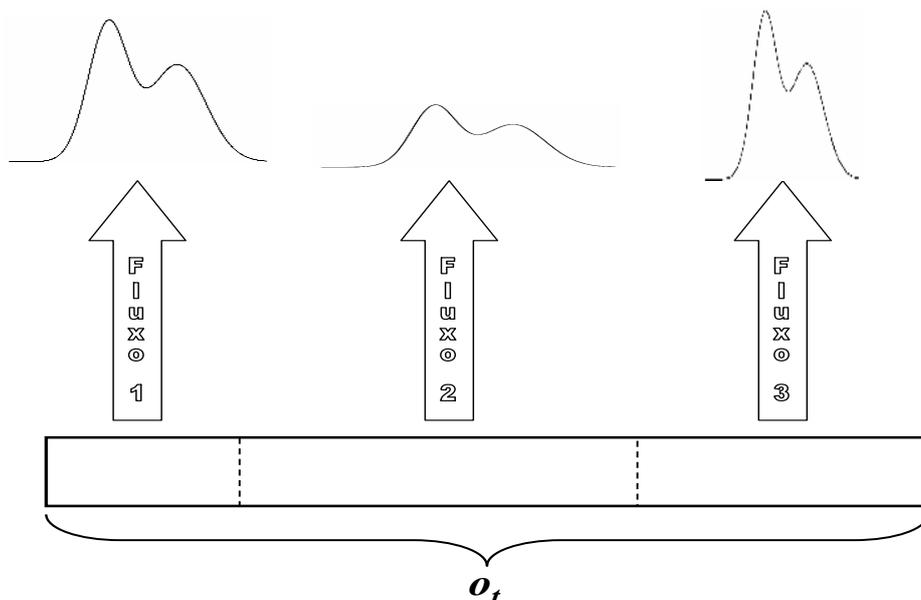


Figura 7.5: Função densidade de probabilidade Multi-Fluxo.

que maximiza a probabilidade $P(O|\lambda)$ para uma determinada sequência de observação O .

$$\lambda_{max} = \arg \max_{\lambda} P(O|\lambda) \quad (7.13)$$

Como um HMM possui parâmetros não-observáveis, ou seja, a sequência de estados é escondida, λ_{max} não pode ser obtida de forma fechada. Sendo assim, em vez de maximizar $P(O|\lambda)$, parâmetros do modelo de máxima probabilidade são estimados através do método de Baum-Welch [99, 100], que vai maximizar localmente $P(O|\lambda)$, através de um procedimento iterativo.

7.3 Síntese de voz com emoções baseada em HMM

A Figura 7.6 apresenta o sistema de síntese de voz com emoções. Ele possui três estágios: o treinamento, a adaptação e a síntese, que vão realizar a reprodução da fala com estados emotivos variados. No estágio de treinamento, parâmetros de espectro (coeficientes mel-cepestrais) e excitação (F_0) são extraídos frame a frame como características estáticas da base de dados e as características dinâmicas (parâmetros delta e delta-delta) são calculadas daqueles. As observações dos parâmetros espectrais e frequência fundamental são combinadas, frame a frame, em um vetor e HMMs dependentes do contexto são treinados utilizando-se estes vetores. O sistema não modela apenas espectro e excitação, mas também duração ao modelar a estrutura temporal da voz. Sendo assim, o sistema modela simultaneamente espectro, pitch

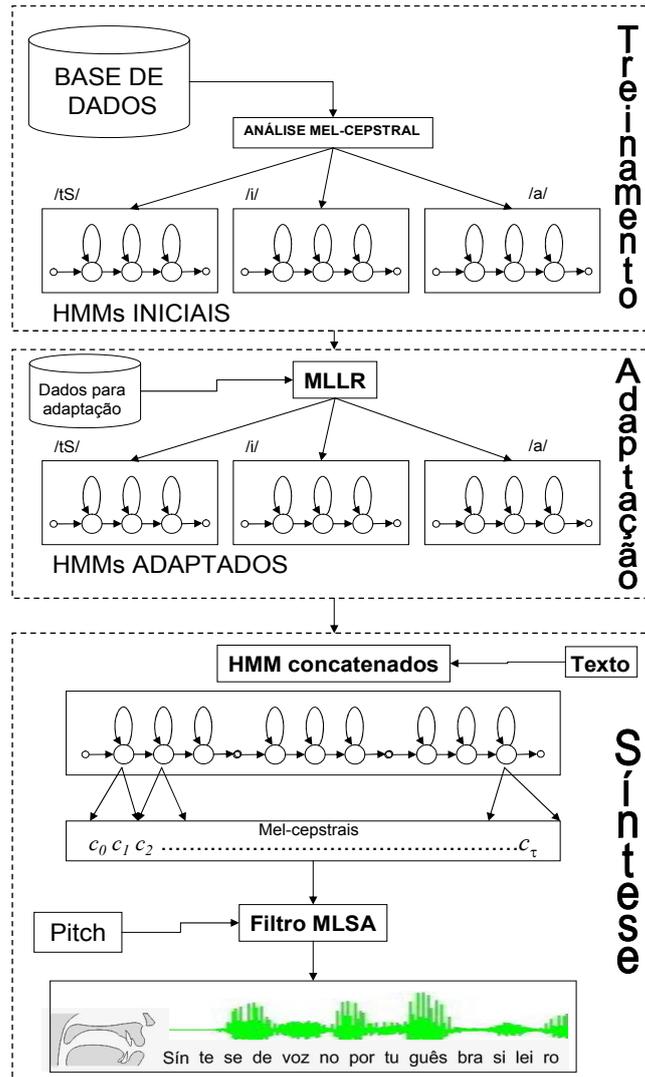


Figura 7.6: Sistema de Síntese de Voz baseado em HMM.

(frequência fundamental - F_0) e duração em uma única estrutura, considerando fatores contextuais linguísticos e fonéticos. Espectro e F_0 são modelados por HMM baseado na distribuição de probabilidade multi-espaco (MSD-HMM) [101], que inclui HMMs discretos e contínuos como casos especiais, e a duração de estados é modelada por distribuição de gaussianas multi-dimensionais. As distribuições de parâmetros para espectro, pitch e duração são agrupadas independentemente usando uma técnica de árvore de decisão baseada em agrupamento do contexto [102, 103], o que torna o sistema capaz de sintetizar arbitrariamente um dado texto ao selecionar as distribuições de saída cujas unidades HMMs não estavam nos dados de treino. A Figura 7.7 mostra a composição do vetor de observação para o treinamento dos HMMs, onde N é a ordem da análise mel-cepstral. A composição também considera a GV (Variância global - *Global Variance*) sobre os coeficientes mel-cepstrais e o

logaritmo da frequência fundamental. Este conceito foi introduzido em [104] como uma forma de aliviar a falta da variância global, utilizando os modelos existentes.

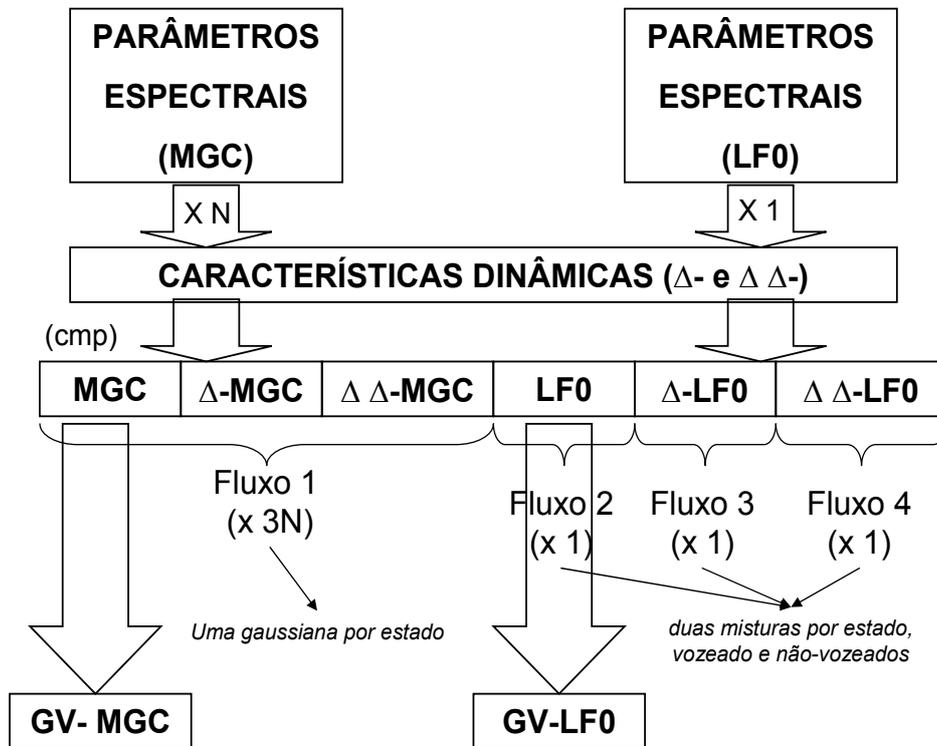


Figura 7.7: Composição do vetor de observação para treino dos HMMs.

Uma vez treinados os dados da base de treino, os modelos HMMs gerados possuem as mesmas características do locutor ou locutores daquela base, no nosso caso as emoções previamente treinadas. Um segundo estágio é apresentado, onde uma pequena quantidade de dados, diferentes da usada no treino, é utilizada para transformar as características das emoções treinadas para a emoção alvo. As características são calculadas dos dados de adaptação e os modelos iniciais são transformados nos modelos da emoção alvo através de uma técnica de adaptação, no caso CMLLR (*Constrained Maximum-Likelihood Linear Regression*), sobre os vetores de médias e as matrizes de covariância das probabilidades de saída, como demonstrado na Figura 7.8. Como a quantidade de dado de adaptação é limitada, nem todos os exemplos treinados serão vistos nos dados de adaptação. Mesmo assim, podemos transformar parâmetros de modelos não vistos, através do compartilhamento da matriz de transformação entre grupos de parâmetros. Estes grupos são chamados classes de regressão, pois a adaptação é essencialmente uma regressão linear, onde a quantidade de classes vai depender da quantidade de dados disponíveis.

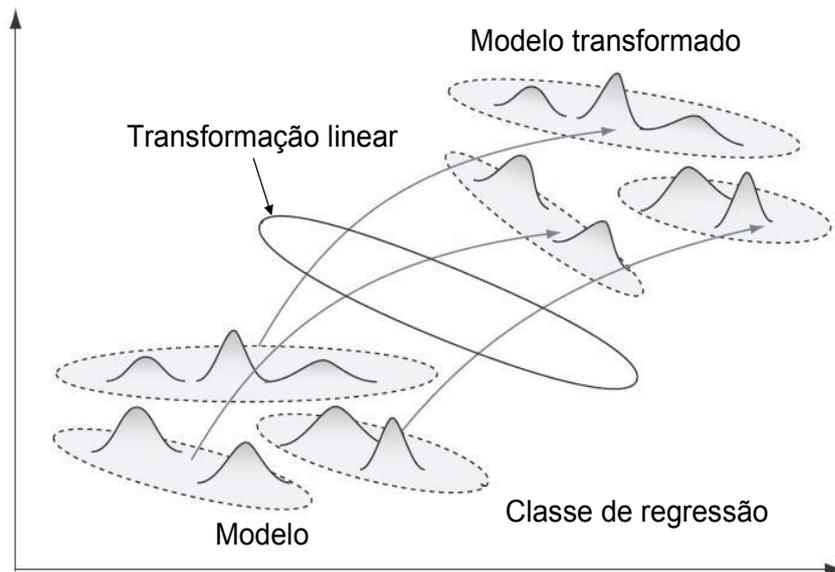


Figura 7.8: Transformação linear no espaço das emoções.

MLLR (*Maximum-Likelihood Linear Regression*) é uma técnica de adaptação cujas médias e, opcionalmente, as matrizes de covariância de um HMM são transformadas através de uma transformação linear com o objetivo de maximizar a função de probabilidade dado novos parâmetros para adaptação e o modelo, como definido a seguir:

$$\hat{\mu} = \mathbf{A}\mu + \mathbf{b} \quad (7.14)$$

$$\hat{\Sigma} = \mathbf{U}\Sigma\mathbf{U}^T \quad (7.15)$$

onde μ é um vetor de médias do modelo, Σ é sua matriz de covariância, $\hat{\mu}$ e $\hat{\Sigma}$ o vetor de médias e a matriz de covariância do modelo adaptado, respectivamente. (\mathbf{A}, \mathbf{b}) é a transformação linear para adaptar a média e \mathbf{U} , a transformação para adaptar a matriz de covariância. O Algoritmo EM (*Expectation-Maximization*) [105] é normalmente utilizado para encontrar os parâmetros ótimos, primeiro estimando a transformação \mathbf{U} , depois a transformação da média através de \mathbf{A} e \mathbf{b} . Uma variante deste método, CMLLR (*Constrained Maximum-Likelihood Linear Regression*) força

a transformação ser a mesma tanto para $\hat{\mu}$ como para $\hat{\Sigma}$, conforme a seguir:

$$\hat{\mu} = \mathbf{A}_c \mu + \mathbf{b}_c \quad (7.16)$$

$$\hat{\Sigma} = \mathbf{A}_c \Sigma \mathbf{A}_c^T \quad (7.17)$$

Aqui, a estimativa da transformação também é obtida por meio de otimização iterativa, conforme EM e a CMLLR permite a transformação ser aplicada a nível de características conforme $\hat{o}_t = \mathbf{A}_c^{-1} o_t + \mathbf{A}_c^{-1} b_c$, onde \hat{o}_t é o vetor de observação no tempo t .

Na parte de síntese, os HMMs são concatenados de acordo com o texto a ser sintetizado. O texto a ser sintetizado é transformado em uma sequência de rótulos de fonemas dependentes do contexto. Conforme a sequência de rótulos, uma sentença HMM, que representa o texto a ser sintetizado, é construída pela concatenação de HMMs dependentes do contexto. A partir das sentenças HMMs, durações dos fonemas são determinadas baseadas nas distribuições de duração de estados. Parâmetros de espectro e F_0 são gerados usando-se o algoritmo de geração de parâmetros [106], diretamente dos HMMs com características dinâmicas. Por fim, a forma de onda é gerada a partir destes parâmetros, usando-se o filtro MLSA (*Mel Log Spectrum Approximation*), que sintetiza a voz a partir da sequência gerada de coeficientes mel-cepestrais e F_0 [107].

7.4 Testes realizados e resultados obtidos

O sistema foi testado com as sentenças foneticamente balanceadas gravadas no Capítulo 6, no treinamento e na adaptação. Foram utilizados 10 locutores, cinco de cada sexo, que gravaram 1000 frases com fala neutra e 161 com fala emotiva, nas três definidas naquele capítulo (alegria, raiva e tristeza). O conteúdo gravado por cada locutor corresponde a 90 minutos em média, incluindo as regiões de silêncio. Cada locução dura entre 2,5 segundos e 8,5 segundos em média, dependendo da emoção. Na representação linguística da base foram utilizados 38 fones, conforme os apresentados na Tabela 5.1, no Capítulo 5, além de silêncio (sil), pausa (pau), os ditongos /ow/, /ej/ e /e~j~/, e as representações dos conjuntos fonéticos /js/ e /j~s/.

A base, inicialmente gravada a 48kHz, estéreo, 24 bits, foi subamostrada para 16kHz, mono, 16 bits. A janela utilizada foi a de Hanning com 25ms e overlap de 5ms. Foram extraídos coeficientes mel-cepestrais e valores de pitch, além de suas características dinâmicas, delta e delta-delta de cada frame. Os coeficientes mel-cepestrais foram obtidos através de uma análise de ordem 25. Foram utilizados HMM esquerda-direita sem saltos e 5 estados, adaptação tipo CMLLR e 32 classes

de regressão.

Tabela 7.1: Testes realizados na Síntese com emoções.

	Treinamento			Adaptação			Síntese
	Nº de locuções	Locutor	Emoção	Nº de locuções	Locutor	Emoção	(MLSD em dB)
1	2000	EDU+FMA	NEU	161	EDU	ALE	4,78
2	5000	EDU+FMA+MSE+MAN+FER	NEU	10	EDU	ALE	4,65
3	5805	EDU+FMA+MSE+MAN+FER	NEU+RAI	20	FMA	ALE	4,28
4	8000	FMA+MSE+MAN+FER JUL+MIR+MRO+SAR	NEU	62	FMA	ALE	4,42
5	4000	JUL+MIR+MRO+SAR	NEU	10	MIR	ALE	4,74
6	2000	MAN+FMA	NEU	161	MAN	ALE	4,33
7	3000	EDU+FMA+MAN	NEU	161	FER	RAI	4,99
8	5000	EDU+FMA+MSE+MAN+FER	NEU	10	FER	RAI	4,89
9	4000	JUL+MIR+MRO+SAR	NEU	100	EDU	RAI	4,46
10	9000	FMA+MSE+MAN+FER+EDU JUL+MIR+MRO+SAR	NEU	20	MRO	RAI	5,30
11	5805	EDU+FMA+MSE+MAN+FER	NEU+ALE	20	EDU	TRI	4,31
12	4000	JUL+MIR+MRO+SAR	NEU	10	MRO	TRI	5,02
13	3966	JUL+MRO+SAR	NEU+ALE+RAI	62	MRO	TRI	5,01

Foram realizados 13 testes, seis com adaptação para Alegria, quatro para raiva e três para tristeza, conforme mostrado na Tabela 7.1, sempre variado os locutores e as quantidades de dados para treino e adaptação. A primeira coluna (da esquerda para a direita) enumera a sequência de testes realizados, agrupando-os de acordo com a emoção alvo. A segunda coluna, apresenta a quantidade de dados utilizados no treinamento. A terceira e quarta colunas, contém, respectivamente, o trigramas dos locutores e das emoções envolvidos no treino. Cada locutor que aparece na tabela está associado a cada emoção. Por exemplo, no último treinamento com 3966 locuções, cada um dos três locutores envolvidos contribuíram com 1000 locuções neutras (NEU), 161 com alegria (ALE) e 161 com raiva (RAI), da seguinte forma:

$$\begin{aligned}
 \text{Total de locuções} &= 3 \text{ locutores} * (1000 \text{ NEU} + 161 \text{ ALE} + 161 \text{ RAI}) \\
 &= 3966 \text{ locuções}
 \end{aligned}$$

O mesmo critério foi adotado nas colunas seguintes, para adaptação (no. de locuções, locutores e emoção).

7.4.1 Distância Mel-log espectral

Na última coluna da Tabela 7.1, mostramos uma medida de comparação entre a fala emotiva sintetizada e a fala emotiva real, do mesmo locutor, sentença e emoção. Essa medida denomina-se MLSD (*Mel Log Spectral Distance*) e vai representar a distância entre os coeficientes mel-log espectrais $\mathbf{C} = [\mathbf{c}'_1, \dots, \mathbf{c}'_T]'$ e $\bar{\mathbf{C}} = [\bar{\mathbf{c}}_1, \dots, \bar{\mathbf{c}}_T]'$,

da fala real e sintetizada, respectivamente, onde $\mathbf{c}_t = [c_t(0), \dots, c_t(M)]$ e MLSD, representada por $D(\mathbf{C}, \bar{\mathbf{C}})$, é definida como:

$$D(\mathbf{C}, \bar{\mathbf{C}}) = \frac{1}{T} \sum_{t=1}^T \sqrt{\sum_{n=1}^M (c_t(n) - \bar{c}_t(n))^2} \quad (7.18)$$

onde $c_t(n)$ representa o n -ésimo coeficiente mel-log espectral no tempo t . Os valores da MLSD são apresentados em dB e estão normalmente compreendidos entre e 4 e 5 dB de forma satisfatória, quando comparados com adaptação em sistemas dependentes e independentes do locutor [108].

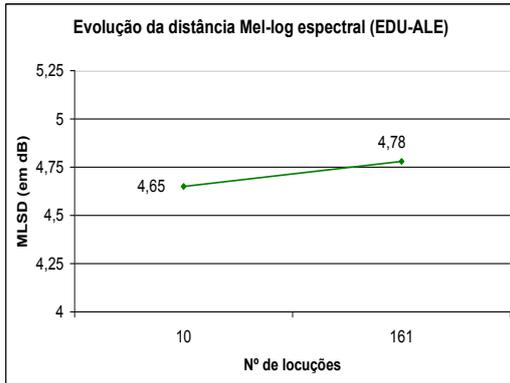
A Figura 7.9 apresenta alguns dos resultados da Tabela 7.1, usando a MLSD como medida de comparação entre a fala emotiva real e a sintetizada. Os dados são relativos a sentença 101 (“As duas análises apontam quais as praias do município próprias para banho.”), da Seção C.2, no Apêndice C. Em 7.9(a), temos a evolução da MLSD para alegria com a variação da quantidade de dados de adaptação (10 para 161 locuções) para o locutor EDU. Em 7.9(b), temos o mesmo teste para o locutor FMA, variando dado de adaptação de 20 para 62 locuções. 7.9(c) ilustra a MLSD para variação da quantidade de dado de adaptação de 10 para 161 locuções, considerando o locutor FER, porém na emoção alegria. Finalmente, em 7.9(d), temos o estado da MLSD para tristeza com variação dos dados para adaptação do locutor MRO, de 10 para 62 locuções. Podemos observar que em todos os casos ficou comprovado que a distância tende a ficar menor com a diminuição do número de dados de adaptação, exceto no último exemplo, onde a MLSD permaneceu estável.

7.4.2 Espectros gerados

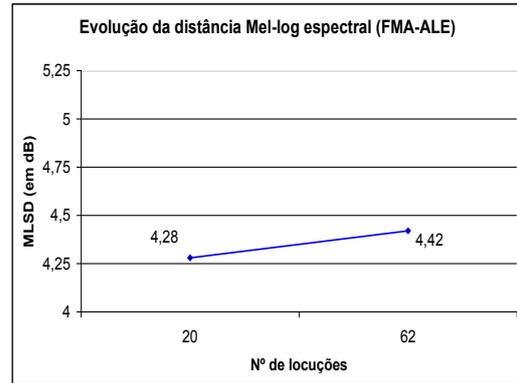
Nesta subseção, da Figura 7.10 até 7.18, apresentamos todos as sequências espectrais dos testes com síntese com emoções apresentados na Tabela 7.1. Apresentamos sempre o espectro da fala com emoção original, seguido de um ou dois espectros de fala emotiva sintetizada. Podemos observar nos testes realizados, praticamente todos os espectros da fala emotiva sintetizada estão bem próximos do espectro da fala emotiva original, principalmente aqueles demonstrados nas Subfiguras 7.10(c), 7.11(b), 7.14(c) e 7.18(b), relativos aos testes 2, 3, 8 e 12, respectivamente, que apresentaram resultados bem satisfatórios quanto a emoção sintetizada.

7.5 Conclusões e trabalhos futuros

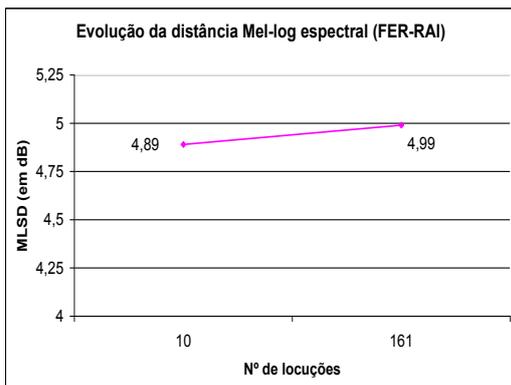
Um sistema de síntese de voz com emoções baseado em HMM foi apresentado. O sistema foi treinado com sentenças foneticamente balanceadas, gravadas por 10 locutores, cinco de cada sexo, oriundas da nova base de dados de voz para síntese e



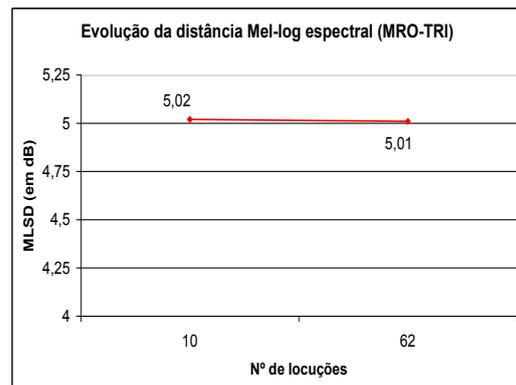
(a) MLS D nos testes 1 e 2 (Alegria)



(b) MLS D nos teste 3 e 4 (Alegria)



(c) MLS D nos testes 7 e 8 (Raiva)



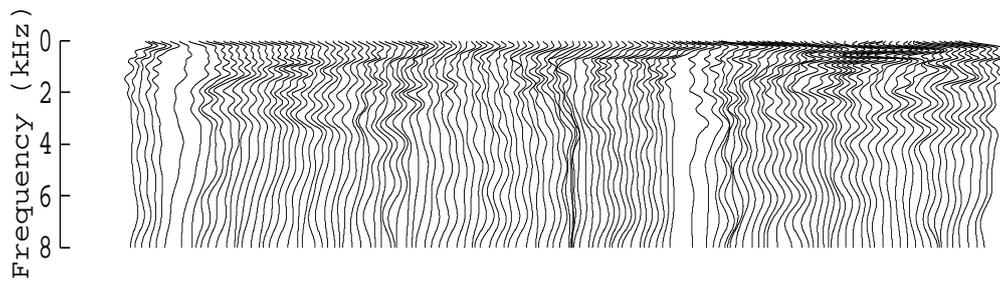
(d) MLS D nos testes 12 e 13 (Tristeza)

Figura 7.9: Evolução da MLS D, fixando o locutor e variando dado de adaptação

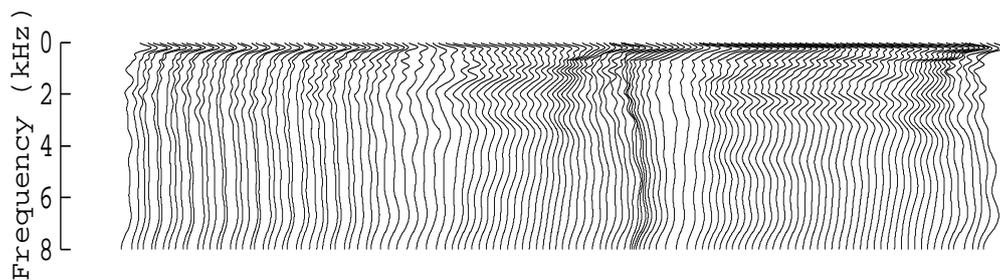
reconhecimento, descrita no Capítulo 6. Foram utilizados dados de voz com 16kHz, mono, 16 bits. Na análise do sinal, foi utilizada a janela de Hanning com 25ms e overlap de 5ms. Foram extraídos 24 coeficientes mel-cepestrais e F_0 , bem como seus delta e delta-delta, HMM esquerda-direita sem saltos e 5 estados, adaptação tipo CMLLR e 32 classes de regressão. Foram realizados 13 testes, variando-se a quantidade de dados para adaptação, além da emoção alvo. Os resultados foram analisados de duas formas: através da MLS D e da comparação espectral. Em ambas as análises, pôde ser observado que, para conseguirmos bons resultados, não foi necessária uma quantidade grande de dados para adaptação.

Como trabalhos futuros, pretendemos utilizar outras técnicas de adaptação, bus-

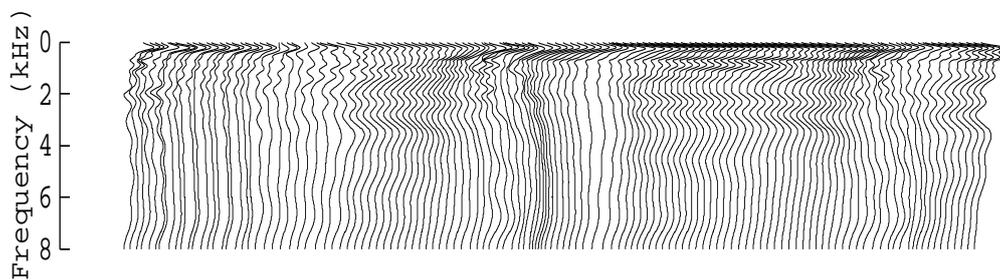
cando uma maior definição da emoção sintetizada, bem como aumentar o número de emoção, incluindo medo e indiferença. Pretendemos, também buscar uma alternativa para o efeito “buzzy” do sintetizador, devido ao uso de um vocoder, base mel-cepestral, com trem de pulso periódico.



(a) Fala com alegria original (EDU)

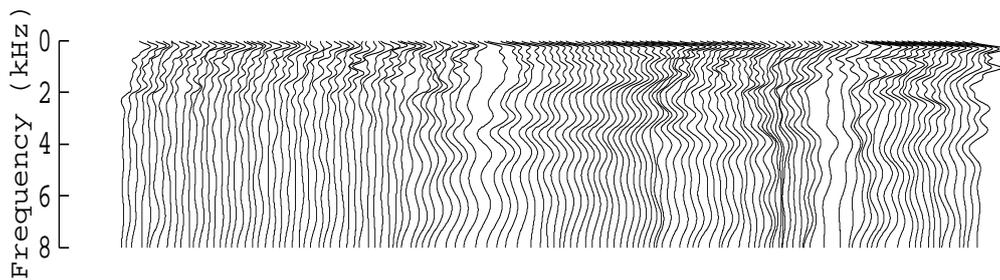


(b) Fala com alegria sintetizada: Teste 1

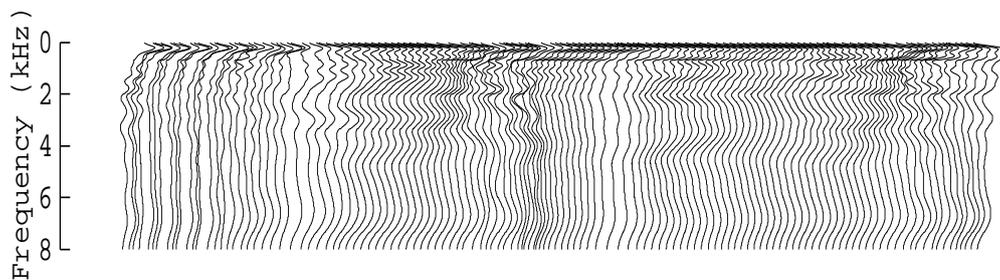


(c) Fala com alegria sintetizada: Teste 2

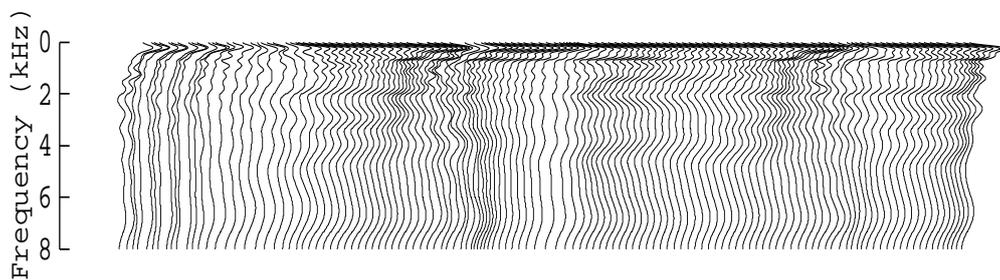
Figura 7.10: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (EDU)”



(a) Fala com alegria original (FMA)

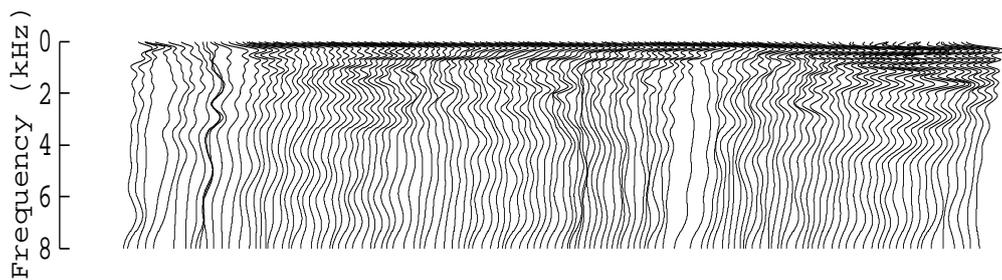


(b) Fala com alegria sintetizada: Teste 3

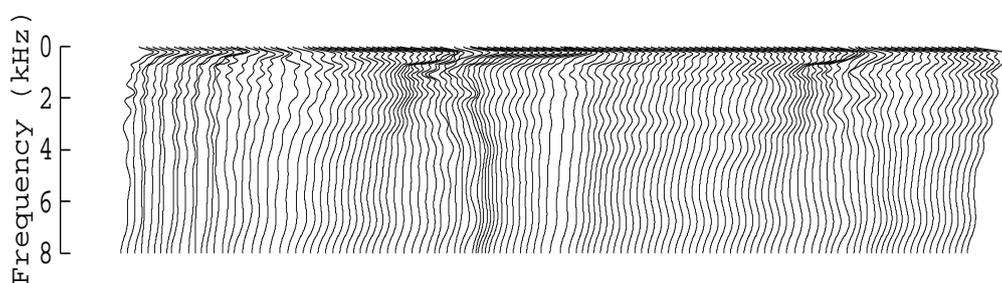


(c) Fala com alegria sintetizada: Teste 4

Figura 7.11: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (FMA)”

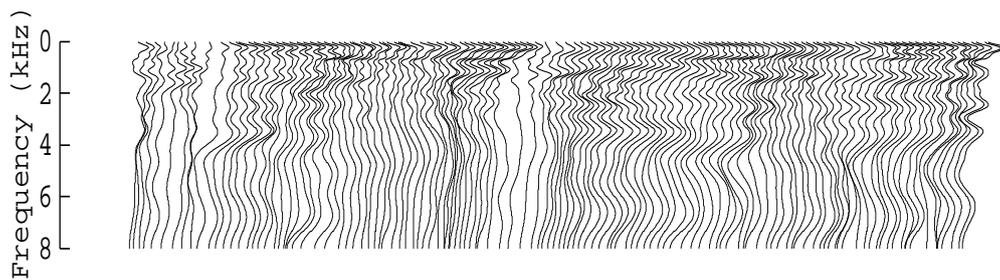


(a) Fala com alegria original

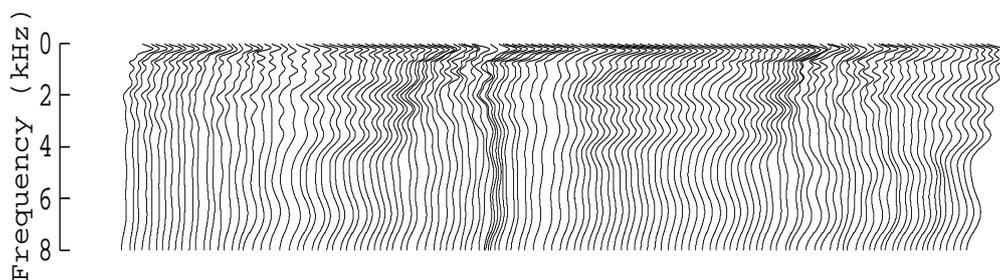


(b) Fala com alegria sintetizada: Teste 5

Figura 7.12: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MIR)”

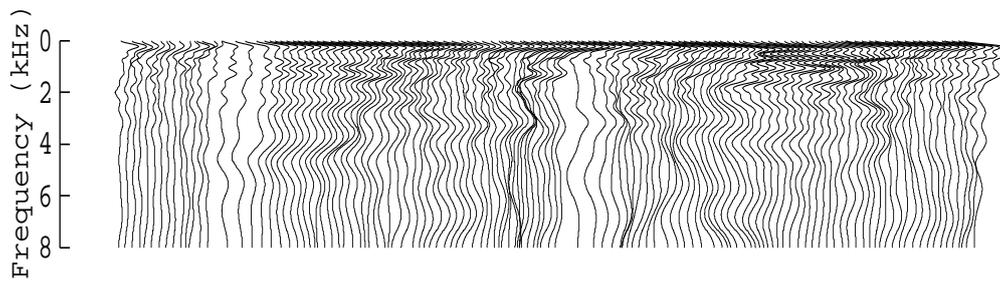


(a) Fala com alegria original

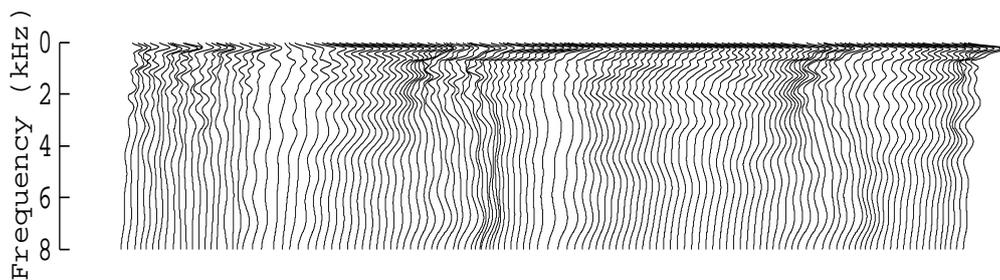


(b) Fala com alegria sintetizada: Teste 6

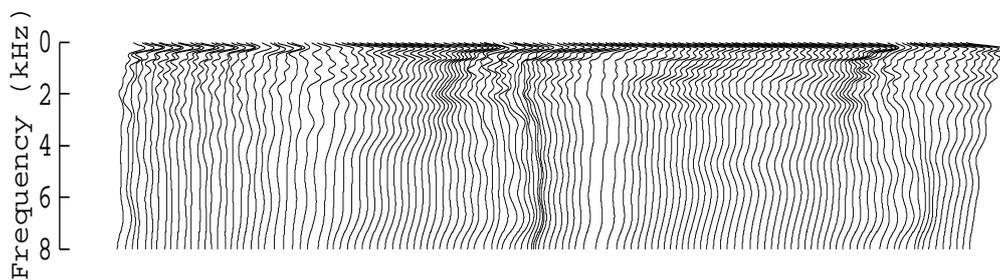
Figura 7.13: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MAN)”



(a) Fala com raiva original (FER)

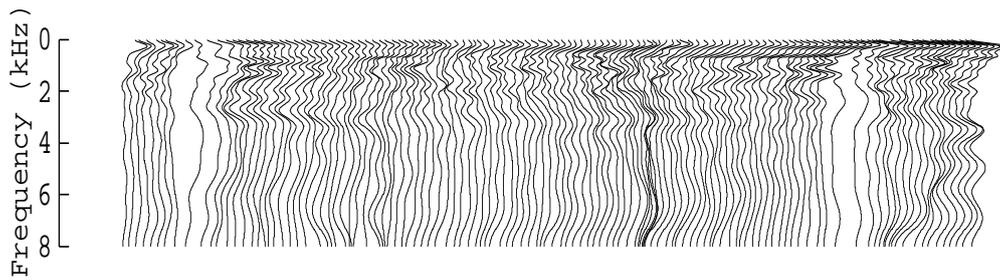


(b) Fala com raiva sintetizada: Teste 7

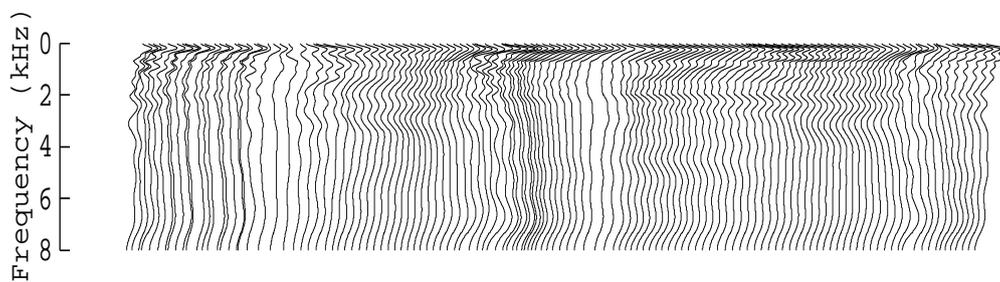


(c) Fala com raiva sintetizada: Teste 8

Figura 7.14: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (FER)”

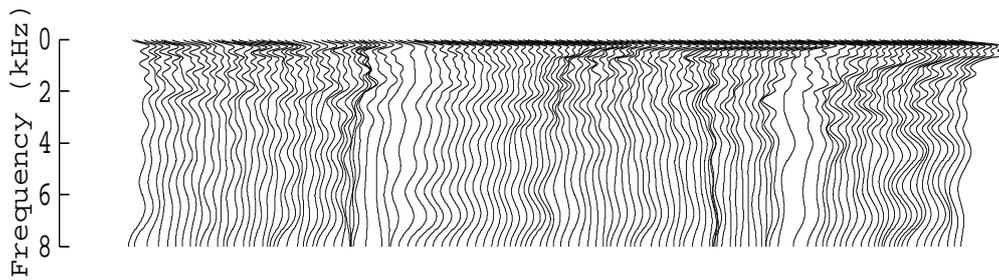


(a) Fala com raiva original (EDU)

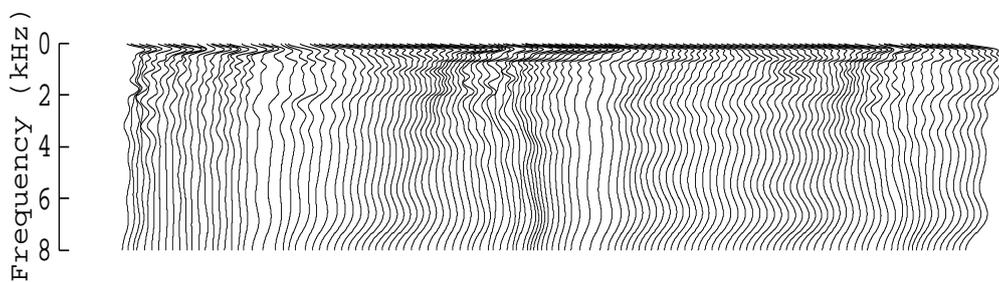


(b) Fala com raiva sintetizada: Teste 9

Figura 7.15: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (EDU)”

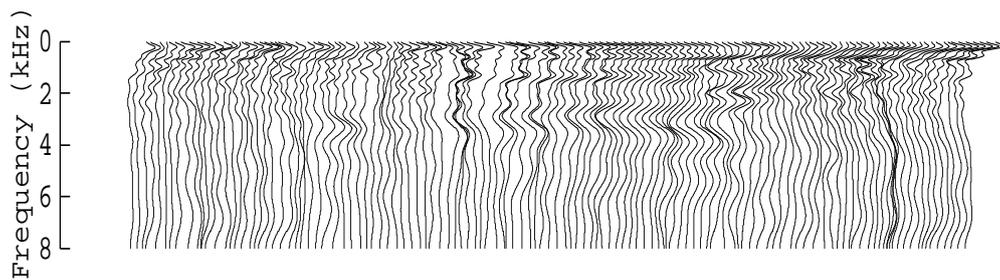


(a) Fala com raiva original (MRO)

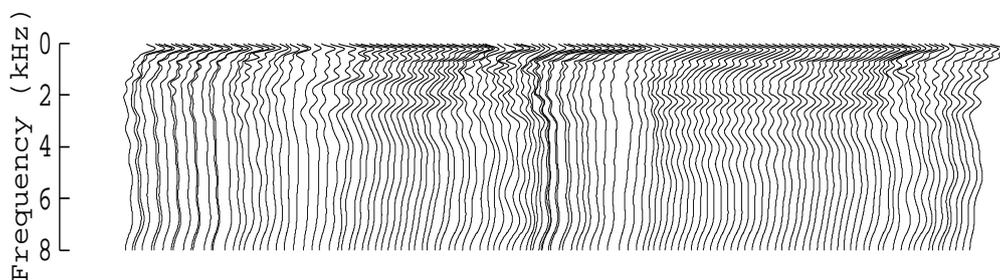


(b) Fala com raiva sintetizada: Teste 10

Figura 7.16: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MRO)”

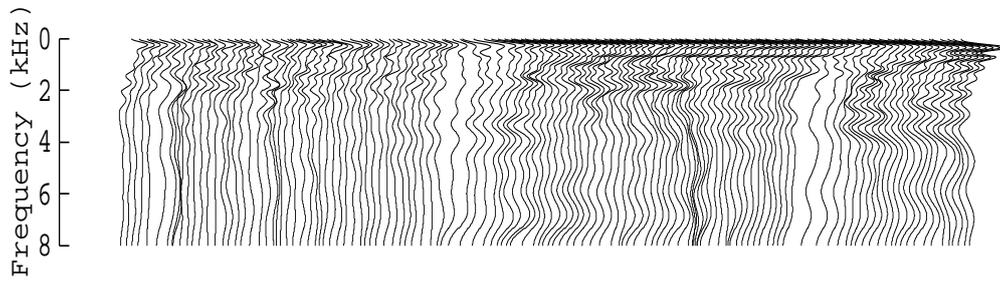


(a) Fala com tristeza original (EDU)

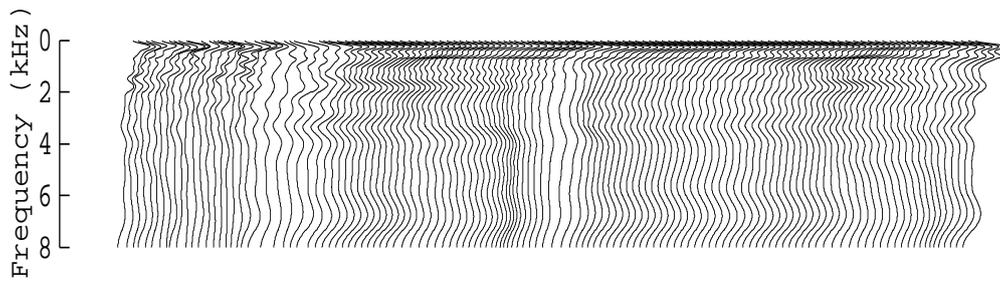


(b) Fala com tristeza sintetizada: Teste 11

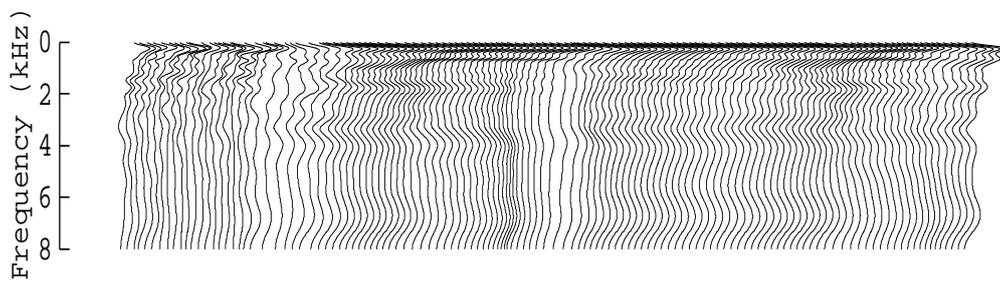
Figura 7.17: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (EDU)”



(a) Fala com tristeza original (MRO)



(b) Fala com tristeza sintetizada: Teste 12



(c) Fala com tristeza sintetizada: Teste 13

Figura 7.18: Espectro da sentença: “As duas análises apontam quais as praias do município próprias para banho. (MRO)”

Capítulo 8

Conclusões da tese e trabalhos futuros

8.1 Conclusões da tese

Nesta tese são apresentados quatro módulos aplicados ao processamento da linguagem em um TTS baseado em HMM, bem como o projeto de uma nova base de dados de voz para síntese e reconhecimento e algoritmos para síntese com emoções. Os módulos de processamento foram discriminados ao longo dos capítulos.

Primeiramente, foi discutida a questão da ambiguidade entre homógrafos do PB (Capítulo 2). Foi desenvolvido e implementado um algoritmo, baseado em regras linguísticas, para promover a desambiguação de palavras homógrafas heterófonas existentes no português falado no Brasil. O conjunto de regras é composto de 23 algoritmos, separados por tipo, capazes de solucionar o problema da ambiguidade de 111 pares de homógrafos existentes na base de dados. As regras implementadas permitem uma análise morfossintática, se as palavras homógrafas pertencem a diferentes classes gramaticais, e uma análise semântica, se as palavras pertencem a mesma classe gramatical. O conjunto de regras foi testado com extratos de texto, com naturezas distintas, provenientes de três bases de dados. Essa pluralidade na natureza dos textos em teste foi introduzida para verificar generalização e robustez do desambiguador. O primeiro teste foi realizado contendo 1.564.591 palavras, provenientes da base de dados CETEN-Folha. Foram identificados 20.308 homógrafos (1,30% do texto analisado) e foi alcançado um índice de acerto na desambiguação de 99,00%. Partindo para um teste com uma linguagem mais elaborada, foi utilizado como fonte de dados a Bíblia Sagrada. Este teste apresentou um conteúdo com mais de 750.000 palavras e 7.904 homógrafos foram identificados (1,05% do texto processado), tendo sido verificada uma taxa de acerto de 99,00%. Por fim, foi feito um terceiro teste com a obra literária de Machado de Assis, “Dom Casmurro”,

com uma linguagem também elaborada, mas impregnada de regionalismo de época (ocorre no Rio de Janeiro do Segundo Império), além de ser regida em primeira pessoa, sendo recorrente o aparecimento de homógrafos dos tipos 1 e 2 (Verbos). O conteúdo apresenta quase 70.000 palavras, onde 510 homógrafos (0,73% do texto) foram verificados, totalizando uma taxa de acerto de 99,02%.

Em seguida, no Capítulo 3, foi abordada a questão da marcação de tonicidade, fator preponderante para a correta transcrição fonética, juntamente com a questão da desambiguação de homógrafos que, de certa forma, já realiza uma marcação prévia da tonicidade em palavras com mesma grafia, numa proporção que gira em torno de 1,00%, como verificado nesta tese. O conjunto de algoritmos é baseado em regras linguísticas, num total de 19 regras, que mostraram-se capazes de realizar a identificação da vogal tônica das palavras do PB. O algoritmo proposto foi implementado e testado, também com texto extraído do CETEN-Folha, composto de 15.974 palavras e uma taxa de acerto de 98,90% foi atingida. Um percentual de acerto considerável, atentando para o tamanho do teste. Além do mais, foi verificado que os erros do conjunto de regras implementado, na verdade poderiam ser considerados ainda menores (em torno de 0,1%), se pudéssemos considerar apenas os erros de falhas na marcação, pois os demais erros envolvem nomes próprios, estrangeirismos e acrônimos, que na sua maioria, não seguem normas e regras linguísticas.

No Capítulo 4, foi apresentado um algoritmo baseado em regras linguísticas, num total de 25 regras, que mostraram uma boa eficiência na capacidade de realizar a separação em sílabas de palavras do PB. O algoritmo que propomos foi implementado e testado com extrato de texto da base de dados CETEN-Folha, composto de 18.037 sílabas e uma taxa de acerto de 99,41% foi alcançada. Aqui, também é importante considerar que o erro real do algoritmo pode ser reduzido a menos de 0,5% (relativos às falhas de separação), já que estrangeirismos e acrônimos, em geral, não tem uma regência regular pelas normas da linguística portuguesa.

No Capítulo 5 foi também implementado um algoritmo capaz de realizar a transcrição fonética de palavras do PB, realizando a conversão G2P aplicada ao sistema TTS. O algoritmo proposto foi implementado e testado com um extrato de texto da base de dados CETEN-Folha, composto por cerca de 40.000 fones, onde apenas 574 foram incorretamente transcritos. Esse número representou uma taxa de acerto global de 98,57%, constatado que a aplicação de regras linguísticas é uma boa escolha para a conversão G2P no PB.

O Capítulo 6 apresenta um breve comentário sobre a questão das emoções da fala no contexto da síntese de voz. É comentado também a respeito do nosso projeto, financiado pelo CNPq, para desenvolvimento da base de dados para síntese e reconhecimento com emoções. Este projeto constituiu-se elemento essencial no estudo das emoções através da fala, já que não existe nenhum material documentado

a esse respeito disponível para o PB. Uma nova base de dados de voz para síntese e reconhecimento foi gravada, contendo cerca de 300 horas de gravação. A base, composta por 20 locutores renomados na mídia nacional, 10 do sexo masculino e 10 do sexo feminino, foi parcialmente utilizada nesta tese, cinco do sexo masculino e cinco do sexo feminino. Um teste de validação foi realizado através da internet e uma taxa de percepção emotiva de 81,8% foi verificada. Os maiores índices de confusão ocorreram com a fala neutra sendo confundida com as emoções. A justificativa é o fato do conteúdo textual ser jornalístico e, naturalmente, impregnado com alguma emoção intrínseca, apenas percebida pelo locutor no ato da gravação. A confusão reversa, das emoções com a fala neutra, também pode ser justificada pelo conteúdo textual, pois muitas das emoções não eram contextualizadas e, depois de algumas horas de gravação, o locutor perdia sua referência emotiva.

Finalmente, no Capítulo 7, apresentamos os detalhes do sistema de síntese de voz com emoções implementado. O sistema desenvolvido é composto de três estágios, treinamento, adaptação e síntese. O treinamento foi realizado com sentenças foneticamente balanceadas, gravadas por 10 locutores, cinco de cada sexo, oriundas da nova base de dados de voz para síntese e reconhecimento, descrita no Capítulo 6. Os dados de voz foram utilizados com 16kHz, mono, 16 bits. Na composição dos vetores de características, a análise do sinal foi realizada utilizando-se a janela de Hanning com 25ms e overlap de 5ms. Foram extraídos 24 coeficientes mel-cepstrais e $\log F_0$, além dos seus delta e delta-delta. HMM do tipo esquerda-direita sem saltos e 5 estados, adaptação tipo CMLLR e 32 classes de regressão. Os testes envolveram 13 experimentos, variando-se a quantidade de dados de adaptação e a emoção alvo. Os resultados foram analisados através da comparação da MLSLSD entre a fala emotiva real e a sintetizada, bem como através da comparação espectral dos mesmos. Nas duas análises, observamos que com poucos dados de adaptação das emoções, bons resultados já são atingidos e tendem a melhorar com menores quantidades.

8.2 Trabalhos futuros

Atentando para os resultados obtidos em cada módulo, vemos que uma manipulação adequada de estrangeirismos trará melhorias significativas para este trabalho, dada a quantidade de palavras já inseridas no cotidiano da língua portuguesa. Isso é preponderante para a síntese com naturalidade e inteligibilidade.

Alguns pequenos ajustes precisam ainda ser realizados, nos algoritmos que precedem a transcrição fonética, tendo em vista que erros nestes módulos vão acarretar, novos erros no G2P, principalmente com relação a questão da tonicidade das vogais [e] e [o].

Com relação a base de dados gravada, a edição primária de metade da base ainda

precisa ser concluída, além das edições secundárias e terciárias de todo conteúdo. Pretendemos, também, dar maior amplitude na avaliação, utilizando mais análises perceptivas ao expandirmos o número de avaliadores, bem como acrescentar mais duas opções de avaliação ao ouvirem as locuções: “não identificada” e “outras”.

Quanto ao nosso sistema de síntese de voz com fala emotiva, pretendemos analisar o desempenho do sistema com novas técnicas de adaptação, buscando uma maior definição da emoção sintetizada, bem como aumentar o número de emoções, incluindo “medo” e “indiferença”. Pretendemos, também buscar uma alternativa para o efeito “buzzy” do sintetizador, devido ao uso de um vocoder, base mel-cepestral, com trem de pulso periódico.

Referências Bibliográficas

- [1] DARWIN, C. *The Expression of Emotion in Man and Animals*. New York, Project Gutenberg EText p38 - electronic reprint of the 1899 re-release of Darwin's 1872 publication by D. Appleton and Co. of New York/London., 1899.
- [2] FAIRBANKS, G., PRONOVOST, W. "Vocal Pitch During Simulated Emotion", *Science*, v. 88, n. 2286, pp. 382–383, out. 1938.
- [3] CAHN, J. E. "From sad to glad: emotional computer voices". In: *Proc. of Speech Tech 1988, Voice Input/Output Applications Conference and Exhibition*, pp. 35–37, 1988.
- [4] BLACK, A. W. "Unit selection and emotional speech". In: *European Conference on Speech Communication and Technology*, pp. 1649–1652, 2003.
- [5] CHU, M., LU, S. "High Intelligibility and Naturalness Chinese TTS System and Prosodic Rules". In: *Proc. of XIII International Congress of Phonetic*, pp. 334–337, 1995.
- [6] PEARSON, S., MORAN, H., HATA, K., et al. "Combining concatenation and formant synthesis for improved intelligibility and naturalness in text-to-speech systems". In: *Second ISCA/IEEE Workshop on Speech Synthesis*, pp. 69–72, 1994.
- [7] KLATT, D. H. "The Klattalk text-to-peech conversion system". In: *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1589–1592, 1986.
- [8] BIRKHOLZ, P. "Articulatory Synthesis of Singing". In: *Proc. of Interspeech*, 2007.
- [9] TEIXEIRA, A. J. S., MARTINEZ, R., SILVA, L. N., et al. "Simulation of human speech production applied to the study and synthesis of European Portuguese", in *EURASIP Journal on Applied Signal Processing*, v. 2005, pp. 1435–1448, 2005.

- [10] URBANCZYK, S. C., EADY, S. J. “Assignment of syllable stress in a demisyllable-based text-to-speech synthesis system”, in *Proc. of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp. 467–470, 1989.
- [11] EIDE, E., AARON, A., BAKIS, R., et al. “A corpus-based approach to <AHEM/> expressive speech synthesis”. In: *Proc. 5th ISCA Speech Synthesis Workshop*, 2004.
- [12] DOSVOX. “Projeto DOSVOX”. <http://intervox.nce.ufrj.br/dosvox>, Acessado em 2010.
- [13] MBROLA. “The MBROLA Project”. <http://tcts.fpms.ac.be/synthesis/mbrola.html>, Acessado em 2010.
- [14] DUTOIT, T. “High Quality Text-to-Speech Synthesis: A Comparison of Four Candidate Algorithms”. In: *Proc. of IEEE*, pp. 565–568, 1994.
- [15] HOLZAPFEL, M., HOFFMANN, R., HOGE, H. “A wavelet-domain PSOLA approach”. In: *Third ESCA/COCOSDA Workshop on Speech Synthesis*, pp. 283–286, 1998.
- [16] YU, J., ZHANG, M., TAO, J., et al. “A novel HMM-based TTS system using both continuous HMMs and discrete HMMs”. In: *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 709–712, 2007.
- [17] BLACK, A. W., ZEN, H., TOKUDA, K. “Statistical parametric speech synthesis”. In: *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1229–1232, 2007.
- [18] KRSTULOVIC, S., HUNECKE, A., SCHROEDER, M. “A HMM-based speech synthesis system applied to German and its adaptation to a limit set of expressive football announcements”. In: *Proc. of Interspeech*, 2007.
- [19] ABDEL-HAMID, O., ABDON, S., RASHWAN, M. “Improving Arabic HMM based speech synthesis quality”. In: *Proc. of Interspeech*, pp. 1332–1335, 2006.
- [20] MARTINCIC-IPSIC, S., IPSIC, I. “Croatian HMM-based speech synthesis”, *Journal of Computing and Information Technology*, v. 14, n. 4, pp. 307–313, 2006.
- [21] WU, Y.-J., WANG, R. H. “HMM-based trainable speech synthesis for Chinese”, *Journal of Chinese Information Processing*, v. 20, n. 4, pp. 75–81, 2006.

- [22] BARROS, M., MAIA, R., TOKUDA, K., et al. “HMM-based European Portuguese speech synthesis”. In: *Proc. of Interspeech*, pp. 2581–2584, 2005.
- [23] ZEN, H., TODA, T. “An overview of Nitech HMM-based speech synthesis for Blizzard Challenge 2005”. In: *European Conference on Speech Communication and Technology*, 2005.
- [24] BULUT, M., LEE, S., NARAYANAN, S. “A Statistical Approach for Modeling Prosody Features using POS Tags for Emotional Speech Synthesis”, *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, v. 4, pp. 1237–1240, 2007.
- [25] BULUT, M., LEE, S., NARAYANAN, S. “Analysis of emotional speech prosody in terms of part of speech tags”. In: *Proc. of Interspeech*, 2007.
- [26] BURKHARDT, F. “Emofilt: the simulation of emotional speech by prosody-transformation”. In: *Proc. of Interspeech*, pp. 509–512, 2005.
- [27] SCHRÖDER, M. “Emotional Speech Synthesis - A Review”. In: *European Conference on Speech Communication and Technology*, pp. 561–564, 2001.
- [28] MAIA, R. S., ZEN, H., TOKUDA, K., et al. “A HMM-based Brazilian Portuguese Speech Synthesizer and its Characteristics.” *Revista da Sociedade Brasileira de Telecomunicações*, v. 21, n. 2, pp. 58–71, ago. 2006.
- [29] PINTO, G. O., BARBOSA, F. L. F., RESENDE JR., F. G. V. “A Brazilian Portuguese TTS Based on HMMs”. In: *Proc. of IEEE Int. Telecomm. Symposium (ITS)*, 2002.
- [30] SILVA, D. C., LIMA, A. A., MAIA, R., et al. “A rule-based grapheme-phone converter and stress determination for Brazilian Portuguese natural language processing”, *Proc. of the International Telecommunications Symposium*, pp. 992–996, 2006.
- [31] COLOTTE, V., BEAUFORT, R. “Linguistic features weighting for a text-to-speech system without prosody model”. In: *Proc. of the 9th European Conference on Speech Communication and Technology*, pp. 2549–2552, 2005.
- [32] MATSUOKA, K., TAKEISHI, E., ASANO, H., et al. “Natural language processing in a Japanese text-to-speech system for written-style texts”. In: *Proc. of the 3th IEEE Workshop on Interactive Voice Technology for Telecommunications Applications*, pp. 33–36, 1996.

- [33] WEERASINGHE, R., WASALA, A., GAMAGE, K. “A Rule Based Syllabification Algorithm for Sinhala”. In: *Lecture Notes in Computer Science*, pp. 438–449, 2005.
- [34] ZOVATO, E., PACCHIOTTI, A., QUAZZA, S., et al. “Towards emotional speech synthesis: a rule based approach”. In: *5th ISCA Speech Synthesis Workshop*, pp. 219–220, 2004.
- [35] KLATT, D. H. “Structure of a phonological rule component for a synthesis-by-rule program”, *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, v. 24, n. 5, pp. 391–398, out. 1976.
- [36] BRAGA, D. *Algoritmos de processamento da linguagem natural para sistemas de conversão texto-fala em português*. Tese de Doutorado, Departamento de Galego-Português, Francés e Lingüística, Facultade de Filoxía da Universidade da Coruña, A Coruña, Espanha, fev. 2008.
- [37] BRAGA, D., FREITAS, D., FERREIRA, H. “Processamento Linguístico Aplicado à Síntese da Fala”. In: *3o. Congresso Luso-Moçambicano de Engenharia*, 2003.
- [38] “Speech Assessment Methods Phonetic Alphabet (SAMPA)”. <http://www.phon.ucl.ac.uk/home/sampa/home.htm>, Acessado em 2010.
- [39] YAROWSKY, D. “Homograph disambiguation in text-to-speech synthesis”. In: *In Jan van Santen, Richard Sproat, Joseph Olive, and Julia Hirschberg, editors, Progress in Speech Synthesis*, pp. 157–172, 1997.
- [40] TESPASIT, V., CHAROENPORNSAWAT, P., SORNLERTLAMVANICH, V. “A Context-Sensitive Homograph Disambiguation in Thai Text-to-Speech Synthesis”. In: *In Proc. of HLT-NAACL'03 - Human Language Technology and North American Chapter of the Association for Computational Linguistics Conference*, 2003.
- [41] DONG, H., TAO, J., XU, B. “Grapheme-to-phoneme conversion in Chinese TTS system”. In: *International Symposium on Chinese Spoken Language Processing*, pp. 165–168, 2004.
- [42] RIBEIRO, R., OLIVEIRA, L. C., TRANCOSO, I. “Using Morphosyntactic Information in TTS Systems: Comparing Strategies for European Portuguese”. In: *PROPOR'2003- 6th Workshop on Computational Processing of the Portuguese Language*, pp. 143–150, 2003.

- [43] RIBEIRO, R., OLIVEIRA, L. C., TRANCOSO, I. “Morphosyntactic Disambiguation for TTS Systems.” In: *Proc. of the 3rd Intl. Conf. on Language Resources and Evaluation*, v. 5, pp. 1427–1431, 2002.
- [44] SEARA, I. C., KAFKA, S. G., KLEIN, S., et al. “Alternância vocálica das formas verbais e nominais do Português Brasileiro para aplicação em conversão Texto-Fala”, *Revista da Sociedade Brasileira de Telecomunicações*, v. 17, n. 1, pp. 79–85, 2002.
- [45] SEARA, I. C., KAFKA, S. G., KLEIN, S., et al. “Considerações sobre os Problemas de Alternância Vocálica das Formas Verbais do Português Falado no Brasil para Aplicação em um Sistema de Conversão Texto-Fala”. In: *Anais do XIX Simpósio Brasileiro de Telecomunicações*, 2001.
- [46] FERRARI, L., BARBOSA, F., RESENDE JR., F. G. V. “Construções gramaticais e sistemas de conversão texto-fala: o caso dos homógrafos.” In: *Proc. of the International Conference on Cognitive Linguistics*, 2003.
- [47] BARBOSA, F., FERRARI, L., RESENDE JR., F. G. V. “A methodology to analyze homographs for a Brazilian Portuguese TTS system”. In: *PROPOR'2003 - 6th Workshop on Computational Processing of the Portuguese Language*, Springer-Verlag, 2003.
- [48] BARBOSA, F., FERRARI, L., RESENDE JR., F. G. V. “A distinção entre homógrafos heterófonos em sistemas de conversão texto-fala”. In: *Processamento da Linguagem, Cultura e Cognição: estudos de lingüística cognitiva*, 2003.
- [49] “Corpus de Extractos de Textos Eletrônicos NILCS/Folha de São Paulo (CETEN-Folha)”. <http://acdc.linguateca.pt/cetenfolha>, Acessado em 2010.
- [50] Bíblia. “A Bíblia Sagrada”. <http://www.culturabrasil.pro.br/zip/biblia.rtf>, Acessado em 2008.
- [51] DomCasmurro. “Dom Casmurro - Machado de Assis”. <http://www.machadodeassis.org.br/>, Acessado em 2009.
- [52] JSpell. “The JSpell Project”. <http://natura.di.uminho.pt/wiki/index.cgi?jspell>, Acessado em 2009.
- [53] Wordnet. “Wordnet”. <http://wordnet.princeton.edu/>, Acessado em 2006.

- [54] SILVA, D. C., BRAGA, D., RESENDE JR., F. G. V. “Desambiguação de Homógrafos Heterófonos no Português Brasileiro”, *Anais do Simpósio Brasileiro de Telecomunicações (SBrT)*, 2009.
- [55] AULL, A. M., ZUE, V. W. “Lexical stress determination and its application to large vocabulary speech recognition”, *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, v. 10, pp. 1549–1552, 1985.
- [56] YING, G. S., JAMIESON, L. H., CHEN, R., et al. “Lexical stress detection on stress-minimal word pairs”, *Proc. of the Int. Conf. on Spoken Language Processing (ICSLP)*, pp. 1612–1615, 1996.
- [57] XIE, H., ANDREAE, P., ZHANG, M., et al. “Detecting stress in spoken English using Decision Trees and Support Vector Machines”, *ACSW Frontiers '04: Proc. of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation*, v. 32, pp. 145–150, 2004.
- [58] OANCEA, E., BADULESCU, A. “Stressed Syllable Determination for Romanian Words within Speech Synthesis Applications”, *International Journal of Speech Technology*, v. 5, n. 3, pp. 237–246, 2002.
- [59] MADUREIRA, S., BARBOSA, P., FONTES, M., et al. “Post-stressed syllables in Brazilian Portuguese as markers”, in *Proc. ICPHS'99*, pp. 917–920, 1999.
- [60] ARANTES, P., BARBOSA, P. A. “Secondary stress in Brazilian Portuguese: the interplay between production and perception studies”, in *Proc. of the Speech Prosody*, pp. 73–76, 2006.
- [61] WERTZNER, H. F., SCHREIBER, S., AMARO, L. “Análise da frequência fundamental, jitter, shimmer e intensidade vocal em crianças com transtorno fonológico”, *Revista Brasileira de Otorrinolaringologia*, v. 71, n. 5, pp. 582–588, set. 2005.
- [62] TEIXEIRA, J. P., PAULO, E. R., FREITAS, D., et al. “Acoustical characterisation of the accented syllable in Portuguese; a contribution to the naturalness of speech synthesis”. In: *European Conference on Speech Communication and Technology*, 1999.
- [63] BECHARA, E. *Moderna Gramática Portuguesa*. Rio de Janeiro, RJ, Brazil, Lucerna, 2002.

- [64] TEIXEIRA, J. P., FREITAS, P. G. D. “Divisão silábica automática do texto escrito e falado”. In: *Proc. of PROPOR’2000*, 2000.
- [65] OLIVEIRA, C., MOUTINHO, L., TEIXEIRA, A. “On European Portuguese Automatic Syllabification”. In: *Proc. of Interspeech’2005*, 2005.
- [66] OLIVEIRA, C., MOUTINHO, L. C., TEIXEIRA, A. “On Automatic European Portuguese Syllabification”. In: *9th European Conference on Speech Communication and Technology*, pp. 2933–2936, 2005.
- [67] SEARA JR., R., SEARA, I. C., KAFKA, S., et al. “Parâmetros Lingüísticos utilizados para a geração automática de prosódia em sistemas de síntese de fala”. In: *Anais do Simpósio Brasileiro de Telecomunicações (SBrT)*, 2004.
- [68] KAFKA, S. G., PACHECO, F. S., SEARA, I. C., et al. “Utilização de segmentos transicionais homorgânicos em síntese de fala concatenativa”. In: *Anais do Congresso Brasileiro de Automática*, pp. 2742–2747, 2002.
- [69] MOBIUS, B. “Word and syllable models for German text-to-speech synthesis”. In: *Proc. of the Third ESCA Workshop on Speech Synthesis*, 1998.
- [70] O’DELL, M., LENNES, M., WERNER, S., et al. “Looking for rhythm in conversational speech”. In: *Proc. of the International Congress of Phonetic Sciences (ICPhS)*, 2007.
- [71] RAMUS, F., DUPOUX, E., MEHLER, J. “The psychological reality of rhythm classes: perceptual studies”. In: *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, pp. 337–342, 2005.
- [72] SILVA, D. C., BRAGA, D., RESENDE JR., F. G. V. “Separação das Sílabas e Determinação da Tonicidade no Português Brasileiro”, *Anais do Simpósio Brasileiro de Telecomunicações (SBrT)*, 2008.
- [73] BARBOSA, F., PINTO, G., RESENDE JR., F. G. V., et al. “Grapheme-phone transcription algorithm for a Brazilian Portuguese TTS”. In: *Proc. of 6th. Workshop on Computational Processing of Written and Spoken Portuguese (PROPOR)*, 2003.
- [74] “The Festival Speech Synthesis System”. <http://www.festvox.org/festival/>, Acessado em 2011.
- [75] RABINER, L., SCHAFER, R. W. *Digital Processing of Speech Signals*. New Jersey, USA, Prentice-Hall, 1978.

- [76] ALBANO, E. C., MOREIRA, A. A. “Archisegment-based letter-to-phone conversion for concatenative speech synthesis in Portuguese”. In: *Proc. of the Int. Conf. on Spoken Language Processing (ICSLP)*, 1996.
- [77] SPOLTECH. “Advancing human language technology in Brazil and the United States through collaborative research on Portuguese spoken language systems”. Final Report, jul. 2001. Federal University of Rio Grande do Sul, University of Caxias do Sul, Colorado University, and Oregon Graduate Institute.
- [78] BARBOSA, F., MAIA, R., RESENDE, F. G. “Análise comparativa do impacto da classe gramatical em sistemas TTS baseados em HMMs”. In: *Anais do Simpósio Brasileiro de Telecomunicações (SBrT)*, 2004.
- [79] BURKHARDT, F., PAESCHKE, A., ROLFES, M., et al. “A Database of German Emotional Speech”. In: *Proc. of Interspeech*, pp. 1517–1520, 2005.
- [80] CAMPBELL, N. “Databases of Emotional Speech”, *Proc ISCA (International Speech Communication Association) ITRW on Speech and Emotion*, pp. 34–38, 2000.
- [81] VIOLA, I. C., MADUREIRA, S. “Voice quality and Speech Expressiveness”. In: *Paralinguistic Speech07 - 16th International Conference on Phonetic Sciences*, pp. 65–70, 2007.
- [82] MATTE, A. C. F. “Relating emotional content to speech rate in Brazilian Portuguese”. In: *Proc. of the Speech Prosody*, 2004.
- [83] OUDEYER, P.-Y. “The production and recognition of emotions in speech: features and algorithms”, *International Journal in Human-Computer Studies*, v. 59, pp. 157–183, 2002.
- [84] BARBOSA, P. A. “From syntax to acoustic duration: A dynamical model of speech rhythm production”, *Speech Communication*, v. 49, n. 9, pp. 725–742, 2007.
- [85] ALBANO, E. C., AQUINO, P. A. “Linguistic criteria for building and recording units for concatenative speech synthesis in Brazilian Portuguese”. In: *European Conference on Speech Communication and Technology*, pp. 725–728, 1997.
- [86] TEIXEIRA, J. P., FREITAS, D., BRAGA, D., et al. “Phonetic events from the labeling the European Portuguese database for speech synthesis,

Proc.FEUP/IPB-DB”. In: *European Conference on Speech Communication and Technology*, 2001.

- [87] ALCAIM, A., SOLEMICZ, J. A., MORAIS, J. A. “Frequência de ocorrência dos fones e listas de frases foneticamente balanceadas para o português falado no Rio de Janeiro”, *Revista da Sociedade Brasileira de Telecomunicações*, v. 7, n. 1, pp. 23–41, dez. 1992.
- [88] CAMPBELL, N. “The Recording of Emotional Speech; JST/CREST database research”, *Proc. of LREC2002 (Language Resources & Evaluation Conference)*, v. 6, pp. 2029–2032, 2002.
- [89] CIRIGLIANO, R. J. R., MONTEIRO, C., BARBOSA, F. L., et al. “Um Conjunto de 1000 Frases Foneticamente Balanceadas para o Português Brasileiro Obtido Utilizando a Abordagem de Algoritmos Genéticos”. In: *Anais do XXII Simpósio Brasileiro de Telecomunicações (SBrT 2005)*, 2005.
- [90] MARTINS, R. A. C. “*Projeto, Gravação e Edição de Base de Voz para Aplicações em Síntese e Reconhecimento da Fala*, Trabalho de Conclusão de Curso, Departamento de Engenharia Elétrica, UFRJ, Rio de Janeiro, Brasil”. mar. 2011.
- [91] protocols. “Pro Tools 9 Software”. <http://www.avid.com/US/products/Pro-Tools-Software>, Acessado em 2010.
- [92] YOSHIMURA, T., TOKUDA, K., MASUKO, T., et al. “Speaker interpolation for HMM-based speech synthesis system”, *Acoustical Science and Technology*, v. 21, n. 4, pp. 199–206, 2000.
- [93] SHICHIRI, K., SAWABE, A., YOSHIMURA, T., et al. “Eigenvoices for HMM-based speech synthesis”, *Proc. of the Int. Conf. on Spoken Language Processing (ICSLP)*, 2002.
- [94] YAMAGISHI, J., TACHIBANA, M., MASUKO, T., et al. “Speaking style adaptation using context clustering decision tree for HMM-based speech synthesis”, *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 2004.
- [95] TOKUDA, K., ZEN, H., BLACK, A. “An HMM-based speech synthesis applied to English”, in *Proc. of IEEE Workshop in Speech Synthesis*, 2002.
- [96] MAIA, R., ZEN, H., TOKUDA, K., et al. “Towards the development of a Brazilian Portuguese text-to-speech system based on HMM”, in *Proc.*

of the *European Conf. on Speech Communication and Technology (EUROSPEECH)*, 2003.

- [97] ZEN, H., TOKUDA, K., MASUKO, T., et al. “Hidden semi-Markov model based speech synthesis”, *Proc. of the Int. Conf. on Spoken Language Processing (ICSLP)*, 2004.
- [98] JURAFSKY, D., MARTIN, J. H. *SPEECH and LANGUAGE PROCESSING: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. New Jersey, USA, Prentice-Hall, 2006.
- [99] RABINER, L., JUANG, B.-H. *Fundamentals of Speech Recognition*. New Jersey, USA, Prentice-Hall, 1993.
- [100] DELLER, J. R., PROAKIS, J. G., HANSEN, J. H. L. *Discrete-time Processing of Speech Signals*. New Jersey, USA, Prentice-Hall, 1987.
- [101] TOKUDA, K., MASUKO, T., MIYAZAKI, N., et al. “Hidden markov models based on multi-space probability distribution for pitch pattern modeling”. In: *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 229–232, 1999.
- [102] YOUNG, S. J., ODELL, J., WOODLAND, P. “Tree-based state tying for high accuracy acoustic modeling”. In: *In Proc. ARPA Human Language Technology Workshop*, pp. 307–312, 1994.
- [103] SHINODA, K., WATANABE, T. “Acoustic modeling based on the MDL criterion for speech recognition”. In: *In Proc. EUROSPEECH-97*, pp. 99–102, 1997.
- [104] TODA, T., TOKUDA, K. “Speech Parameter Generation Algorithm Considering Global Variance for HMM-Based Speech Synthesis”. In: *In Proc. EUROSPEECH-05*, pp. 2801–2804, 2005.
- [105] GALES, M. J. F., WOODLAND, P. C. “Mean and Variance Adaptation within the MLLR Framework”, *Computer Speech and Language*, v. 10, n. 4, pp. 249–264, out. 1996.
- [106] TOKUDA, K., KOBAYASHI, T., IMAI, S. “Speech parameter generation from HMM using dynamic features”, *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 660–663, maio 1995.
- [107] FUKADA, T., TOKUDA, K., KOBAYASHI, T., et al. “An adaptive algorithm for mel-cepstral analysis of speech”, *Proc. of the IEEE Int. Conf.*

on Acoustics, Speech, and Signal Processing (ICASSP), pp. 137–140, mar. 1992.

- [108] TAMURA, M., MASUKO, T., TOKUDA, K., et al. “Speaker adaptation for HMM-based speech synthesis system using MLLR”, *Proc. ESCA/COCOSDA Workshop on Speech Synthesis*, pp. 273–276, nov. 1998.
- [109] HTS. “HMM-based Speech Synthesis System (HTS) - Home”. <http://hts.sp.nitech.ac.jp/>, Acessado em 2010.
- [110] SPTK. “Speech Signal Processing Toolkit (SPTK)”. <http://sptk.sourceforge.net/>, Acessado em 2010.
- [111] ENGINE. “HTS engine APP”. <http://hts-engine.sourceforge.net/>, Acessado em 2010.
- [112] OpenFST. “OpenFst Library”. <http://www.openfst.org/>, Acessado em 2010.
- [113] Active. “ActiveTcl with Snack”. <http://www.activestate.com/>, Acessado em 2010.
- [114] TEVAH, R. T. *Implementação de um Sistema de Reconhecimento de Fala Contínua com amplo vocabulário para o Português Brasileiro*. Tese de Mestrado, Programa de Engenharia Elétrica, UFRJ/COPPE, Rio de Janeiro, Brasil, jun. 2006.

Apêndice A

Publicações

A.1 Artigo em Revista

- “*A Rule-Based Method for Homograph Disambiguation in Brazilian Portuguese Text-to-Speech Systems*” - Artigo, de autoria de Denilson C. Silva, Daniela Braga e Fernando G. V. Resende Jr., submetido para publicação no “Journal of Communication and Information Systems” (JCIS).

A.2 Artigos em Conferências

- “*Conjunto de Regras para Desambiguação de Homógrafos Heterófonos no Português Brasileiro*” - Artigo, de autoria de D. C. Silva, Daniela Braga e Fernando G. V. Resende Jr., foi apresentado no XXVII Simpósio Brasileiro de Telecomunicações (SBrT2009).
- “*Separação das Sílabas e Determinação da Tonicidade no Português Brasileiro*” - Artigo, de autoria de D. C. Silva, Daniela Braga e Fernando G. V. Resende Jr., foi apresentado no XXVI Simpósio Brasileiro de Telecomunicações (SBrT2008).
- “*A rule-based grapheme-phone converter and stress determination for Brazilian Portuguese natural language processing*” - Artigo, de autoria de D. C. Silva, A. A. de Lima, R. Maia, D. Braga, J. F. de Moraes, J. A. de Moraes e Fernando G. V. Resende Jr., foi apresentado no IV International Telecommunications Symposium (ITS2006).
- “*Detecção dos extremos da voz em presença de ruído através de uma análise no domínio wavelet*” - Artigo, de autoria de D. C. Silva e Fernando G. V. Resende Jr., foi apresentado no XXII Simpósio Brasileiro de Telecomunicações (SBrT2005).

- “*Identificação da natureza do ruído baseada em HMM*” - Artigo, de autoria de D. C. Silva e Fernando G. V. Resende Jr., foi apresentado no XXI Simpósio Brasileiro de Telecomunicações (SBrT2004).
- “*Detecção de extremos robusta baseada na identificação da natureza do ruído*” - Artigo, de autoria de D. C. Silva e Fernando G. V. Resende Jr., foi apresentado na II Semana da Eletrônica da UFRJ (SE2004).
- “*Reconhecimento robusto de palavras isoladas usando multi-bandas*” - Artigo, de autoria de D. C. Silva e Fernando G. V. Resende Jr., foi apresentado na I Semana da Eletrônica da UFRJ (SE2003).

Apêndice B

Algoritmos para Treinamento do HBPTTS

Neste Apêndice, mostramos todos os detalhes relativos ao treinamento do HBPTTS-6.1.1, tanto no que diz respeito a preparação dos dados para sua execução, bem como o treinamento em si ao longo de todas as suas etapas.

B.1 Ferramentas necessárias

O treinamento do HBPTTS-6.1.1 é realizado segundo o HTS (*HMM-based Speech Synthesis System*) em [109], adaptado para o PB. Inicialmente, algumas ferramentas são necessárias na preparação do processo de treino em plataforma Linux:

- HTS de [109];
- SPTK (*Speech Signal Processing Toolkit*) de [110];
- HTS engine API de [111];
- OpenFst Library de [112];
- ActiveTcl com Snack de [113].

No processo de treinamento realizado no HBPTTS foram utilizados o HTS-2.1.1 para HTK-3.4.1 (HTK - *Hidden Markov Model Toolkit*), SPTK-3.3, HTS engine API 1.03, OpenFst 1.1 e ActiveTcl 8.4, pois esta versão possui Snack requerido para o script de extração de F0.

Todas estas ferramentas deverão estar instaladas e funcionais para não prejudicar o processo de geração dos modelos durante o treinamento.

B.2 Preparação do módulo de treinamento

Uma vez disponíveis as ferramentas necessárias para auxiliar o processo de treinamento dos modelos, dispomos de um módulo de treinamento no HBPTTS que realizará tal operação.

O módulo de treino possui a seguinte estrutura:

```
training
|
|-- Makefile
|-- config.sh
|-- configure
|-- data
|   |-- Makefile
|   |-- labels
|   |   |-- gen
|   |   |   |-- test_001.lab
|   |   |   |-- test_002.lab
|   |   |-- mono
|   |       |-- data_abo_000.lab
|   |-- questions
|   |   |-- questions_qst001.hed
|   |   |-- questions_utt_qst001.hed
|   |-- raw
|   |   |-- data_000.raw
|   |-- scripts
|   |   |-- addhtkheader.pl
|   |   |-- fst2lab.pl
|   |   |-- fst2sym.pl
|   |   |-- getf0.tcl
|   |   |-- makelabel7-4.pl
|   |   |-- window.pl
|   |-- uttinfo
|   |   |-- data_000.uttinfo
|   |-- win
|   |   |-- lf0.win1
|   |   |-- lf0.win2
|   |   |-- lf0.win3
|   |   |-- mgc.win1
|   |   |-- mgc.win2
```

```

|          '-- mgc.win3
'-- scripts
  |-- Config.pm
  '-- Training.pl

      phone      syll      stress      word      class
sil
s      si~      0      si~tezi      content
i~
t      te      1
e
z      zi      0
i
dZ     dZi      0      dZi      function
i
v      v0s     1      v0s      content
0
s
k      ko~     0      ko~      content
o~
e      e      0      emoso~j~  content
m      mo      0
o
s      so~j~   1
o~
j~
p      pa      1      para      function
a
r      ra      0
a
u      u      0      u      function
p      poX     0      poXtugejs content
o
X
t      tu      0
u
g      gejs   1
e
js
b      bra     0      brazileru content
r
a
z      zi      0
i
l      lej     1
ej
r      ru      0
u
sil
sil
sil

```

Figura B.1: Exemplo de arquivo .uttinfo.

Para que o módulo possa gerar os modelos treinados, precisamos ter os conjuntos de dados de treino nos seus respectivos locais. Os dados necessários são:

- Arquivos .raw - Deverão ser alocados na pasta raw, onde os dados de áudio estão em formato RAW com 16kHz, 16 bits e canal simples;
- Arquivos .uttinfo - Deverão ser alocados na pasta uttinfo, onde o conteúdo do arquivo está disposto conforme exemplo da Figura B.1. A organização deste arquivo é feita conforme o estabelecido em [28];
- Arquivos .lab - Deverão ser alocados na pasta labels/mono, onde um exemplo do conteúdo do arquivo está disposto a seguir. Neste arquivo temos as marcas de início e fim de cada monofone, inclusive das marcas de pausas e silêncio.

```

# text000.lab #
0          34500000    sil
34500000   35600000    s
35600000   37500000    i~
37500000   38800000    n
38800000   39600000    t
39600000   40800000    e
40800000   41400000    z
41400000   41800000    i
41800000   42500000    dZ
43000000   44400000    i
44400000   45000000    v
45000000   45800000    o
45800000   46400000    s
46400000   47500000    k
47500000   48500000    o~
48500000   49000000    e
49000000   49400000    m
49400000   50700000    o
50700000   51400000    s
51400000   52600000    o~
52600000   53400000    j~
53400000   54200000    p
54200000   54500000    a
54500000   55100000    r
55100000   55800000    a
55800000   58000000    u
58000000   58500000    p
58500000   59700000    o
59700000   60100000    X
61600000   62800000    t
62800000   63100000    u
63100000   63900000    g
63900000   64400000    e
64400000   65600000    js
65600000   66600000    b
66600000   66900000    r
66900000   68000000    a
68000000   68300000    z
68300000   70200000    i
70200000   71600000    l
71600000   72400000    ej
72400000   73000000    r
73000000   74200000    u
74200000   74500000    sil
74500000   75400000    sil
75400000   76100000    sil

```

Quando o conteúdo textual é muito extenso, podemos fazer uso do próprio módulo de processamento de texto do HBPTTS para criar os arquivos .uttinfo, fazendo manualmente as correções necessárias para o treinamento dos novos dados.

Da mesma forma, criar os arquivos .lab com as marcas de início e fim de conteúdos textuais muito extensos poderia demorar dias e até semanas. Para agilizar este processo de marcação, utilizamos um segmentador automático forçado baseado num reconhecedor de fonemas com o HTK. Este processo será descrito na Seção B.3

B.3 Segmentação automática da base de dados

Uma tarefa preponderante na construção de uma base de dados de voz é a anotação dos seus dados com o seu conteúdo fonético e o alinhamento entre eles. Estes procedimentos são extremamente desejáveis e úteis na síntese, já que estas informações são utilizadas na classificação das unidades, auxiliando a selecionar e concatenar corretamente as unidades em relação às características acústicas e linguísticas. Para reduzir tempo e esforço gastos com a segmentação manual, o procedimento aqui descrito foi adotado, visando o treinamento do HBPTTS. Na segmentação fonética automática forçada, utilizamos um método através do HTK, adaptando HMMs baseados num reconhecedor fonético com 12 coeficientes mel-cepestrais e log energia normalizada, bem como suas derivadas de primeira e segunda ordens, totalizando 39 coeficientes. Foi adotada uma taxa de frame com 10ms com janelamento de Hamming de 25ms. O conjunto de HMMs é baseado em monofones independentes do contexto, com uma gaussiana por estado. Os modelos utilizam a técnica “flat start” [114] e possuem 3 estados emitentes com transição esquerda-direita e nenhum salto. Os HMMs resultantes são reestimados em 15 interações.

```
segmentation
|-- add
|-- aligned.out
|-- bin
|   |-- codetrain.scp
|   |-- model
|   |   |-- hmmdefs
|   |   '-- tiedlist
|-- prompts
|-- '-- scripts
|   |-- config
|   |   |-- config
|   |   |-- global.ded
|   |   |-- macros
|   |   |-- mkphones0.led
|   |   |-- mkphones1.led
|   |   |-- mktri.led
|   |   |-- proto
|   |   |-- sil.hed
|   |   |-- tree1.hed
|   |   |-- vFloors
|   |   |-- wav_config
```

```

|     |-- wav_config_0dn
|     |-- wav_config_dnz0
|-- create_trainscp.pl
|-- fixfulllist.pl
|-- hmms
|     |-- aligned.mlf
|     |-- dict
|     |-- dict-tri
|     |-- dict1
|     |-- fulllist
|     |-- fulllist-original
|     |-- hmm0
|     |     |-- hmmdefs
|     |     |-- macros
|     |     |-- proto
|     |     |-- vFloors
|     |-- hmm1
|     |     |-- hmmdefs
|     |     |-- macros
|     |-- hmm10
|     |     |-- hmmdefs
|     |     |-- macros
|     |-- hmm11
|     |     |-- hmmdefs
|     |     |-- macros
|     |-- hmm12
|     |     |-- hmmdefs
|     |     |-- macros
|     |-- hmm13
|     |     |-- hmmdefs
|     |     |-- macros
|     |-- hmm14
|     |     |-- hmmdefs
|     |     |-- macros
|     |-- hmm15
|     |     |-- hmmdefs
|     |     |-- macros
|     |-- hmm2
|     |     |-- hmmdefs

```

```

|     |   |   '-- macros
|     |   |-- hmm3
|     |   |   |-- hmmdefs
|     |   |   '-- macros
|     |   |-- hmm4
|     |   |   |-- hmmdefs
|     |   |   '-- macros
|     |   |-- hmm5
|     |   |   |-- hmmdefs
|     |   |   '-- macros
|     |   |-- hmm6
|     |   |   |-- hmmdefs
|     |   |   '-- macros
|     |   |-- hmm7
|     |   |   |-- hmmdefs
|     |   |   '-- macros
|     |   |-- hmm8
|     |   |   |-- hmmdefs
|     |   |   '-- macros
|     |   |-- hmm9
|     |   |   |-- hmmdefs
|     |   |   '-- macros
|     |   |-- mfcc
|     |   |-- mktri.hed
|     |   |-- monophones0
|     |   |-- monophones1
|     |   |-- phones0.mlf
|     |   |-- phones1.mlf
|     |   |-- stats
|     |   |-- tiedlist
|     |   |-- train.scp
|     |   |-- tree.hed
|     |   |-- trees
|     |   |-- triphones1
|     |   |-- wintri.mlf
|     |   |-- wlist
|     |   '-- words.mlf
|   |-- logs
|   |-- perlsort.pl

```

```

|      '-- segmentation.sh
|-- etext2wlistmlf.pl
|-- get_alignment.pl
|-- lexicon
|  '-- lexicon.lex
|-- mono.list
|-- pre_lex.pl
|-- scripts
|  |-- maketrihed
|  |-- mkclscript.prl
|  |-- prompts2mlf
|  '-- prompts2wlist
|-- segmentation.txt
|-- splitlab.sh
|-- train
|  |-- mfcc
|  '-- wav
|-- wavlist
'-- wavlist.txt

```

Onde MFCC é o Coeficiente mel-cestral ou coeficiente cestral da mel-frequência (*Mel-frequency cepstral coefficient*).

B.3.1 Preparação dos dados de voz

Após gravados todos os dados definidos no Apêndice C, iniciamos um processo de preparação para proceder à segmentação forçada do conteúdo. O formato desejado dos arquivos após o alinhamento é semelhante ao exemplificado a seguir:

```

# text001.lab #
0          5600000    sil
5600000    9000000    a
9000000    9600000    t
9600000    10100000   r
10100000   11200000   a~
11200000   12400000   m
12400000   14300000   a
14300000   15100000   E
15100000   15900000   u
...        ...        ...

```

25700000	26300000	d
26300000	27000000	u
27000000	27500000	l
27500000	29100000	i
29100000	30500000	v
30500000	30800000	r
30800000	31800000	u
31800000	38100000	sil

Onde a primeira e segunda colunas referem-se, respectivamente, as amostras iniciais e finais dos fones descritos na terceira coluna, transcritos da frase “A trama é o ... do livro.”.

Os arquivos necessários para proceder a preparação da segmentação automática são:

1. Os arquivos de áudio (.wav) e a lista com o conteúdo textual correspondente (.txt);
2. O léxico (dicionário com transcrições das palavras que constituem o vocabulário do sistema dispostas alfabeticamente);

Ao arquivo que contem o léxico, devemos acrescentar duas novas entradas que se referem ao início (SENT-START) e fim de frase (SENT-END), conforme a seguir:

```
# lexicon.lex #
...           ...           ...
SENDO         [SENDO]        s e~ d u
SENSAÇÃO     [SENSAÇÃO]    s e~ s a s a~ w~
SENT-END      []           sil
SENT-START   []           sil
SENTE        [SENTE]      s e~ tS i
SENTEM       [SENTEM]     s e~ t e~ j~
...           ...           ...
```

As waves e o seu respectivo conteúdo já são conhecidos, tendo em vista que foram utilizados no processo de gravação da base de dados.

O léxico deverá ser criado com o maior número de palavras possíveis para posterior geração do dicionário de palavras no alinhamento das waves, ou, no mínimo, com todas as palavras existentes nas waves gravada.

Dois scripts foram criados para realizar tais operações. O primeiro, denominado **pre_lex.pl**, cujo o comando é *perl pre_lex [origem] [sr_in] [destino] [sr_out] [nwaves]*, onde:

- “origem” especifica o local onde estão gravadas a base de dados a ser segmentada;
- “sr_in” é a frequência de amostragem do áudio da base;
- “destino” é o local especificado para as waves subamostradas;
- “sr_out” é a frequência de amostragem desejada na segmentação das waves;
- “nwaves” é o número de arquivos de áudio envolvidos no processo de segmentação.

O resultado do processo será o conjunto de waves gravadas na pasta “segmentation/train/wav/”, já na formatação requerida, os arquivos **codetrain.scp**, na pasta “segmentation/bin/”, e o arquivo **wavlist**, na pasta “segmentation/”.

Esses arquivos são exemplificados a seguir e sua formatação é fundamental para o processo de automatização da segmentação:

```
# codetrain.scp #
...
../../train/wav/text010      ../../train/mfcc/text010.mfc
../../train/wav/text011      ../../train/mfcc/text011.mfc
...
...

# wavlist #
...
*/text010
*/text011
...
```

O segundo script é denominado por **get_lex.pl**, cujo o comando é *perl get_lex.pl [wavlist] [textlist]*, onde:

- “wavlist” é a lista gerada por **pre_lex.pl**;
- “textlist” é a lista com o conteúdo textual gravado, conforme visto a seguir:

```
# textlist #
...
As situações ditas embaraçosas são resolvidas com os dados.
O presidente assumiu o poder com total apoio político.
...
```

Este script `get_lex.pl` deve estar localizado na raiz do HBPTTS, pois ele utiliza o analisador de texto do sintetizador para capturar a lista de fonemas existentes no conteúdo textual.

O resultado do script é um arquivo `prompts` na pasta “segmentation/bin/”, e o arquivo `lexicon.lex`, na pasta “segmentation/lexicon/”.

Neste ponto, temos a lista de frases com o conteúdo textual gravado, a lista de waves correspondentes e o léxico construído com o maior número de palavras possíveis, bem como suas respectivas transcrições fonéticas.

B.3.2 Criação do Dicionário

Para criarmos o dicionário, necessitamos do arquivo `prompts`, com a lista de frases anteriormente gravadas, do arquivo `wlist` gerado a partir de `prompts`, onde temos uma lista variada de palavras requeridas pela base de dados. O dicionário é criado pela extração das informações fonéticas do léxico

As ferramentas do HTK são utilizadas, porém, para que seja capaz de treinar os arquivos de áudio gravados e as transcrições fonéticas, o dicionário gerado necessita ser foneticamente balanceado com aproximadamente 30 ou 40 frases de 8 a 10 palavras cada. Se o nosso conteúdo textual tiver menos do que isto, ou se não for balanceado foneticamente (se alguns fonemas somente ocorrem uma ou duas vezes), então necessitamos adicionar novas palavras para garantir que temos cerca de 3 a 5 ocorrências de cada fonema em nosso dicionário.

O arquivo `prompts` inclui nosso conteúdo textual para criar o nosso dicionário foneticamente balanceado. Este arquivo basicamente contém na primeira coluna os nomes dos arquivos de áudio gravados e os referidos locais onde estão armazenados, seguidos da lista das frases correspondentes, na segunda coluna, conforme a seguir:

```
# prompts #
...      ...
*/text010 As situações ditas embaraçosas são resolvidas com os dados.
*/text011 O presidente assumiu o poder com total apoio político.
...      ...
```

Uma lista das palavras do conteúdo textual, ordenadas alfabeticamente, também é requerida. Esta lista é tomada a partir da segunda coluna do arquivo `prompts` e organizada no arquivo `wlist`, como a seguir:

```
# wlist #
apoio
as
assumiu
```

com
 dados
 ditas
 embaraçosas
 o
 os
 poder
 político
 presidente
 resolvidas
 são
 situações
 total
 ...

A ferramenta “HDMan” cria o dicionário das palavras do conteúdo textual gravado, a partir do léxico inicialmente criado, além da lista completa de fones presentes no dicionário e das suas respectivas quantidades encontradas.

O comando utilizado é: *HDMan -A -D -T 1 -m -w wlist -e ./ -n monophones1 -i -l dict lexicon.lex*, onde “monophones1” é a lista de monofones, “dict” é o dicionário gerado e “lexicon.lex” é o léxico.

O dicionário, exemplificado logo abaixo, é semelhante ao léxico, porém conterà apenas as palavras presentes no conteúdo textual gravado.

...
LENTO	[LENTO]	l e~ t u sp
LEONTINA	[LEONTINA]	l e o~ tS i~ n a sp
LER	[LER]	l e X sp
LESIVO	[LESIVO]	l e z i v u sp
LESSA	[LESSA]	l e s a sp
LESTE	[LESTE]	l E s t i sp
...

B.3.3 Criação da transcrição fonética

Para treinar os modelos, os dados devem estar associados a sua transcrição fonética. Como não foi realizada nenhuma transcrição manual para inicializar os modelos, uma técnica denominada “flat start” é requerida.

O conteúdo é armazenado em dois arquivos MLF (Master Label File):

- **words.mlf**

- **phones.mlf**

a O script “prompts2mlf” é usado para criar **words.mlf** de **prompts**;

```
#!MLF!#
"/filewave_001.lab"
PESQUISA
É
UMA
COISA
QUE
MUDA
A
TODA
HORA
.
.
"/filewave_002.lab"
...
```

b A ferramenta “HLEd” é utilizada para criar **phones0.mlf**, que não possui pausas curtas (sp) após cada transcrição fonética das palavras. O comando *HLEd -A -D -T 1 -l '*' -d dict -i phones0.mlf mkphones0.led words.mlf* é utilizado.

c A ferramenta “HLEd” é utilizada novamente criando **phones1.mlf**, através do comando *HLEd -A -D -T 1 -l '*' -d dict -i phones1.mlf mkphones1.led words.mlf*, onde **mkphones1.led** é o arquivo que descreve sobre silêncio e pausas curtas (short-pause - sp) entre palavras.

Um trecho dos arquivos **phone0.mlf** e **phone1.mlf** são demonstrados a seguir com a transcrição fonética do conteúdo textual. Observamos que a diferença entre eles é a presença da sinalização de pausas curtas.

```
#!MLF!#
"/filewave_001.lab"
sil
p
e
s
k
i
```

```

z
a
E
...
o
r
a
sil
.
"/filewave_002.lab"
...

#!MLF!#
"/filewave_001.lab"
sil
p
e
s
k
i
z
a
sp
E
sp
...
sp
o
r
a
sp
sil
.
"/filewave_002.lab"
...

```

Análise acústica dos dados

O último estágio de preparação dos dados é parametrizar a forma de onda em seqüências de vetores. A configuração dos parâmetros para extração das características são dadas pelo arquivo “config”.

No arquivo de configuração, TARGETKIND = MFCC_0_D_A.

Para cada frame do sinal, os seguintes coeficientes são extraídos:

- O 0-ésimo coeficiente $[c_0]$, que é proporcional a energia total do frame (sufixo “_0” em TARGETKIND);

- Os 12 coeficientes MFCC [c_1, \dots, c_{12}] (NUMCEPS = 12);
- Os 13 delta coeficientes, estimados da derivada de primeira ordem de [c_0, c_1, \dots, c_{12}] (sufixo “_D” em TARGETKIND);
- Os 13 delta-delta coeficientes, estimados da derivada de segunda ordem de [c_0, c_1, \dots, c_{12}] (sufixo “_A” em TARGETKIND).

O vetor de coeficientes será composto por 39 elementos extraídos de cada frame do sinal de voz.

O comando para extração das características é *HCopy -A -D -T 1 -C config -S codetrain.scf*, onde “codetrain.scf” é um arquivo que contem os caminhos de origem e destino das características.

B.3.4 Treinamento dos modelos acústicos

Inicialização e treinamento dos modelos (HMM)

Primeiramente escolhemos *a priori* a estrutura para cada modelo:

1. Número de estados;
2. Forma das observações associadas a cada estado;
3. Disposição das transições entre estados.

```
~o <VecSize> 25 <MFCC_0_D_N_Z>
~h "proto"
<BeginHMM>
  <NumStates> 5
  <State> 2
    <Mean> 25
      0.0 0.0 0.0 ...
    <Variance> 25
      1.0 1.0 1.0 ...
  <State> 3
    <Mean> 25
      0.0 0.0 0.0 ...
    <Variance> 25
      1.0 1.0 1.0 ...
  <State> 4
    <Mean> 25
      0.0 0.0 0.0 ...
```

```

<Variance> 25
    1.0 1.0 1.0 ...
<TransP> 5
0.0 1.0 0.0 0.0 0.0
0.0 0.6 0.4 0.0 0.0
0.0 0.0 0.6 0.4 0.0
0.0 0.0 0.0 0.7 0.3
0.0 0.0 0.0 0.0 0.0
<EndHMM>

```

<NumStates> 5: Define o numero total de estados, incluindo os 2 estados não-emissivos (1 e 5);

<State> 2: Introduce a descrição das observações do estado 2. Aqui escolhemos usar observações com gaussianas simples e matriz diagonal, já que a função é completamente descrita pela média e variância (Os elementos da diagonal da matriz autocorrelação). Os estados 1 e 5 não são descritos, visto que não possuem nenhuma observação.

<Mean> 25: 0.0 0.0 (...) 0.0 (x 25) dá o vetor de médias (em um espaço 25-dimensional) da observação corrente. Todos os elementos são arbitrariamente inicializados por 0. O arquivo apenas fornece o “prototipo” do HMM (sua estrutura global). Estes coeficientes serão treinados mais tarde.

<Variance> 25: 1.0 1.0 (...) 1.0 (x 25) dá o vetor de variâncias da observação corrente. Todos os elementos são arbitrariamente inicializados por 1.

<TransP> 5: Dá a matriz de transições 5x5 do HMM. Onde a_{ij} é a probabilidade de transição do estado i para o estado j . Valores nulos indicam que as correspondentes transições não são permitidas. Os outros valores são arbitrariamente inicializados (porém cada linha da matriz deve somar 1). Elas serão modificadas posteriormente, durante o treinamento.

A inicialização dos modelos é feita utilizando-se a técnica “Flat Start”, para a segmentação automática da base de dados, onde três etapas básicas são realizadas:

- as pausas existentes entre palavras são desconsideradas, para gerar uma primeira estimativa dos modelos dos fonemas;
- em seguida, é criado um modelo de pausa baseado no modelo de silêncio de início e fim de frase e reestimam-se os modelos;
- executa-se um realinhamento dos dados em função dos modelos e estes são novamente estimados.

A ferramenta “HCompV” fará uma varredura nos arquivos de dados, calcula a média e variância global e define todas as gaussianas em um determinado modelo

com a mesma média e variância. Considerando que a lista de todos os dados de treino estão armazenados em `train.scp`, o comando utilizado é : `HCompV -A -D -T 1 -C config -f 0.01 -m -S train.scp -M hmm0 proto`, onde `train.scp` é a lista de arquivos MFCC que estão participando do treinamento. `hmm0` é o nome do diretório de saída. `proto` é o arquivo com a definição básica do HMM. Assim teremos na saída os arquivos `macros`, `vFloors` e `hmmdefs`. Com esta operação do `proto` fornecido e `train.scp`, é gerado o `hmm0` com um novo `proto`.

```

macros

-o
  <VecSize> 25
  <MFCC_0_D_A>
-v "varFloor1"
  <Variance> 25
0.0012 0.0003 ...

hmmdefs

-h "aa"
  <BeginHMM> ...
  <EndHMM>
-h "eh"
  <BeginHMM> ...
  <EndHMM>
... etc

```

Os monofones armazenados no diretório `hmm0` são reestimados através da ferramenta “HERest” através do comando: `HERest -A -D -T 1 -C config -I phones0.mlf -t 250.0 150.0 1000.0 -S train.scp -H hmm0/macros -H hmm0/hmmdefs -M hmm1 monophones0`. A reestimação é utilizada para gerar o `hmm1`. Novamente a ferramenta HERest é utilizada por duas vezes seguidas, para uma reestimação, onde é criado o arquivo `hmm2`, na primeira, e `hmm3`, na segunda.

Cada vez que é executado o comando “HERest”, ele realiza uma simples reestimação dos modelos anteriormente gerados. Cada novo conjunto de modelos é armazenado em um novo diretório. A execução de “HERest” deverá ser repetida tantas vezes quanto necessário, alterando o nome dos diretórios de entrada e saída a cada interação, até o diretório `hmm10` (se o número de interações = 10) que contem o conjunto final dos modelos dos monofones inicializados.

Realinhamento dos dados

Os modelos de fones criados são utilizados para realinhar os dados treinados e criar novas transcrições. Isto é feito através da ferramenta “HVite” através do co-

mando: *HVite -A -D -T 1 -l '*' -o SWT -b silence -C config -H hmms/hmm7/macros -H hmms/hmm7/hmmdefs -i hmms/aligned.mlf -m -t 250.0 150.0 1000.0 -y lab -a -I hmms/words.mlf -S hmms/train.scp hmms/dict1 hmms/monophones1*. O comando aplicado usa HMMs armazenados na última interação de reestimação, para transformar a transcrição, a nível de palavras, presente em **words.mlf** para o nível de fone, em “aligned.mlf”. Um arquivo “aligned.out” é gerado com todas as palavras já segmentada. Num passo seguinte, todas as frases transcritas são retiradas deste arquivo e separadas para o treino.

```
# text000.lab #
0          34500000      sil
34500000   35600000      s
35600000   37500000      i~
37500000   38800000      n
38800000   39600000      t
39600000   40800000      e
40800000   41400000      z
41400000   41800000      i
41800000   42500000      dZ
43000000   44400000      i
44400000   45000000      v
45000000   45800000      o
45800000   46400000      s
46400000   47500000      k
47500000   48500000      o~
48500000   49000000      e
49000000   49400000      m
49400000   50700000      o
50700000   51400000      s
51400000   52600000      o~
52600000   53400000      j~
53400000   54200000      p
54200000   54500000      a
54500000   55100000      r
55100000   55800000      a
55800000   58000000      u
58000000   58500000      p
58500000   59700000      o
59700000   60100000      X
61600000   62800000      t
62800000   63100000      u
63100000   63900000      g
63900000   64400000      e
64400000   65600000      js
65600000   66600000      b
66600000   66900000      r
66900000   68000000      a
68000000   68300000      z
68300000   70200000      i
70200000   71600000      l
71600000   72400000      ej
72400000   73000000      r
73000000   74200000      u
74200000   74500000      sil
# text001.lab #
...
```

Apêndice C

Uma nova base de dados de voz para reconhecimento e síntese.

C.1 1000 frases foneticamente balanceadas.

Conjunto de 1000 frases gravadas com fala neutra:

- 1) Pesquisa é uma coisa que muda a toda hora.
- 2) No total, serão chamados vinte e seis mil candidatos.
- 3) O número de convocados por vaga é de doze candidatos.
- 4) Atualmente, esse abatimento é limitado a setenta por cento dos gastos.
- 5) Sandra Regina Machado: acho que ela enfim criou juízo.
- 6) Eles estão colocando armadilhas nas fazendas onde já ocorreram os ataques.
- 7) Dessas, somente umas trezentas e vinte foram inauguradas em território americano.
- 8) No total, sete mísseis foram disparados contra o encrave.
- 9) Em Florianópolis, foi registrado dois graus Celsius na manhã de domingo.
- 10) As situações ditas embaraçosas são resolvidas com os dados.
- 11) Itamar tem razão de estar exultante como nunca desde que virou presidente.
- 12) A mãe de todas as reformas é a reforma política.
- 13) Conseguiram eliminar áreas supérfluas ou que antes eram desperdiçadas.
- 14) Uma lata de leite em pó integral vale um ingresso.
- 15) A maioria dos passageiros do barco naufragado era de crianças.
- 16) A provável causa do acidente foi excesso de lotação a bordo.
- 17) Não prometo nada, porque não adianta eu prometer e não cumprir.
- 18) Se for eleito, vocês vão ver o meu trabalho.
- 19) Ele era um dos poucos atores negros que tinham espaço.
- 20) A secretaria estadual de saúde distribuirá cem mil preservativos no carnaval.

- 21) São essas qualidades que inspiraram o plano real desde a sua criação.
- 22) Todos os batizados são consagrados a Deus e devem tender à santidade.
- 23) O homem Sem Qualidades, não é um livro comum.
- 24) Os problemas surgem nas importações diretas via catálogos, por exemplo.
- 25) Lula chegou ao Rio com duas horas de atraso.
- 26) No ano, a taxa é de três vírgula sete por cento.
- 27) Batizado de Heitor, o trabalho traz treze faixas compostas pelo guitarrista.
- 28) Já Viola levou o terceiro cartão amarelo domingo passado.
- 29) Inflação volta a subir no Rio e em São Paulo.
- 30) CINCO: freqüentemente responde perguntas antes que elas sejam concluídas.
- 31) Quero estar no grupo e ajudar sempre que necessário.
- 32) Expedito Andreino de Souza, cinqüenta e um anos, encarregado de almoxarifado
- 33) Isso mostra que nossa flutuação cambial suja e voluntariosa foi apenas tolice.
- 34) Nós somos uma espécie de contingente escolhido do povo brasileiro.
- 35) Não é o primeiro carregamento de armas para Angola.
- 36) O dinheiro para pagamento dos policiais militares sairia de verba da prefeitura.
- 37) Até agora, na televisão, eles já somam quatro ministros.
- 38) Pode ser o início de uma campanha de contestação da vitória.
- 39) Da mesma forma, empataram nas finalizações: um a um.
- 40) O jogador do Palmeiras foi o único que lançou três vezes.
- 41) Quanto ao seu próprio time o diagnóstico é menos generoso.
- 42) Não há desculpas para o mau jogo que fizemos contra a Arábia.
- 43) Salário real médio do varejo aumenta doze por cento no mês passado
- 44) Em Tóquio, o dólar fechou cotado a noventa e oito ienes.
- 45) Os sobreviventes cremaram seus mortos e procuram reconstruir suas vidas.
- 46) Os dois times perderam uma chance atrás da outra.
- 47) A geração de Zico não seria campeã do mundo.
- 48) O curso é gratuito e tem a duração de seis meses.
- 49) Assaltava desde os oito anos e meus amigos estão quase todos presos.
- 50) O mercado de ações vive momentos de plena euforia.
- 51) A imprensa não se incomodou, nem denunciou esse detalhe.
- 52) Primeiro arredondou a idade do percussionista para cinqüenta anos.
- 53) Os médicos ainda não sabem quanto tempo ele deverá ficar no hospital.
- 54) O feminino fica em quarto, sua melhor colocação em olimpíadas.
- 55) O modelo foi reestilizado e ganhou linhas mais agressivas.
- 56) Seu interior, no entanto, lembra o de automóveis japoneses.
- 57) Sete membros do grupo que se opõe à paz continuam presos.
- 58) Os jogadores reclamam do calor, que consideram excessivo dentro do estádio.

- 59) O limite anterior para estes títulos bancários era de noventa dias.
- 60) Também preparam a transição para o real no caso das aplicações financeiras.
- 61) Há algumas coisas que não podem deixar de serem vistas em Paris.
- 62) Há demanda por real, não por dólar, cuja cotação cai.
- 63) Há cento e setenta bilhões de moedas de um centavo em circulação.
- 64) Cada uma delas custa oito décimos de um centavo.
- 65) O mercado fica de alto risco a curto prazo.
- 66) A perspectiva continua otimista para o médio e longo prazos.
- 67) O desafio agora é controlar, com medicamentos, essa secreção.
- 68) O jogo contra a Suécia deixou todo o país preocupado.
- 69) Está instalado na casa do avô de Lúcia Flexa de Lima.
- 70) Funcionários do governo do estado dão orientação técnica aos voluntários.
- 71) O ministério diz que não foi consultado sobre a transação.
- 72) Este ano a empresa deve perder cerca de sete milhões de dólares.
- 73) Ela começou a tirar fotos ontem em Buenos Aires.
- 74) Ele se comprometeu a ler e responder parte da minha correspondência.
- 75) Os brasileiros no exterior ganharam um poder aquisitivo sem precedentes.
- 76) O objetivo é abordar todas as fases da ópera.
- 77) O curso é diário e tem duração de três meses.
- 78) Os jogadores santistas deixaram o gramado sob vaias da torcida.
- 79) A torcida voltou a criticar a atual administração e não economizou críticas.
- 80) A Infraero ainda não sabe em qual dos dois aeroportos Alexandre embarcou.
- 81) Eles têm interesse na redução das taxas nos empréstimos rurais.
- 82) Ele podia dar explicações práticas para sua preferência por faroestes.
- 83) Longe dos estúdios, também estava longe do controle dos produtores.
- 84) Conversões pelo dólar turismo a mil seiscentos e oitenta reais.
- 85) O preço do primeiro estágio é de duzentos reais.
- 86) Tentou usar a cruz mas era curta para isto.
- 87) Temos de melhorar a qualidade do ensino, especialmente na escola pública.
- 88) Tenho muita vontade de aprender, mas baixo poder aquisitivo.
- 89) Taborda pegou as assinaturas na manhã do dia vinte e cinco.
- 90) Também não são recomendáveis os temas de acidentes de automóveis ou alcoolismo.
- 91) Para discutir sobre ecologia e projetos para gerações futuras.
- 92) Um velho trem circula entre os bairros de Capivari e Abernésia.
- 93) Gilda Portugal Gouvêa será delegada do Ministério da Educação em São Paulo.
- 94) A integração é crucial quando se pensa no mercado externo.
- 95) Programa criado com a ajuda de Luxemburgo chega em novembro
- 96) Dinamizar tais mercados significa coordenar toda a cadeia da carne.

- 97) O resultado é menor do que a soma das aptidões.
- 98) O sistema brasileiro se inspirou principalmente no modelo utilizado na França.
- 99) Já as cadernetas têm apenas hoje como um bom dia para aplicar.
- 100) Viva o discreto presidente do Palmeiras, Alberto Helena Junior.
- 101) Nada pior que a aplicação distorcida de uma idéia.
- 102) Rua Padre João Manuel, número cento e cinqüenta e quatro.
- 103) A cópia distribuída demonstra que as correções foram feitas às pressas.
- 104) Os dados são coletados nas juntas comerciais dos estados.
- 105) É muito interessante ver as mudanças do estilo das mulheres.
- 106) Ele está com o relator, o juiz Célio Benevides.
- 107) As importações de maçã argentina não preocupam os plantadores.
- 108) Os laboratórios questionam os cálculos mas não negam aumentos reais.
- 109) Apesar disso, continua ligado ao seu país de origem.
- 110) O registro do hospital dizia que ele havia fugido.
- 111) Várias agências de São Paulo oferecem cursos de língua no Japão.
- 112) Juro para capital de giro chega a cento e quarenta por cento ao ano.
- 113) O Tribune sugere que o acordo foi lesivo para o governo.
- 114) Não posso me imaginar trabalhando em algo sempre igual.
- 115) O Mercosul é um compromisso irreversível do governo Brasileiro.
- 116) A mamata da Alfândega foi cinco vezes mais cara.
- 117) Chega a secretário da cultura com amplo apoio de intelectuais e artistas.
- 118) Petista não pode ver orelha de eleitor que dá o maior aluguel.
- 119) Atenção: a partir de sexta estarei no “Caderno Copa noventa e quatro”.
- 120) Pelo menos dezenove pessoas ficaram feridas nos dois ataques.
- 121) Os manuais serão vendidos do dia dez a vinte de outubro.
- 122) A legislação eleitoral em vigor nesses anos vetava contribuições de pessoas jurídicas.
- 123) O principal segredo para fazer uma boa salada é caprichar no visual.
- 124) No inverno, esta variação é menor, mas ainda existe.
- 125) Silêncio custa nove mil dólares ao técnico da Romênia.
- 126) Fiz o que a maioria procurou fazer: evitei provocar marola.
- 127) Uma versão da história de amor de Tristão e Isolda.
- 128) O livro é ilustrado por Zélia Suassuna, mulher do autor.
- 129) A maior alta foi registrada em Salvador, de dezoito por cento.
- 130) Com a anistia, voltou a prática política da década de cinqüenta.
- 131) Cristiano tem um excepcional trabalho de luz e enquadramento.
- 132) O quorum qualificado dificultaria mudanças circunstanciais nos futuros governos.
- 133) A contínua exposição à poluição intensa favorece o envelhecimento precoce da

pele.

- 134) O Brasil está representado por uma imensa delegação de duas pessoas.
- 135) Inclui mais de noventa minutos de vídeo, fotos e ilustrações.
- 136) Nos próximos dias, devem ser enviadas as outras notificações.
- 137) Evite correr ao banco para trocar o seu dinheiro.
- 138) O cruzeiro real continua valendo até o dia quinze de julho.
- 139) Há quem garanta que um craque não desaprende a arte de jogar.
- 140) O autor argumenta que a medida fere a legislação sanitária.
- 141) A rua Florida sempre foi a grande passarela dos turistas.
- 142) Isso sem contar os negócios gerados pelo licenciamento de produtos ou franquias.
- 143) Durante a revista eles mandaram que eu me comportasse como homem.
- 144) Sou uma pessoa normal e me comporto do jeito que quiser.
- 145) É só sair nas ruas que os soldados chegam junto.
- 146) Não deixe de entrar na Galeria Pacífico, recém restaurada e reformada.
- 147) É preciso reformular o parlamento para que ele possa atender a sociedade.
- 148) A ocorrência foi registrada como homicídio culposo não intencional.
- 149) Na avaliação pelo desempenho individual, cada executivo tem cinco metas anuais.
- 150) Na recepção, às vezes falhamos em momentos importantes do jogo.
- 151) Portuguesa pode terminar, hoje, jogo inacabado com o Santo André
- 152) O populismo em ação é um dos grandes espetáculos da política.
- 153) É a oportunidade para o mercado observar os talentos mais promissores.
- 154) Os maridos não têm reclamações a fazer um do outro.
- 155) Carlos afirma que faz de conta que não vê Eduardo e Oscar.
- 156) Eduardo diz que não reclama porque se sente o mais querido.
- 157) Vamos esquecer o passado, vamos pensar no presente e no futuro.
- 158) O América está em quinto lugar, com vinte e seis pontos.
- 159) O próprio fundo de ações retém o imposto eventualmente devido no resgate.
- 160) A classe média paulistana vai às lojas neste Natal com mais apetite.
- 161) Em grande fase, é a nova sensação do time.
- 162) Os jornais têm aberto enorme espaço aos vários tipos de roubalheira.
- 163) Grandes empresas começam a investir maciçamente em seus jovens
- 164) No entanto, medidas de segurança muito mais rigorosas deverão ser adotadas.
- 165) Algumas oficinas de Parceiros da Mata Atlântica estão sendo chamadas de estações.
- 166) Para participar dessas atividades, é preciso se inscrever.
- 167) Costuma-se dizer que não se pode discutir com a vitória.
- 168) Faltam exatas duas semanas para que comece de fato a corrida eleitoral.

- 169) Parece o Covas beijando o Covas ou a Erundina beijando a Erundina.
- 170) Diz que ele levanta a Bíblia o tempo todo na televisão.
- 171) O time de Jair Pereira não conseguia criar oportunidades de gol.
- 172) Respondeu que a sua frase havia sido tirada do contexto.
- 173) Sou doce de coração, a paz é que me importa.
- 174) Na Universidade de Paris Três é só isso que se faz.
- 175) É porque falo do que é real, enquanto os outros...
- 176) A Coréia exporta um bilhão de dólares em móveis por ano.
- 177) Ainda não apareceu um desgraçado aqui para montar um pólo moveleiro.
- 178) O governo do Rio não terá representante no encontro.
- 179) O conceito ainda soa meio estranho para o brasileiro.
- 180) A polícia apreendeu nos últimos vinte dias sessenta carros a seu serviço.
- 181) Nuno Cobra está em fase de revisão do texto.
- 182) Seu livro deve sair em outubro ou novembro deste ano.
- 183) O bebê estava em posição pélvica invertida no interior do útero.
- 184) Problemas encontrados nos diversos subprocessos organizacionais possuem causas próprias.
- 185) Ao eliminarmos as causas de um problema, automaticamente eliminamos o problema.
- 186) Dizem que a disciplina é importante e se embriagam às escondidas.
- 187) Fotógrafo da revista Caras quase é linchado por querer fotografar os feridos.
- 188) As três já pensam em voltar a Londres no ano que vem.
- 189) A maioria dos times paulistas está bem na competição.
- 190) Não acho que exista uma resposta única para os problemas das mulheres.
- 191) Nenhum estava muito entusiasmado e um deles foi claro.
- 192) Há vários restaurantes que oferecem feijoada, carnes e até comida baiana.
- 193) É um filme com um impacto vigoroso sobre a vida nacional.
- 194) O abatimento vai baixando gradativamente até o final do mês.
- 195) A vigilância iniciou um trabalho de combate à doença em dois bairros.
- 196) Os detentos serão transferidos das delegacias que estiverem lotadas.
- 197) As negociações com a Áustria continuavam na noite de ontem.
- 198) Esta tese é motivo de preocupação constante entre os militares.
- 199) A festa foi financiada pela indústria de cigarros Cibrasa.
- 200) Durante a festa, a empresa lançou a marca de cigarros Tetra.
- 201) Dos duzentos imóveis novos oferecidos no mês, apenas quinze foram comercializados.
- 202) Resta quase um mês para aprovar e sancionar a nova lei.
- 203) Biscaia diz que gravações podem ligar bicheiros ao tráfico de drogas
- 204) Segundo Biscaia, esses negócios não são relacionados ao jogo do bicho.

- 205) Apenas dez por cento dos funcionários em todo Brasil trabalharam ontem.
- 206) Hoje, é a autora viva mais estudada nas universidades francesas.
- 207) Segundo ele, foi apenas um almoço rápido entre amigos.
- 208) Sua densidade teatral ficou prejudicada pelos freqüentes intervalos comerciais.
- 209) Exibida em duas ou três partes, renderia muito mais.
- 210) Estes últimos só podem ser resolvidos com mudanças na constituição.
- 211) O envolvimento dos empregados é fundamental para conquistar clientes.
- 212) Patrocínio rende à cê bê éfi vinte e três vírgula seis milhões de dólares.
- 213) Até mesmo na Alemanha a maior parte dos jornais foi muito positiva.
- 214) Eu acabo de receber uma crítica muito elogiosa aqui no Canadá.
- 215) É muito raro que algum documentário esteja na lista deles.
- 216) No passeio, Ciro também pôde ver a impopularidade do colega argentino.
- 217) O Jornal Nacional entrou na história do partido que recebeu dinheiro sujo.
- 218) Augusto de Toledo tem oitenta e três anos de vida.
- 219) Nesta partida, o lateral da Seleção Brasileira Jorginho contundiu o pé.
- 220) Abandonou a França e se fixou em São Paulo.
- 221) Então chegou a hora desta gente bronzeada mostrar seu valor.
- 222) Por mais que reze, só aparece assombração na minha frente.
- 223) Concluíram que é melhor aprofundar o processo do Mercosul.
- 224) Com isso, o preço passou para vinte e cinco reais.
- 225) Itamar abriu a reunião com uma pergunta de fato pertinente.
- 226) A Criolita já está em vias de projeto de mineração.
- 227) No primeiro e terceiro, recusou as ofertas feitas pelos bancos.
- 228) O carro rodopiou na pista molhada e bateu no caminhão.
- 229) Cada um pagou multa de setenta reais por porte de maconha.
- 230) Gabriel, o pensador, é um dos nomes confirmados nesta lista.
- 231) Só sete por cento na pesquisa se disseram satisfeitos com suas explicações.
- 232) DOIS: o piloto teve morte cerebral instantânea, no momento do acidente.
- 233) Seja como for, Oliveira discorda de João Manuel num ponto.
- 234) Ela disse que ainda decidirá seu voto para a câmara dos deputados.
- 235) O código tributário só permite a indexação dos impostos pela inflação passada.
- 236) A Latasa vinha solicitando ao governo a adoção da medida.
- 237) Quem precisar abastecer a geladeira hoje não vai ter tantas dificuldades.
- 238) Já os supermercados e mercados municipais permanecerão fechados nesta terça-feira.
- 239) Ele sonha protagonizar a final da Copa contra a seleção espanhola.
- 240) Ele atribuiu o empate dos espanhóis à má sorte.
- 241) Por enquanto, os avanços concentram-se na área do reimplante.
- 242) Existem péssimas edições do livro, por isso, use o da Edição Pensamento.

- 243) A unidade taximétrica passa a custar oitocentos e setenta cruzeiros reais.
- 244) Já no dia sete, caiu para quarenta e três por cento.
- 245) Agora, o pobre pode ficar com o par completo.
- 246) Esta espécie de ruído desaparece após uma noite de sono.
- 247) Vibrações do ar balançam a membrana, fazendo-a trepidar.
- 248) Estes ossos captam as vibrações da membrana e balançam.
- 249) O trem partiu com setenta e seis passageiros, dezenove em cada vagão.
- 250) Ela tinha de saber que pobre não foi feito para comer biscoito.
- 251) Segundo ele, a polícia não iria ceder a exigências.
- 252) Dos vinte e um prêmios, quatorze foram para trabalhos jornalísticos.
- 253) O treino de hoje será das treze às dezessete horas em Brasília.
- 254) Só exercitar o quadríceps pode provocar desequilíbrio e forçar mais o joelho.
- 255) Por isso, é importante fortalecer os músculos de trás da coxa.
- 256) O travesseiro aromático relaxa, combate a insônia e atua como sedativo.
- 257) O técnico Cláudio Duarte fez mudanças no time gaúcho.
- 258) O que conta agora é o que o candidato tem a dizer.
- 259) Não bastará que Lula exiba os dentes diante das câmeras.
- 260) Agora, ele despenca do teto em direção ao piso.
- 261) Esse tempo seria reduzido à metade com as duas urnas.
- 262) O novo presidente seria conhecido quatro dias após a votação.
- 263) Ele disse que o medo do congelamento ainda atrapalha o plano.
- 264) Fabrício disse que nunca administrou mas também nunca roubou um sorvete.
- 265) “Guimarães Rosa: o alquimista do coração”, de José Maria Martins.
- 266) Análise de viés psicológico da obra de Guimarães Rosa.
- 267) Os sérvios sempre defenderam a divisão, que consolidaria seus ganhos militares.
- 268) Segundo Almeida, é praticamente impossível controlar o fogo sem ajuda exterior.
- 269) A imagem do terceiro decanato de Sagitário é o titã.
- 270) Eu queria fazer minhas próprias coisas e logo caí na real.
- 271) Sem texto é impossível interpretar com o mínimo de credibilidade.
- 272) Os libretos se tornam meus à medida que os internalizo.
- 273) Hoje o índio só responde à constituição se solicitar sua emancipação.
- 274) Daria na ocasião para quitar quase metade do débito.
- 275) Já dá para comer chocolate sem peso na consciência.
- 276) O valor venal dos imóveis vai ser ajustado para a realidade.
- 277) Os países mais avançados vêm investindo pesadamente em tecnologias poupadoras de água.
- 278) O mercado financeiro só remunera o dinheiro nos dias úteis.

- 279) O Banco Garantia é o mais agressivo participante do mercado financeiro hoje.
- 280) Veloz nas decisões, impiedoso com os adversários, joga para matar.
- 281) As escolas sozinhas não conseguem formar um bom profissional.
- 282) Outros produtos também lembrados: feijão, café e derivados de soja.
- 283) Dos dez primeiros, apenas a Noruega foi eliminada da Copa.
- 284) Os investidores foram para a Argentina e para a Venezuela.
- 285) A própria demora no julgamento do processo já incomoda.
- 286) O seqüestro de bens serviria para garantir o pagamento.
- 287) Feministas, negros e homossexuais fazem reivindicações retrógradas e tendem a exigir privilégios.
- 288) A justiça brasileira considerou o fato apenas um crime comum.
- 289) A decisão está para ser tomada em fevereiro ou março.
- 290) Uma ação de molestamento sexual pode levar anos para se concluir.
- 291) A bola sai com a equipe que tenta fazer o gol.
- 292) A cada sessão, Parreira se preocupa especificamente com um setor.
- 293) A escolha tentou dar o máximo de representatividade aos escolhidos.
- 294) Xavier descarta a hipótese de o caixão ter sido violado.
- 295) Uma parte importante da nossa vida acaba quando morre alguém querido.
- 296) Reúne ainda depoimentos de profissionais que já montaram seu próprio negócio.
- 297) Dinheiro apresentado pelos meios de comunicação como financiamento eleitoral.
- 298) As mulheres podem ser lindas, perfumadas e adoráveis mesmo.
- 299) Este parece ser o lema dos franceses, ou dos filmes franceses.
- 300) Tudo parecia assegurado diante de visíveis sinais de agonia alheia.
- 301) Pode-se optar por comprar cada programa em separado.
- 302) Isso de que o importante é competir é balela.
- 303) Posso tirar mais uma semana para trabalhar o time principal.
- 304) Em setenta e sete todos eles tinham lançado discos.
- 305) O Brasil tem pressa e o governo está pronto para agir.
- 306) Nesse instante, ouvi meu filho sufocando em seu próprio sangue.
- 307) Para ele, mesmo a carne está sem perspectiva de alta.
- 308) O grupo chegou quando os amotinados já tinham saído.
- 309) Tudo será transparente e gradual, fruto de conciliações e consensos.
- 310) Ele descarta armazéns a céu aberto, como ocorreu em oitenta e nove.
- 311) É muito importante divulgar a idéia da qualidade para o público.
- 312) É uma questão de cidadania, que não se restringe a produtos.
- 313) A melhoria da qualidade não interessa só a técnicos ou empresários.
- 314) Nos países desenvolvidos, os grandes jornais publicam cadernos sobre o tema.

- 315) Estou muito cheio de pepinos para ficar tratando dos pepinos dos outros.
- 316) As pessoas se divertiram tanto que aplaudiram de pé.
- 317) É o primeiro contato do grupo com uma produção mais apurada.
- 318) Na Idade Média, os mapas voltam-se para o leste.
- 319) A programação da Arte exclui seriados, esportes e programas de auditório.
- 320) É a maior campanha da empresa realizada até hoje.
- 321) O problema é que a inflação demora um pouco para cair.
- 322) Os pratos são à base de peixe e frutos do mar.
- 323) A Papel Presente aceita encomendas de embalagens até dia vinte.
- 324) É muito emocionante os homens se encontrarem à noite no bosque.
- 325) No final era um monte de homens se abraçando e chorando.
- 326) A pesquisa do Datafolha é um levantamento por amostragem estratificada .
- 327) A margem de erro é de três pontos percentuais.
- 328) Pelo acordo, Israel mantém responsabilidade pela defesa e política externa.
- 329) O importante é que tenho entusiasmo para fazer as coisas.
- 330) Às vezes você pode ter um cargo e não contribuir coisíssima nenhuma.
- 331) A pia batismal de mármore tem dois metros de diâmetro.
- 332) Esse povo tem leis e segredos que só quem é cigano sabe.
- 333) Visitados todos, na ocasião, por delegações de grandes empresários.
- 334) Ontem, mais seis exilados voltaram à faixa de Gaza.
- 335) A paciente, de vinte e dois anos, passa bem.
- 336) O dinheiro é arrecadado pelos templos através do dízimo.
- 337) Tem-se uma receita mensal de trezentos e quarenta mil dólares.
- 338) Juliana disse que o banco a ser escolhido é surpresa.
- 339) Este mês os termômetros descem facilmente abaixo de zero.
- 340) Vale a pena também fazer uma visita ao museu arqueológico.
- 341) Dos oitenta e um senadores, apenas cinqüenta e seis votaram.
- 342) A maternidade agora está capacitada para realizar cento e cinqüenta partos.
- 343) Crescem as baixas civis na república separatista, de população majoritariamente muçulmana.
- 344) Esta época de criatividade caiu um pouco no final da década.
- 345) O artilheiro Viola recebeu na intermediária e avançou para o gol.
- 346) O atacante comemorou o gol imitando um diabo, símbolo do América.
- 347) Uma regra fundamental é definir um limite máximo de consumo.
- 348) Ela é a pessoa mais mercenária que já conheci.
- 349) O time cometeu vinte e seis faltas durante a partida.
- 350) Foi o que afirmaram sessenta e dois por cento dos oitocentos entrevistados.
- 351) Seu desejo é de que apareça logo um doador.
- 352) Agora, continua sendo necessário que a Assembléia Legislativa aprove o nome.

- 353) Depois, foi a vez do baiano retribuir as homenagens.
- 354) Se conseguirmos uma parte do que ele alcançou, será demais.
- 355) Basta uma vitória para o time se classificar para as semifinais.
- 356) Os críticos diziam que bastava um dos dois para cumprir esse papel.
- 357) O Brasil perdeu e puseram nele a culpa pela derrota.
- 358) Este lento processo de recuperação evidencia o cuidado com o consumidor.
- 359) Outro exemplo é a beterraba, uma boa alternativa para a cenoura.
- 360) Quem vencer estará classificado para as oitavas de final.
- 361) Aragão acabou inocentado devido ao baixo número de senadores presentes.
- 362) Eram necessários quarenta e um votos para Aragão perder o mandato.
- 363) Não estou pensando em voltar para o São Paulo.
- 364) O parque é considerado uma ilha de tranqüilidade em plena Avenida Paulista.
- 365) Esta é a opinião de trinta e dois por cento do público.
- 366) Ele, corintiano, afirmou que às seis da tarde iria buscar a namorada.
- 367) Às vezes, numa crise ou numa derrota se aprende mais.
- 368) As vendas por catálogo já atingem cinqüenta e oito bilhões de dólares.
- 369) As estrelas internacionais do violão chegam esta semana a São Paulo.
- 370) Pessoas mais velhas têm pele menos elástica que os jovens.
- 371) Eu sou fiel aos meus amigos e eles são fiéis a mim.
- 372) Os juízes poderão argumentar que na música também é assim.
- 373) O jogo foi marcado pela disputa no meio campo.
- 374) No Grêmio, o meia Carlos Miguel centralizava o jogo.
- 375) O objetivo da ação foi protestar contra o corte ilegal de árvores.
- 376) Todos os cheques abaixo de cem dólares são desprezados.
- 377) Tá no ar e nas telas mais uma calúnia do macaco Simão.
- 378) Acertou o alvo em quarenta e três por cento das suas chances.
- 379) Montadas, as partes do rosto funcionam como portas de entrada.
- 380) Ao clicar sobre os olhos, você entra nas músicas e vídeos.
- 381) Quem sair por último feche a porta e apague a luz.
- 382) A equipe de Americana ocupa a oitava posição.
- 383) A organização atribuiu os dados ao êxito das campanhas de vacinação.
- 384) Botafogo e Guarani iniciaram a segunda fase com performances semelhantes.
- 385) A concentração do bolo de recursos em um pequeno grupo aumentou.
- 386) Foi definido o campeonato estadual, que começa no dia trinta.
- 387) O campeonato terá dois turnos, nos quais jogam todos contra todos.
- 388) Classificam-se dois times de cada chave a um quadrangular final.
- 389) O vencedor de cada grupo ganha um ponto extra.
- 390) O espaço fica, portanto, vago e disponível para o traficante.
- 391) Os dados são da Associação Nacional dos Fabricantes de Veículos Automotores.

- 392) Seus computadores processam até dois milhões de informações por segundo.
- 393) Casagrande substitui Viola e o meia Souza fica no banco.
- 394) A banda tocou nove músicas enquanto passou pela Consolação.
- 395) Durante todo o ano passado, os três tiveram apenas uma falta.
- 396) Para noventa minutos houve uma revolução debaixo das traves.
- 397) O mais desculpável em Itamar são os seus vexames de passarela.
- 398) Para Branco, este é um jogo como outro qualquer.
- 399) Até agosto, o valor era de sessenta e quatro reais.
- 400) Quanto mais avançarmos nas reformas, menor será esse ônus.
- 401) Começa hoje em Ilhéus o terceiro seminário nacional de turismo ecológico.
- 402) Eu fui a primeira pessoa a abrir espaço para você.
- 403) A última chuva no Mato Grosso ocorreu em junho.
- 404) A nova diretoria tomará posse em primeiro de Fevereiro.
- 405) O estado registrou redução nos índices de mortalidade infantil.
- 406) O plano terá recursos no valor de cento e dezoito mil.
- 407) A administração é feita pela Federação de Obras Sociais.
- 408) Os produtos vendidos nestas estantes foram doados por empresas particulares.
- 409) O Nacional vai ser cauteloso ao enfrentar o São Paulo hoje.
- 410) Dieta para atletas deve ser rica e de fácil digestão
- 411) O estudo é interessante e sofisticado, mas suas conclusões são óbvias.
- 412) As propostas de política econômica são poucas e imediatas.
- 413) Eles experimentaram vários tipos de comidas exóticas durante a viagem.
- 414) Estava muito calor e não soubemos dosar a energia.
- 415) Nestor Pestana, duzentos e sessenta e seis, Região Central.
- 416) Glória Funaro já colocou sua coleção de outono-inverno nas prateleiras.
- 417) Acreditavam que perfeição, para o homem, era a imortalidade.
- 418) Nossa lua-de-mel não seria mais a mesma, argumenta Cíntia.
- 419) A autoridade palestina libertou ontem quarenta e cinco integrantes.
- 420) Ficou faltando ainda a devolução do dinheiro da parte aérea.
- 421) A grande diferença é que eu estou procurando solidariedade.
- 422) As chances que apareciam eram perdidas pelos atacantes cruzeirenses.
- 423) A grande virtude da nação Brasileira é a tolerância.
- 424) Tuma e Serra empatam em trinta e sete por cento.
- 425) Vou pegar umas cinco macacas do Fernando Henrique e jogar numa panela.
- 426) Esta diferença acabou se refletindo no índice de julho.
- 427) O objetivo das imagens é mostrar uma história da evolução do universo.
- 428) Segundo ele, a cabine não faz parte do edifício.
- 429) O comitê diz que pagará taxas de serviço a estas empresas.
- 430) A duplicidade se estabelece, então, como condição do escritor.

- 431) Isso quer dizer que os estrangeiros não estarão seguros na Bósnia.
- 432) A principal falha de Parreira ocorre no posicionamento dos dois meias ofensivos.
- 433) O acordo é voltado inicialmente para a área bancária.
- 434) Fui para o México e fiquei lá três anos.
- 435) A permanecer a atual situação, será o Brasil que não vingará.
- 436) Só para o curso de inglês é necessário conhecimento da língua.
- 437) O imposto está previsto para vigorar somente até o final deste ano.
- 438) Ouviu setecentos e onze pessoas maiores de dezesseis anos.
- 439) Oitenta por cento desses adolescentes norte-americanos têm um emprego de meio período.
- 440) O fato é que estamos no terreno do irreconciliável.
- 441) Não há obrigatoriedade de intervalo para jornadas de trabalho.
- 442) Almeida informou que a empresa demitiu quarenta funcionários nos últimos vinte dias.
- 443) O documento não faz qualquer menção aos custos dos assentamentos.
- 444) Observávamos perplexos o vento levar as telhas do galpão.
- 445) Não se conformam que eles, os fatos, sejam fatos.
- 446) Com o debate, o velho trabalhista recuperou a velha agressividade.
- 447) Trata-se do nono subtipo do vírus já identificado no mundo.
- 448) A redução entra na necessidade de gerir a consolidação do real.
- 449) Ana Silva, dezenove anos, está grávida de sete meses.
- 450) Mais uma vez as chuvas de verão atrapalharam a vida dos paulistanos.
- 451) Ele é considerado por seus companheiros de tropa como um oficial moderado.
- 452) Almeida teria sido comunicado que irá deixar o cargo.
- 453) Ela também afirma que só faz programas com camisinha.
- 454) A apuração passou a ser feita antes do sorteio.
- 455) Para lá se dirigem os refugiados, na maioria africanos.
- 456) No Santo André, a luta é para fugir do rebaixamento.
- 457) Eles não têm seguro médico e seu futuro é incerto.
- 458) Como vocês puderam ver na Copa, faltam atacantes na Itália.
- 459) Na temporada, o trânsito na avenida da praia é muito intenso.
- 460) Eles ganharam as primeiras páginas dos jornais do final da semana passada.
- 461) O que aconteceu depois da abertura da bandeira não se sabe.
- 462) Alguns olhos de lince chegaram a enxergar uma cusparada.
- 463) Outros ainda, nada viram, além da abertura da bandeira.
- 464) Eu não quero e não vou falar sobre câmbio e juros.
- 465) Na quinta-feira eu já estava indo para Vitória jogar.
- 466) Dessa forma, os números mais apostados são retirados da disputa.

- 467) Só se pode ser membro do clube enquanto se é solteiro.
- 468) É evidente que este evento é extremamente complexo de se organizar.
- 469) Ser moderno é ser profissional e não inventar desculpas.
- 470) Nos desfiles que assisti, fizeram o público de idiota.
- 471) Moda regional só existe se você se isolar do mundo.
- 472) O desemprego é particularmente grave para mulheres jovens na Espanha.
- 473) As cadernetas rendem vinte e oito por cento no dia quinze.
- 474) Lila Covas ainda não sabe o que vai fazer no fundo
- 475) A “Arte de Amar” e “A Metamorfose” foram publicadas em Roma.
- 476) Sou uma espécie de amor secreto do povo brasileiro.
- 477) Basta que sua remuneração não seja menor que setenta reais.
- 478) Desde então, seu quadro clínico tem deteriorado a cada dia.
- 479) Flu vai à igreja antes de jogar contra o Bahia
- 480) A canção popular é um veículo indiscutível de informação.
- 481) Esse nosso sintetizador de voz vai ficar uma beleza.
- 482) Se engana quem pensa que só vale enfeitar presentes normais.
- 483) A inflação alta também desestimulou as compras a prazo.
- 484) Claro que a burrice é maior do que a Pedra da Gávea.
- 485) Paulino esperava que a reunião proposta por ele fosse marcada para hoje .
- 486) Ele acha difícil, no entanto, que a reunião ocorra.
- 487) O corpo fica vermelho vivo, à medida que sua temperatura sobe.
- 488) Às vezes o preso pega fogo, especialmente quando transpira muito.
- 489) vinte bilhões de dólares já teriam retornado, restando quarenta bilhões.
- 490) Uma exceção foi a crítica de arte Lélia Coelho Frota.
- 491) Como levavam uma vida diferente, eram perseguidos onde chegavam.
- 492) Eles têm razão e se preparam para entrar em greve.
- 493) Com isso se ampliou o espectro da manipulação genética.
- 494) O momento econômico e político deflagrado pelo real é instigante desta reflexão.
- 495) O filme foi lançado pela primeira vez em mil novecentos e setenta.
- 496) Passava o dia vendendo, comprando, barganhando com os fregueses.
- 497) A alta foi de três vírgula vinte e sete por cento.
- 498) O time só caiu nos pênaltis, para a França.
- 499) Empresa que serve ao Ministério da Fazenda aluga comitê aos tucanos.
- 500) Queria ficar, mas não pude; o governo não permitia mais.
- 501) Foi secretário de planejamento do município de São Paulo.
- 502) Dá vontade de eu meter o microfone na sua cabeça.
- 503) Eu fecho o microfone dele, ele abre de novo, poxa.
- 504) Eu já tinha falado pro Ciro que ele foi reclamar.

- 505) Uma página inteira será dedicada a notas e curiosidades.
- 506) Barros é o sexto na classificação geral, com cinquenta pontos ganhos.
- 507) Ele representava uma receita anual de trezentos e cinquenta milhões de dólares.
- 508) Na sua opinião, a prova de inglês deve trazer mais problemas.
- 509) Se eu pensasse assim, não escreveria sobre a Índia.
- 510) O presidente mandou chamá-lo, segundo a imprensa internacional.
- 511) Prefira sempre um especialista plástico a outra modalidade de cirurgia.
- 512) Inicialmente, a revisão deveria terminar em quinze de março.
- 513) A corregedoria investiga se Silva foi omissa no caso da caixinha.
- 514) Em seu lugar, foi nomeado o coronel Alberto Tarjas Neto.
- 515) Pretendo usar um barco pesqueiro Brasileiro na minha instalação na Bienal.
- 516) Os grandes bancos estão procurando apoiar as pequenas instituições.
- 517) A burocracia ineficiente que comanda o futebol Brasileiro levou quase cem jogos.
- 518) A trama é o ponto mais alto do livro.
- 519) Sem a cirurgia, humorista poderia morrer em seis meses
- 520) Esclarecimento à população: Brasil e Camarões só comento amanhã.
- 521) A confederação Brasileira de futebol decretou luto oficial de três dias.
- 522) Com a mesma diferença de gols, a classificação sairá nos pênaltis.
- 523) Cada filhote é portador da metade do patrimônio genético dele próprio.
- 524) A única chance do Bahia aconteceu aos quarenta e três minutos.
- 525) Outro obstáculo às relações afetivas é a competitividade interna.
- 526) Veio para o Brasil em quarenta e nove, fixando-se em Porto Alegre.
- 527) Seu trabalho é extremamente variado no uso de materiais e técnicas.
- 528) Mesmo assim, a revista só pode acontecer de dia.
- 529) Comparado com o mês anterior, registra queda de dez por cento.
- 530) A primeira consiste na recusa da filosofia e no recolhimento ao silêncio.
- 531) A medida valeria após trinta dias desde a sua aprovação.
- 532) O governo poderá adotar ações penais por crime de omissão de socorro.
- 533) Os funcionários estimam uma defasagem salarial em torno de quatrocentos por cento.
- 534) Essas são as operações típicas do mercado flutuante de dólar.
- 535) Eu procuro estar em dia, saber o que se passa.
- 536) O Brasil está na semifinal do décimo segundo campeonato mundial feminino.
- 537) Havia possibilidade de uma nova reunião ser convocada para hoje.
- 538) Sem a retirada, os muçulmanos não compareceram às negociações.
- 539) Esses títulos de nobreza são concedidos pela rainha sob indicação do premiê.
- 540) Está sendo feito um investimento de dois bilhões de dólares.
- 541) Tinha gente que preferia andar quilômetros só para casar em outra igreja.

- 542) Ela estudou teologia espiritual durante dois anos na Itália.
- 543) As crianças serão alimentadas com merenda e irão a pontos turísticos.
- 544) Nós temos um lema: água no umbigo, sinal de perigo.
- 545) Apenas nove por cento afirmam que vão recorrer a empréstimos.
- 546) Disse ter citado os três porque são exemplos de moralidade.
- 547) A pesquisa é feita pelo Procon em convênio com o nosso departamento.
- 548) O australiano vem logo em seguida com quarenta e um pontos.
- 549) Essa possibilidade já havia sido dada aos policiais militares.
- 550) Ele estava no final do colegial quando sentiu as primeiras alterações.
- 551) O sintoma sumiu em poucos dias e não deixou nenhuma sequela.
- 552) O Santos errou sessenta e quatro dos trezentos passes feitos.
- 553) Relator acusa fraude de dois bilhões só neste ano.
- 554) As bancas de jornais do Paraná vão vender preservativos depois do carnaval.
- 555) Rua Venceslau Braz, duzentos e quinze, Botafogo, Zona Sul do Rio.
- 556) Você foi sem calcinha para o camarote do presidente no sambódromo.
- 557) Só fiquei chateada com a história de que tem uma foto.
- 558) O advogado do filósofo, Mário Simas, vai recorrer da decisão.
- 559) Camarões e Rússia não foram adversários à altura para o Brasil.
- 560) O time Brasileiro ainda causa apreensão, passa por altos e baixos.
- 561) Procuro estar atualizado, esquecer o passado e não pensar no futuro.
- 562) Parece que estamos torcendo pelo “brasário” e não pelo Brasil.
- 563) O terreno de grande declividade exige um acompanhamento de engenharia de solos.
- 564) O primeiro pavimento seria destinado à residência do casal de idosos.
- 565) A casa principal estaria localizada no segundo e terceiro pavimentos.
- 566) Pesquisa do Datafolha divulgada ontem revela que Enéas subiu um ponto percentual.
- 567) Enéas afirmou que não se interessa pelos resultados das pesquisas eleitorais.
- 568) Não se comete o mesmo erro várias vezes impunemente.
- 569) O grupo de moda São Paulo promove o seu primeiro encontro.
- 570) Covas passou a campanha inteira dizendo que os policiais ganhavam mal.
- 571) O Brasil tem um enorme potencial em relação à arte pública.
- 572) O prefeito Haroldo Bezerra distribuiu cestas básicas a novecentas famílias desalojadas.
- 573) O passatempo predileto de Leontina será esnobar a pobre Olga.
- 574) Dali poderá surgir uma grande valorização da cultura nacional.
- 575) A câmara teria quinhentas e cinquenta e cinco cadeiras.
- 576) Os custos fixos caíram pela metade no mesmo período.
- 577) As duas análises apontam quais as praias do município próprias para banho.

- 578) O comitê iniciou na cidade uma campanha de combate ao desemprego.
- 579) O melhor resultado foi registrado na região metropolitana de Salvador.
- 580) Cutuque-o e proíba-o de dormir de barriga para cima.
- 581) O censo demográfico feito pelo governo os agrupa em setores distintos.
- 582) Se os problemas vêm para mim é porque eu tinha que agüentar.
- 583) Ela disse que o Alcorão deveria ser revisado cuidadosamente.
- 584) Programa de domingo, oito da noite, aqui na Manchete.
- 585) Em mil e oitocentos, a Irlanda tornou-se parte do Reino Unido.
- 586) O diagnóstico foi revelado posteriormente através de um levantamento nacional na Finlândia.
- 587) No dia treze, haverá concurso para contratar oitenta professores.
- 588) Só falta o jogador concordar para que o negócio seja feito.
- 589) Segundo os organizadores, a municipalização do ecoturismo é a principal proposta.
- 590) Ele, ainda meio confuso, volta para perto da Constituição.
- 591) O gol da vitória saiu ainda no primeiro tempo.
- 592) Substitui Fábio Magalhães, que pediu demissão na semana passada.
- 593) Sua intenção central é tocar os projetos que o Fábio deixou armados.
- 594) Já estava me conscientizando da idade, e me preparando psicologicamente.
- 595) Foi um grande desgaste porque ela ficou dois anos doente.
- 596) Acho que a gastrite foi uma descarga da parte psicológica no corpo.
- 597) Márcia Pinheiro confere as passarelas da estação na Europa.
- 598) Essa comissão será composta por membros da FIFA e da AFA.
- 599) A lei já colocou mais de noventa pessoas na cadeia.
- 600) Atribuiu-se a façanha de ter segurado o Real no grito.
- 601) No dia vinte e dois, foram cinqüenta e quatro mil consultas.
- 602) Mariz está na segunda metade do seu primeiro mandato de senador.
- 603) Resposta: nunca falo de meus romances, e menos ainda deste.
- 604) Em um quadro no canto esquerdo da tela aparecem cenas locais.
- 605) O Chile até hoje está nos treze por cento.
- 606) A abertura de novos grupos foi suspensa pelo governo.
- 607) Voltamos porque essa relação com o público ainda não se esgotou.
- 608) Itamar pode faltar a Mercosul para ir à posse do presidente colombiano
- 609) Os primeiros resultados são esperados para a manhã de hoje.
- 610) No caso da tabela, a correção mensal está prevista na constituição.
- 611) O objeto relacional some no interior imaginário do corpo.
- 612) Segundo ele, tramitam na justiça setenta e dois mil processos desse tipo.
- 613) Tem lucro de um bilhão de dólares por ano.
- 614) Talvez por causa das mudanças que eu busco na minha carreira.

- 615) Lei antitruste gera crise no governo da sucursal de Brasília
- 616) Busco o que meu coração, minha alma e minha inteligência querem.
- 617) O grupo vem de uma temporada no Rio de Janeiro.
- 618) A duração é de cinquenta minutos, a sete reais cada ingresso.
- 619) Anteontem, astronautas realizaram experiências científicas a bordo do ônibus espacial Columbia.
- 620) Os belgas reclamaram, mas o juiz não deu o pênalti.
- 621) Assis também relativiza os riscos de uma explosão do consumo.
- 622) É cedo para se pensar em medidas de controle.
- 623) Isso traria efeitos sobre os financiamentos habitacionais feitos com recursos da poupança.
- 624) Márcio Santos mostrou estar recuperado da tensão da estréia.
- 625) Agora tudo isso acabou e eu fico tendo idéias estranhas.
- 626) Isto aqui não é um colégio de moças e sim de adultos.
- 627) Nas estações espaciais, são feitas experiências de diversos tipos.
- 628) A verdade é que não estou interessado em muitos deles.
- 629) É certo que o desfile pode sobreviver à Mangueira.
- 630) O namoro de férias tem algumas diferenças dos outros tipos de relacionamento.
- 631) É um fator mais importante que o próprio custo do anúncio.
- 632) Os intervalos dos telejornais são hoje os mais valorizados.
- 633) Ao espremer o saco de pano, a casca solta seu sumo.
- 634) Souza será responsável por coordenar o programa de governo do tucano.
- 635) A empresa oferece um serviço de vendas pelo telefone abaixo.
- 636) Serão três espetáculos de teatro e três de dança.
- 637) O balanço também terá que contemplar referências do auditor independente.
- 638) Muitos homens nunca transaram com uma grávida e me procuram.
- 639) O avanço da tecnologia fez esta divisão perder o sentido.
- 640) Não posso obrigar o Rivaldo a jogar fixo na frente.
- 641) O truque, segundo ele, é aproveitar bem a madrugada.
- 642) Não é por acaso que o detetive dessas histórias é um padre.
- 643) O receio era de que eles fossem descobertos pelos soldados.
- 644) Ao mesmo tempo, consagrou pintores apenas por mérito temático.
- 645) Mário de Andrade o converte em protótipo do artista nacional.
- 646) A operação é feita através da emissão de notas fiscais frias.
- 647) dois irmãos gêmeos jogam na seleção da Holanda, adversária do Brasil.
- 648) O Brasil está perdendo pouco a pouco a sua identidade.
- 649) É mais importante saber o que fazem os jogadores no campo.
- 650) Os produtores é que andam muito tímidos, não assumem riscos.
- 651) Muito foi retirado embaixo e outro tanto acrescido em cima.

- 652) A atividade agrícola, aliás, é de um filho de Bisol.
- 653) Nos seus deslocamentos para ir ao dentista, Itamar aparentava mau humor.
- 654) Comprei um jogo de canetas e fui visitá-lo em sua casa.
- 655) Meu tipo inesquecível é o advogado Francisco Rangel Pestana.
- 656) Perdoe o leitor por essas considerações inspiradas no clima de evocação natalino.
- 657) Jogo da Holanda é mais variado que o do Brasil
- 658) Nas finalizações, a Holanda ganha de oitenta a sessenta e três.
- 659) O Brasil só leva vantagem na penetração e na pontaria.
- 660) Empregado do Parlamento Europeu, ele tinha ligações com a bancada socialista.
- 661) Os danos são avaliados em cerca de dois milhões de dólares.
- 662) Devem estar vendo as Américas sob uma ótica totalmente deturpada.
- 663) A maior parte dos títulos raros está em rolos com essa largura.
- 664) Ele controla as principais operações do micro e é considerado seu cérebro.
- 665) Os jogadores treinaram pênaltis pela segunda vez na semana.
- 666) Aí estão alguns endereços para você se divertir muito por lá.
- 667) O vocabulário da sexualidade foi reaparecendo progressivamente há dez anos.
- 668) Junto dela, havia vinte e dois livros com a contabilidade do bicheiro.
- 669) A agência já possui cerca de vinte e cinco profissionais cadastrados.
- 670) Se houver pênalti durante o jogo, o cobrador deverá ser Bebeto.
- 671) Casal de argentinos reclama de hotel de luxo em Salvador
- 672) Nem sempre a hospitalidade baiana faz jus à fama.
- 673) No hotel, eles foram informados que não havia apartamento disponível.
- 674) A lei estabelece maio como referência para este cálculo.
- 675) As entradas podem ser adquiridas pelo telefone acima.
- 676) É a segunda greve convocada pelos aeroportuários em um mês.
- 677) Aliados do tucano prevêem que essas lideranças ficarão neutras.
- 678) Nas cobranças de lateral, só aparecia o pé do jogador.
- 679) No jogo contra Honduras, a única imagem era gerada pela Globo.
- 680) As consultas podem ser feitas por telefone ou por computador.
- 681) O tema saúde voltou a aparecer no momento das fotos.
- 682) Ninguém menos que Pelé classificou a Colômbia como favorita ao título.
- 683) Serve para colocar os dados dos documentos pessoais do portador.
- 684) Isso evitaria as viagens de turismo pelos Estados Unidos.
- 685) Erramos demais o saque porque forçamos de maneira desnecessária.
- 686) Ronaldo conquista a Holanda e persegue os recordes de Pelé
- 687) O julgamento do ex-jogador deve começar em meados de setembro.
- 688) A OTAN já fez cinco outros ataques contra sérvios na guerra.

- 689) Todos eles tiveram alvos limitados, como um tanque específico.
- 690) Quando eu ficar velha, vou me vestir de púrpura.
- 691) Tem nove quartos e em um deles está o mordomo.
- 692) Nos dois dias, os espetáculos começam às vinte horas.
- 693) Trata-se, acima de tudo, de uma questão de sobrevivência.
- 694) São mais comuns em crianças menores de sete anos.
- 695) A taxa de letalidade é de menos de um por cento.
- 696) Se alguém aqui tiver a doença, deve vir ao templo curar-se.
- 697) Estima-se que haja até cem mil ratos morando no templo.
- 698) Firmava a mão na coxa direita sempre que ia subir um batente.
- 699) Não tem necessidade, faz dois meses que eles tiveram férias.
- 700) A técnica poderá ser aplicada se o membro não puder ser reimplantado.
- 701) Nada mais que um jeito matreiro de demonstrar prestígio.
- 702) Será sexta e sábado, às vinte e duas horas.
- 703) O preço varia entre quinze e trinta mil cruzeiros reais.
- 704) Hoje venho quase todos os dias aqui e não perco festas.
- 705) Só vou a São Paulo em aniversário de neto.
- 706) O lateral do Palmeiras afirma que o brinco dá sorte.
- 707) Ele teme a transferência de mercado para o exterior.
- 708) Não é livro para obter sucesso excepcional de público.
- 709) Depois, a conta é encerrada por falta de movimentação.
- 710) Ela não é mais responsável pelo gerenciamento da carreira dos funcionários.
- 711) Pelo jeito, Marieta é uma espécie de mãe de fato.
- 712) Um disquete com dez jogos de domínio público completa o pacote.
- 713) É para encontrar a melhor palavra, disfarça com diplomacia.
- 714) Sua língua com cinto de segurança é um desafio para o repórter.
- 715) Janaina já foi chefona da Associação Municipal dos Estudantes Secundaristas.
- 716) O mesmo acontece nos Estados Unidos e até mesmo na Argentina.
- 717) A causa da morte do feto não foi informada na nota.
- 718) Minatel pediu à Receita Federal as declarações dos últimos anos.
- 719) Ribeiro foi seqüestrado no dia vinte de dezembro de noventa e três.
- 720) Os paulistanos também estão pagando mais caro pelo feijão.
- 721) Em empresas modernas, o funcionário é um colaborador fundamental.
- 722) Seu desejo é de que apareça logo um doador.
- 723) Ele afirmou que não pretende mudar jogadores ou esquema de jogo.
- 724) Se não fosse o Clóvis, eu não teria chegado à presidência.
- 725) O acusado do crime é o ator Guilherme de Pádua.
- 726) Almeida é presidente da liga independente das escolas de samba.
- 727) Perdemos uma ótima oportunidade de fazer a revisão constitucional.

- 728) Segundo ele, os rebeldes estariam salvando os civis de um massacre.
- 729) Pelo menos duzentas mil pessoas morreram na guerra civil ruandesa.
- 730) Paulistana típica, Alessandra está no Rio há três meses.
- 731) Seria morrer com o último tiro da guerra que acabou.
- 732) Para este ano não haverá desequilíbrio nas contas da previdência.
- 733) Mostra o líder chinês em meio a flores e outros padrões decorativos.
- 734) Sem estas, todas as outras invasões são meros paliativos.
- 735) Agora não é hora para isso porque estamos quase na final.
- 736) Primeiro: eu nunca propus a criação de nenhum ministério dos remédios.
- 737) A todo momento ele pedia para ser nomeado presidente da Itaipu Binacional.
- 738) Cinco pessoas já foram presas em conexão com o caso.
- 739) Ontem, Brizola assegurava aos auxiliares que ficaria até o fim.
- 740) Uma eventual desvalorização só ocorreria às vésperas do primeiro turno.
- 741) Não há um departamento de mediadores independente das federações e das agremiações.
- 742) O próprio Lula tenta apresentar razões para o desempenho medíocre.
- 743) Rio de Janeiro: o presidente voltou a cobrar serviço de seus auxiliares.
- 744) Contatos com demais escalões seriam feitos pelo gabinete civil.
- 745) Acho graça chamar esse pequeno ajuste fiscal de plano.
- 746) Se isso fosse plano, qualquer um faria outro numa tarde.
- 747) Ele defende a formulação de um programa mínimo para evitá-la.
- 748) Os que têm cama de solteiro pagam a metade disso.
- 749) Para Pelé, Romário pode dizer as besteiras que quiser.
- 750) Ele é conhecido por nunca dormir duas noites seguidas no mesmo lugar.
- 751) Ele disse acreditar que não deve demorar para chover na região.
- 752) O outro é José Carlos dos Santos, trinta e sete anos.
- 753) A esquerda tem medo de se modificar e deixar de ser dogmática.
- 754) Eu vi apenas dois videoteipes dos amistosos contra Argentina e Alemanha.
- 755) Vou acompanhar melhor os outros jogos do Brasil antes da Copa.
- 756) As opções de lazer são poucas para quem já cansou de futebol.
- 757) Eles nos impuseram a guerra, e decidiram unilateralmente destruir o Iêmen.
- 758) Esses são os nomes de dois pães tradicionalmente consumidos em São Paulo.
- 759) Nem o socialismo nem o neoliberalismo oferecem uma resposta.
- 760) A fuga de capitais e a interrupção dos investimentos alejariam a economia.
- 761) No setor defensivo, o problema é a mudança na escalação.
- 762) Tantas expulsões são raras no São Paulo de Telê.
- 763) Os recursos externos são atraídos pelos juros praticados no país.
- 764) As vendas no ano passado somaram onze mil unidades.
- 765) Em alguns momentos, sua fisionomia revelou a dor que sentia.

- 766) Daqui a pouco, a violência vai ser a única diversão.
- 767) Existem várias redes do gênero, com acesso gratuito ou pago.
- 768) Tudo isso demonstra a solidariedade que existe no coração do povo
- 769) Faz uso dela porque sabe que ela é necessária.
- 770) O treinador pensa em reforçar o setor de marcação do meio campo.
- 771) Pior mesmo é quem não sabe e acredita que sabe.
- 772) A maior surpresa do real foi e continua sendo a política cambial.
- 773) Em poucas palavras: a oferta de dólares supera a procura.
- 774) A superoferta contínua faz com que a moeda estrangeira se desvalorize.
- 775) É uma política oportuna e bem conduzida até agora.
- 776) Isso, entretanto, não significa que se trate de estratégia isenta de riscos.
- 777) Foi demais ver que tínhamos sido os catalisadores de tudo.
- 778) A escolha errada de prancha e vela põe tudo por água abaixo.
- 779) Sozinha, acha que fica mais forte e acaba amedrontando as pessoas.
- 780) Sou livre, mas não a ponto de fazer produção independente.
- 781) A cabeça muda e só hoje sei falar disso sem ficar supergrilada.
- 782) A polícia suspeita que haja no local um arsenal de armas.
- 783) Tenta-se justificar o ganho com os juros reais elevados.
- 784) Eu sou deputado, e não soldado, respondeu ontem Luís Eduardo.
- 785) Não há previsão de prazo para o julgamento da ação.
- 786) Os jogadores norte-americanos só conseguiram desarmar dois ataques da Romênia.
- 787) O Parlamento recusou-se a renovar as linhas de financiamento.
- 788) Destruiria todo esforço feito até agora para equilibrar as contas públicas.
- 789) A proposta deverá ser votada pelo plenário da câmara ainda nesta semana.
- 790) Alguns quadros serão turísticos, mostrando lugares que todo mundo conhece.
- 791) Está cada vez mais difícil encontrar uma voz que seja sincera.
- 792) Nunca havia ido para um treino com seguranças no ônibus.
- 793) Em Patos, foram apreendidas cinco mil e quinhentas arribaçãs, aves em extinção.
- 794) Eram bóias-frias, mas muitos nem bóia tinham nas suas sacolas.
- 795) A televisão virou um açougue, uma máquina de produzir.
- 796) Há algo de maravilhoso e de assustador nessa idéia.
- 797) Um jogo de noventa minutos pode render anos de debate.
- 798) No entender do advogado Celso Botelho de Moraes não há inconstitucionalidade.
- 799) Acha também que o número de partidos deve diminuir.
- 800) Simultaneamente, a amizade inicialmente manifestada por Laura se desfez.
- 801) Talvez o Brasil tenha falado de forma mais clara que os outros.

- 802) O psiquiatra Xavier também acha que o álcool é uma droga.
- 803) Pela França, sagrou-se campeão europeu em oitenta e quatro.
- 804) Neste momento, o futuro guerreiro olha para o céu.
- 805) Farão um quarto para um casal dinâmico, que preserva a intimidade.
- 806) O executivo afirma ter dúvidas sobre a eficiência desse tipo de comunicação.
- 807) As investigações já começaram e não têm data para terminar.
- 808) Os salários reais médios pagos no comércio também apresentaram evolução em agosto.
- 809) Colete à prova de balas é uma tendência em crescimento.
- 810) À frente do número nove teria sido colocado um outro nove.
- 811) Possibilitar que a população se sinta identificada com o estado.
- 812) A entrevista se revelou constrangedora logo em sua apresentação.
- 813) Jô Soares esclareceu ao público que Enéas estava ali não como convidado.
- 814) Por dezesseis votos contra quatro, ele foi mantido fora.
- 815) A euforia pela pista menos quente causou três incidentes.
- 816) Ele era estudante e técnico de manutenção elétrica e hidráulica.
- 817) A polícia ainda investiga a morte do segundo filho, Fábio.
- 818) Os preços no atacado foram os principais responsáveis pela queda.
- 819) No Paraná, em oitenta e nove, conquistou o título estadual no Curitiba.
- 820) Todos os dias, das dezessete horas ao último cliente.
- 821) Todo cidadão terá direito de ver as imagens que lhe dizem respeito.
- 822) Em dezembro de noventa e um a sociedade da empresa foi desfeita.
- 823) Depois do episódio do sambódromo, Itamar Franco ficou abalado.
- 824) Pena: detenção de um a seis meses, ou multa.
- 825) Os pesquisadores que trabalham na Monsanto não conseguiram solucionar o problema.
- 826) O contraceptivo impede o desenvolvimento de ovos dos insetos.
- 827) O trabalho foi apresentado na conferência européia de endocrinologia comparada.
- 828) Perdem todo o conforto e as características de carro de passeio.
- 829) Diz que o momento mais agressivo do debate foi quando alguém bocejou.
- 830) O favoritismo da Mangueira não resistiu a duzentos metros de pista.
- 831) Se estiver vivo, Sequeiros completou ontem quarenta e dois anos.
- 832) Se alguma coisa falhar, será a falência do programa nacional.
- 833) Jobim divulga hoje seus pareceres sobre o capítulo do poder judiciário.
- 834) É considerado o maior contista e um dos maiores dramaturgos russos.
- 835) Só depois que se descobre é que parece ser claro.
- 836) Sua possível vitória assegura um caminho de otimismo para o país.
- 837) O esforço físico e a atitude são característicos do futebol britânico.

- 838) Este foi o cenário de ontem do mercado financeiro.
- 839) O dólar em queda serve como uma represa para os preços.
- 840) Trabalhamos exclusivamente com tipos que fogem do padrão tradicional de beleza.
- 841) O Mogi vem a campo depois de uma semana tumultuada.
- 842) As outras pessoas que estavam no campo de futebol nada sofreram.
- 843) Não vejo motivo para impugnar uma candidatura presidencial por tão pouco crime.
- 844) Já fui muito amigo do uísque para relaxar, mas depois brigamos.
- 845) O Itamaraty só exige que a decisão seja tomada pelas Nações Unidas.
- 846) Volto ao jogo de logo mais com a Suécia.
- 847) Não há, assim, motivo de temor para o Brasil.
- 848) Não há como controlar o uso dos recursos por parte das equipes.
- 849) Para ele, a concorrência é boa para os consumidores.
- 850) Já o intercomunicador entre motos sai por duzentos e trinta reais.
- 851) Termo genérico que pode representar tanto um custo como uma despesa.
- 852) Vitória: dizem que o Brasil perdeu o bonde da história.
- 853) Valores válidos para apartamentos novos na segunda semana de novembro.
- 854) Quatro trios elétricos vão animar a festa até a próxima terça-feira.
- 855) O Brasil também é cor de rosa e carvão.
- 856) A ética é o domínio do absoluto, do categórico.
- 857) Ninguém me ama ninguém me quer ninguém me chama de meu amor.
- 858) A nova caixa é italiana, semelhante à do Tipo.
- 859) A maioria dos clubes e danceterias da cidade vai estar funcionando normalmente.
- 860) O atacante Romário errou um dos três pênaltis no treino de anteontem.
- 861) O outro atacante titular, Bebeto, acertou suas duas cobranças.
- 862) Nos anos sessenta Marilena viveria uma de suas melhores experiências.
- 863) Pode ser que estas contradições todas ainda perdurem muito.
- 864) O Figueiredo precipitou alguma coisa e não deu para trabalhar com ele.
- 865) No dia de seu aniversário, Rose não saiu de casa.
- 866) A história desses milhares de índios é igual à de muitas tribos.
- 867) A Folha devia dar seguro contra boladas do Branco.
- 868) Somente na última rodada serão definidas as colocações das equipes.
- 869) Finalmente, decidiu-se por um dicionário em um volume.
- 870) Os depósitos realizados foram todos transferidos para o dia primeiro de junho.
- 871) Reúne a coleção nacional de arte britânica desde o século dezessete.
- 872) É um dos eventos mais importantes do país este ano.
- 873) Teve gente que falou que no meu lugar iria se matar.

- 874) O documento foi entregue a Lula, no início da noite.
- 875) Pegar bastante malária é bom porque a pessoa vai ficando imune.
- 876) A autoridade policial tem de estar embasada em provas.
- 877) Pela primeira vez nesta Copa se viu dezesseis seleções lutando pela vitória.
- 878) Por outro lado, muitos remunerados ainda jogavam por amor à camisa.
- 879) Portugal dava-me tudo o que eu podia receber dele.
- 880) O concurso para contratação desses profissionais deverá ser divulgado em outubro.
- 881) Não estava presente porque viu que aquilo era inevitável.
- 882) Quatro homens entraram separadamente pela porta principal da agência.
- 883) Dois outros homens aguardavam a conclusão do assalto no estacionamento.
- 884) Os seguranças e o gerente foram dominados simultaneamente pelos assaltantes.
- 885) Houve princípio de pânico entre os clientes e funcionários do banco.
- 886) Os assaltantes retiraram as armas dos seguranças da agência.
- 887) Sexo, para ela, significa estar amando profundamente um homem.
- 888) Moradores do Grande Rio voltam a viver momentos de violência.
- 889) Seleção russa desfalcada faz dois jogos caça-níqueis no sul do país
- 890) Vive-se, no litoral norte de São Paulo, o verão das sucursais.
- 891) Teatro no centro cultural São Paulo: confusão, confusão, confusão.
- 892) Esse é um dos aspectos que têm que ser revistos.
- 893) O termo caça-níqueis se aplica exatamente aos dois jogos.
- 894) Os desmandos até agora observados são graves e têm de ser apurados.
- 895) Um em cada dez entrevistados disse que o Brasil não teve destaques.
- 896) O Datafolha entrevistou apenas moradores da cidade de São Paulo.
- 897) Esta seção vai abordar as obras literárias exigidas pelos principais vestibulares.
- 898) Em abril a taxa foi de dois vírgula seis por cento.
- 899) Então, é claro que aí o pessoal do partido me procurou.
- 900) A solução foi fazer um sol falso em computador.
- 901) Sem falar nos que ainda não acertaram o fornecimento de motores.
- 902) Ser progressista hoje é ligar o pensamento liberal com a social democracia.
- 903) Mancuso mostrou boa forma e nunca arrastou a perna nos treinos.
- 904) Em Campinas nos inspiramos para concluir as mil frases.
- 905) Bia Lessa leu esse espelho assustador da humanidade e ficou tocada.
- 906) As melhores chances de medalha do Brasil estão na natação.
- 907) Nas margens do rio, não havia ninguém para ver o candidato.
- 908) Segundo ele, ser ou não titular vai depender do treinador Parreira.
- 909) O problema não são as reformas, mas a política econômica.
- 910) As emissoras que aceitaram nossas condições irão receber as imagens.
- 911) De vez em quando ele entrava em contato com os garotos.

- 912) O cara era um armário e brandia um facão persuasivo.
- 913) A expectativa é de que o real reverta a tendência.
- 914) São tantas, que você aprende a conviver com elas.
- 915) O grupo formado por veículos e construção manteve os preços estáveis.
- 916) A Constituição Federal proíbe mudanças nas regras em ano eleitoral.
- 917) Isso é proibido e é um desrespeito ao consumidor.
- 918) Os aumentos foram proporcionados pelo crescimento das vendas no comércio.
- 919) Ele disse que só foi solto depois que antecipou sua volta.
- 920) Já começam a surgir os efeitos na balança comercial.
- 921) Em um ano, papou todos os títulos que disputou.
- 922) Os preços variam entre quinze e sessenta dólares o metro quadrado.
- 923) A Assembléia foi convocada pelo sindicato dos policiais federais no Rio.
- 924) Achava engraçado ser pós-moderna sem saber que diabos isso significava.
- 925) Digo só que a banda e a música são boas.
- 926) Com a suspensão traseira quebrada, Gil abandonou a prova.
- 927) Os empréstimos do Banco Mundial exigem contrapartidas dos governos beneficiados.
- 928) Começou apresentando um Lula de terno e gravata tentando convencer como estadista.
- 929) Ele afirmou também que dará a ordem quando quiser.
- 930) Os juros baixos têm sido a chave da recuperação econômica.
- 931) Os participantes do desfile usavam as cores dos países participantes.
- 932) Foi seguido o estilo tradicional das paradas festivas americanas.
- 933) O time dirigido pelo brasileiro está em décimo lugar no campeonato espanhol.
- 934) É sempre bom alguém do lado para trocar idéia.
- 935) A Argentina foi um dos refúgios favoritos dos nazistas alemães.
- 936) Há um roteiro de atrações que é cumprido à risca.
- 937) Ciro pede que Serra explique em público por que criticou o plano.
- 938) Houve empate em um a um na primeira fase da competição.
- 939) O motivo foi o deslizamento de terra que aconteceu anteontem.
- 940) A explicação é o número de dias úteis da aplicação.
- 941) São as reservas externas em poder do Banco Central.
- 942) O ministro me autorizou a dizer que não há intervenção federal.
- 943) Vamos participar das instâncias partidárias e reforçar a democracia interna.
- 944) Em torno da múmia, havia vasilhas e pratos de cerâmica negra.
- 945) A Seleção Brasileira feminina de basquete é campeã mundial.
- 946) Hortência e Paula, as estrelas, devem abandonar a equipe.
- 947) Leia sobre o título nas páginas oito e nove.
- 948) A atual montagem de Dorotéia traz ao palco seis atrizes negras.

- 949) Rodrigues foi escolhido, pois sua obra tem uma temática negra.
- 950) Depois da vitória, afirmou que Camarões é o mais perigoso.
- 951) Vai ser uma espécie de mapa da cidade artística brasileira.
- 952) Roberto Rodrigues é secretário da agricultura e abastecimento de São Paulo.
- 953) O Brasil passou pela Rússia, sem sustos nem cintilações.
- 954) Venda de veículos no varejo bate novo recorde em maio.
- 955) Já não são aqueles negros alegres e maravilhosos de copas passadas.
- 956) Desconfio que não resistirão à nossa aplicada tática de jogo.
- 957) Observadores políticos acham que ele tem grandes chances de voltar ao cargo.
- 958) Resumindo: raros trabalhadores brasileiros infelizmente conseguem tanta segurança e padrão salarial.
- 959) Daí ter sido publicado nesta coluna que eles estão reclamando de barriga cheia.
- 960) O oficial trabalha como gerente da Loja de Lelé.
- 961) Infeliz de quem errar em alguma decisão: leva vermelho na hora.
- 962) Mesmo assim, a companhia prevê lentidão no trânsito na região.
- 963) Esta resolução entra em vigor na data de sua publicação.
- 964) Torcedores ocuparam o tráfego em cinco quarteirões do centro da cidade.
- 965) Captadas por antenas parabólicas, as palavras do ministro viraram munição petista.
- 966) O futebol serviu à Bolívia como janela para o mundo.
- 967) Fundação Pró Sangue tem estoque só para dois dias de abastecimento
- 968) Esse número é suficiente para apenas dois dias de abastecimento.
- 969) A UEFA também pensa que nós devemos ter mais uma vaga garantida.
- 970) Foi também, em mil novecentos e setenta e nove, prefeito de Santos.
- 971) É a condição para que todos afinal saiam ganhando.
- 972) O cercadinho dos animais é decorado com ovos coloridos e cenouras.
- 973) Na época, os irmãos e primos da vítima juraram vingança.
- 974) Como mulheres e advogadas, nunca mais devemos nos intimidar.
- 975) Empregados participam da sua formulação e se sentem comprometidos com eles.
- 976) Os esforços iniciais se concentrarão na conservação da biodiversidade.
- 977) O partido precisaria, de resto, abrir-se para composições políticas.
- 978) Conversões feitas com câmbio a quatrocentos e oitenta e sete cruzeiros reais.
- 979) Os poetas sempre estiveram namorando as rosas e inventando coisas para elas.
- 980) Hoje, elas saem mais facilmente desse mundo fechado.
- 981) A Universidade Federal da Paraíba está desenvolvendo cartões inteligentes.
- 982) Siqueira e Arruda tentaram fugir e foram detidos pelos policiais.
- 983) A defesa argumenta que os agentes federais não tinham mandado de busca.
- 984) Além do São Paulo, preciso me preocupar com a seleção.

- 985) É o que salta do discurso de ontem aos senadores.
- 986) A grande novidade de sua fala está na ausência de novidades.
- 987) O técnico do Palmeiras é cauteloso, apesar do favoritismo.
- 988) Reservadamente, confessa não entender o comportamento do presidente.
- 989) Ele defende o presidente da Federação do Rio, Eduardo Viana.
- 990) É um público formador de opinião como nenhum outro.
- 991) Os documentos que chegaram a Brasília passam por análise financeira e criminal.
- 992) Percorrem cinco etapas até estarem à disposição dos delegados.
- 993) No momento do crime, trinta alunos assistiam aula na sala.
- 994) Um grupo de senadores está articulando um movimento para fortalecê-lo.
- 995) Brasília convive há quase dois meses com uma constante névoa seca.
- 996) Fez também a personagem Terezinha da Fronteira, na novela Fera Ferida.
- 997) A mãe de Alfredo diz que o marido não participou do crime.
- 998) Depois que a Mangueira passar, o suspense dominará a passarela.
- 999) Algumas analisam a possibilidade de emprestar até trinta e seis meses.
- 1000) A dificuldade é que a captação ainda continua no curto prazo.

C.2 161 frases nas emoções: Alegria, Tristeza e Raiva.

Conjunto de 161 frases gravadas com fala alegre, triste e com raiva:

- 1) Seus computadores processam até dois milhões de informações por segundo.
- 2) O mais desculpável em Itamar são os seus vexames de passarela.
- 3) Para Branco, este é um jogo como outro qualquer.
- 4) Até agosto, o valor era de sessenta e quatro reais.
- 5) O plano terá recursos no valor de cento e dezoito mil.
- 6) As propostas de política econômica são poucas e imediatas.
- 7) Agora, continua sendo necessário que a Assembléia Legislativa aprove o nome.
- 8) Se conseguirmos uma parte do que ele alcançou, será demais.
- 9) Às vezes, numa crise ou numa derrota se aprende mais.
- 10) As vendas por catálogo já atingem cinqüenta e oito bilhões de dólares.
- 11) Dos duzentos imóveis novos oferecidos no mês, apenas quinze foram comercializados.
- 12) Para ele, mesmo a carne está sem perspectiva de alta.
- 13) É muito importante divulgar a idéia da qualidade para o público.

- 14) É a maior campanha da empresa realizada até hoje.
- 15) Esta diferença acabou se refletindo no índice de julho.
- 16) O acordo é voltado inicialmente para a área bancária.
- 17) O imposto está previsto para vigorar somente até o final deste ano.
- 18) Há demanda por real, não por dólar, cuja cotação cai.
- 19) O mercado fica de alto risco a curto prazo.
- 20) Os brasileiros no exterior ganharam um poder aquisitivo sem precedentes.
- 21) Como vocês puderam ver na Copa, faltam atacantes na Itália.
- 22) Eles ganharam as primeiras páginas dos jornais do final da semana passada.
- 23) Ele acha difícil, no entanto, que a reunião ocorra.
- 24) O momento econômico e político deflagrado pelo real é instigante desta reflexão.
- 25) Barros é o sexto na classificação geral, com cinquenta pontos ganhos.
- 26) Em seu lugar, foi nomeado o coronel Alberto Tarjas Neto.
- 27) Os grandes bancos estão procurando apoiar as pequenas instituições.
- 28) A medida valeria após trinta dias desde a sua aprovação.
- 29) Está sendo feito um investimento de dois bilhões de dólares.
- 30) Chega a secretário da cultura com amplo apoio de intelectuais e artistas.
- 31) Camarões e Rússia não foram adversários à altura para o Brasil.
- 32) Pesquisa do Datafolha divulgada ontem revela que Enéas subiu um ponto percentual.
- 33) Não se comete o mesmo erro várias vezes impunemente.
- 34) Covas passou a campanha inteira dizendo que os policiais ganhavam mal.
- 35) Os custos fixos caíram pela metade no mesmo período.
- 36) Se os problemas vêm para mim é porque eu tinha que aguentar.
- 37) Só falta o jogador concordar para que o negócio seja feito.
- 38) O gol da vitória saiu ainda no primeiro tempo.
- 39) Substituí Fábio Magalhães, que pediu demissão na semana passada.
- 40) Sua intenção central é tocar os projetos que o Fábio deixou armados.
- 41) Essa comissão será composta por membros da FIFA e da AFA.
- 42) No dia vinte e dois, foram cinquenta e quatro mil consultas.
- 43) O Chile até hoje está nos treze por cento.
- 44) Segundo ele, tramitam na justiça setenta e dois mil processos desse tipo.
- 45) Tem lucro de um bilhão de dólares por ano.
- 46) Talvez por causa das mudanças que eu busco na minha carreira.
- 47) Busco o que meu coração, minha alma e minha inteligência querem.
- 48) O número de convocados por vaga é de doze candidatos.
- 49) Dessas, somente umas trezentas e vinte foram inauguradas em território americano.
- 50) Em Florianópolis, foi registrado dois graus celsius na manhã de domingo.

- 51) Itamar tem razão de estar exultante como nunca desde que virou presidente.
- 52) O problema é que a inflação demora um pouco para cair.
- 53) Visitados todos, na ocasião, por delegações de grandes empresários.
- 54) Tem-se uma receita mensal de trezentos e quarenta mil dólares.
- 55) Juliana disse que o banco a ser escolhido é surpresa.
- 56) Este mês os termômetros descem facilmente abaixo de zero.
- 57) Agora tudo isso acabou e eu fico tendo idéias estranhas.
- 58) Os produtores é que andam muito tímidos, não assumem riscos.
- 59) Nas finalizações, a Holanda ganha de oitenta a sessenta e três.
- 60) Empregado do Parlamento Europeu, ele tinha ligações com a bancada socialista.
- 61) Devem estar vendo as Américas sob uma ótica totalmente deturpada.
- 62) Apesar disso, continua ligado ao seu país de origem.
- 63) Quanto mais avançarmos nas reformas, menor será esse ônus.
- 64) É a oportunidade para o mercado observar os talentos mais promissores.
- 65) Dizem que a disciplina é importante e se embriagam às escondidas.
- 66) Patrocínio rende à CBF vinte e três vírgula seis milhões de dólares.
- 67) Ela disse que ainda decidirá seu voto para a câmara dos deputados.
- 68) A Latasa vinha solicitando ao governo a adoção da medida.
- 69) Dos oitenta e um senadores, apenas cinqüenta e seis votaram.
- 70) O time cometeu vinte e seis faltas durante a partida.
- 71) Foi o que afirmaram sessenta e dois por cento dos oitocentos entrevistados.
- 72) Esta é a opinião de trinta e dois por cento do público.
- 73) A equipe de Americana ocupa a oitava posição.
- 74) A nova diretoria tomará posse em primeiro de Fevereiro.
- 75) Paulino esperava que a reunião proposta por ele fosse marcada para hoje.
- 76) Uma página inteira será dedicada a notas e curiosidades.
- 77) A corregedoria investiga se Silva foi omissos no caso da caixinha.
- 78) A trama é o ponto mais alto do livro.
- 79) Cada filhote é portador da metade do patrimônio genético dele próprio.
- 80) A única chance do Bahia aconteceu aos quarenta e três minutos.
- 81) Seu trabalho é extremamente variado no uso de materiais e técnicas.
- 82) A medida valeria após trinta dias desde a sua aprovação.
- 83) Essas são as operações típicas do mercado flutuante de dólar.
- 84) O Brasil está na semifinal do décimo segundo campeonato mundial feminino.
- 85) Havia possibilidade de uma nova reunião ser convocada para hoje.
- 86) Esses títulos de nobreza são concedidos pela rainha sob indicação do premiê.
- 87) Está sendo feito um investimento de dois bilhões de dólares.
- 88) Tinha gente que preferia andar quilômetros só para casar em outra igreja.
- 89) Ela estudou teologia espiritual durante dois anos na Itália.

- 90) Apenas nove por cento afirmam que vão recorrer a empréstimos.
- 91) A pesquisa é feita pelo Procon em convênio com o nosso departamento.
- 92) O australiano vem logo em seguida com quarenta e um pontos.
- 93) Essa possibilidade já havia sido dada aos policiais militares.
- 94) Ele estava no final do colegial quando sentiu as primeiras alterações.
- 95) O Santos errou sessenta e quatro dos trezentos passes feitos.
- 96) Relator acusa fraude de dois bilhões só neste ano.
- 97) As bancas de jornais do Paraná vão vender preservativos depois do carnaval.
- 98) Enéas afirmou que não se interessa pelos resultados das pesquisas eleitorais.
- 99) O grupo de moda São Paulo promove o seu primeiro encontro.
- 100) Os custos fixos caíram pela metade no mesmo período.
- 101) As duas análises apontam quais as praias do município próprias para banho.
- 102) O comitê iniciou na cidade uma campanha de combate ao desemprego.
- 103) O melhor resultado foi registrado na região metropolitana de Salvador.
- 104) A lei já colocou mais de noventa pessoas na cadeia.
- 105) Mariz está na segunda metade do seu primeiro mandato de senador.
- 106) A abertura de novos grupos foi suspensa pelo governo.
- 107) Voltamos porque essa relação com o público ainda não se esgotou.
- 108) Os primeiros resultados são esperados para a manhã de hoje.
- 109) O grupo vem de uma temporada no Rio de Janeiro.
- 110) É cedo para se pensar em medidas de controle.
- 111) Márcio Santos mostrou estar recuperado da tensão da estréia.
- 112) Serão três espetáculos de teatro e três de dança.
- 113) O avanço da tecnologia fez esta divisão perder o sentido.
- 114) Não posso obrigar o Rivaldo a jogar fixo na frente.
- 115) Não é por acaso que o detetive dessas histórias é um padre.
- 116) O receio era de que eles fossem descobertos pelos soldados.
- 117) O Brasil está perdendo pouco a pouco a sua identidade.
- 118) Comprei um jogo de canetas e fui visitá-lo em sua casa.
- 119) Jogo da Holanda é mais variado que o do Brasil
- 120) O Brasil só leva vantagem na penetração e na pontaria.
- 121) A maior parte dos títulos raros está em rolos com essa largura.
- 122) Os jogadores treinaram pênaltis pela segunda vez na semana.
- 123) Nem sempre a hospitalidade baiana faz jus à fama.
- 124) A lei estabelece maio como referência para este cálculo.
- 125) É a segunda greve convocada pelos aeroportuários em um mês.
- 126) As consultas podem ser feitas por telefone ou por computador.
- 127) O tema saúde voltou a aparecer no momento das fotos.
- 128) O julgamento do ex-jogador deve começar em meados de setembro.

- 129) A OTAN já fez cinco outros ataques contra sérvios na guerra.
- 130) Ele teme a transferência de mercado para o exterior.
- 131) Ele afirmou que não pretende mudar jogadores ou esquema de jogo.
- 132) Ele é conhecido por nunca dormir duas noites seguidas no mesmo lugar.
- 133) As vendas no ano passado somaram onze mil unidades.
- 134) O treinador pensa em reforçar o setor de marcação do meio campo.
- 135) O psiquiatra Xavier também acha que o álcool é uma droga.
- 136) À frente do número nove teria sido colocado um outro nove.
- 137) A entrevista se revelou constrangedora logo em sua apresentação.
- 138) Os preços no atacado foram os principais responsáveis pela queda.
- 139) O contraceptivo impede o desenvolvimento de ovos dos insetos.
- 140) Jobim divulga hoje seus pareceres sobre o capítulo do poder judiciário.
- 141) Este foi o cenário de ontem do mercado financeiro.
- 142) O dólar em queda serve como uma represa para os preços.
- 143) O Brasil também é cor de rosa e carvão.
- 144) A Folha devia dar seguro contra boladas do Branco.
- 145) Não estava presente porque viu que aquilo era inevitável.
- 146) Esse é um dos aspectos que têm que ser revistos.
- 147) O termo caça-níqueis se aplica exatamente aos dois jogos.
- 148) O Datafolha entrevistou apenas moradores da cidade de São Paulo.
- 149) Em abril a taxa foi de dois vírgula seis por cento.
- 150) As melhores chances de medalha do Brasil estão na natação.
- 151) De vez em quando ele entrava em contato com os garotos.
- 152) A Argentina foi um dos refúgios favoritos dos nazistas alemães.
- 153) Houve empate em um a um na primeira fase da competição.
- 154) Esta resolução entra em vigor na data de sua publicação.
- 155) Esse número é suficiente para apenas dois dias de abastecimento.
- 156) Os poetas sempre estiveram namorando as rosas e inventando coisas para elas.
- 157) A defesa argumenta que os agentes federais não tinham mandado de busca.
- 158) A grande novidade de sua fala está na ausência de novidades.
- 159) Um grupo de senadores está articulando um movimento para fortalecê-lo.
- 160) A mãe de Alfredo diz que o marido não participou do crime.
- 161) A dificuldade é que a captação ainda continua no curto prazo.

C.3 85 frases interrogativas.

Conjunto de 85 frases interrogativas gravadas:

- 1) Os funcionários da fábrica ainda não chegaram por quê?
- 2) Você sabe por que me apaixonei por você?
- 3) Por que ele sempre usa este boné?
- 4) Quem foi contigo ao cinema na semana passada?
- 5) Como se chama o amigo do Pedro que viajou ontem?
- 6) O Antônio já chegou com as malas prontas para ficar?
- 7) Quando foste de férias pela última vez à Europa?
- 8) De quem é aquele livro com gravuras esquisitas na capa?
- 9) Gostas muito de chocolate branco com recheio?
- 10) Que acontecimentos foram registrados no dia de hoje?
- 11) Quem gritou com a criança para ela estar chorando?
- 12) Quantos vêm para a festa de fim de ano aqui da empresa?
- 13) Qual será sua decisão após a apuração dos fatos?
- 14) Ela não disse que dia iria chegar dos emirados árabes?
- 15) Não ficou claro quem ganhou o debate político na segunda-feira?
- 16) Qual é o teu nome?
- 17) Quem escolheu teu nome e por quê?
- 18) O que representa este nome?
- 19) Que sentido podes atribuir ao teu nome hoje?
- 20) Como está tua relação com a saúde?
- 21) Adoeces com frequência?
- 22) Já te perguntaste o que as doenças têm a ver contigo?
- 23) Fazes uso de medicação?
- 24) Como está tua educação alimentar?
- 25) Sente-se bem com o teu sono?
- 26) Como te sentes em relação aos teus pais?
- 27) Participas de grupos sociais, religiosos ou profissionais? Para quê?
- 28) Como estás vivendo o aspecto lazer?
- 29) Reclamas muito da vida? Em que sentido?
- 30) Culpas muito os outros pelos fatos errados ou negativos na tua vida?
- 31) O que significa “respeitar o corpo” pra ti?
- 32) Como lidas com a mentira e a fofoca?
- 33) O que podes fazer pra ser um pouco mais feliz?
- 34) Como vives tua sexualidade? Por que?
- 35) O que podes fazer pra não repetir mais alguma característica da tua personalidade ou comportamento que te incomoda?
- 36) O que podes fazer pra melhorar tua relação com alguma pessoa que lembras agora?
- 37) Qual a característica da tua mãe que é mais negativa?

- 38) Qual a característica do teu pai que é mais negativa?
- 39) Qual característica tua é mais parecida com as da mãe?
- 40) Qual característica tua é mais parecida com alguma que o pai tenha?
- 41) Achas que fostes educado pra ser saudável, feliz e bem sucedido? Então...? O que depende de ti se não colocares nenhum culpado na tua história?
- 42) O que mais gostas de fazer no teu lazer?
- 43) Sentes-te responsável por isso?
- 44) O que de fato te fez sofrer nos últimos anos?
- 45) Estás mais preso as dores e sofrimentos ou aproveitas a tua passagem “por aqui” como única chance de ser feliz?
- 46) O que extrairias de dentro da tua personalidade?
- 47) O que acrescentarias à tua vida hoje a fim de ser mais feliz?
- 48) Se dependesse de ti como começarias a agir a partir de hoje?
- 49) Tens o hábito de contar teu sonho pra alguém e tentas entendê-lo a partir de ti mesmo?
- 50) Que fato, momento ou circunstância podes atribuir como a marca do início da tua adolescência?
- 51) Quem eram os teus amigos?
- 52) Ocupavas o teu tempo com que atividades? O que sentias por isso?
- 53) Tivestes alguma doença?
- 54) Tomastes alguma medicação na adolescência?
- 55) Quais as pessoas da tua família que mais participaram da tua vida dos 12 aos 22 anos?
- 56) O que lembras com muita saudade?
- 57) O que não gostarias que tivesse acontecido?
- 58) Quais os colegas de brincadeiras e festas que tu lembras?
- 59) Como teus pais se relacionavam contigo?
- 60) Como tua mãe se relacionava contigo?
- 61) Se tinhas irmãos, como eles se relacionavam contigo?
- 62) Foi uma fase mais feliz ou triste? O que justifica isso?
- 63) Qual a característica da personalidade mais marcante em ti nesta fase?
- 64) Que costume ou atitude tens hoje que lembras ser igual ou diferente da tua adolescência?
- 65) Em que cidade tu nasceste?
- 66) Quantos habitantes moravam lá na época?
- 67) Existia alguma pessoa “folclórica” que todos lembram dela por atos excêntricos ou engraçados nesta cidade?
- 68) Os pais de teus pais moravam na mesma cidade?
- 69) Quais aspectos físico-geográficos, que tu lembras na tua infância, da cidade em

que nasceste?

- 70) Brincavas com quem e que brincadeiras gostavas?
- 71) Como te relacionavas com os amigos e vizinhos?
- 72) Como te sentias com a tua família?
- 73) Como te relacionavas com o pai?
- 74) Como te relacionavas com a mãe?
- 75) Como te relacionavas com os irmãos?
- 76) Caso não tinhas irmãos o que sentias por ser filho único?
- 77) Caminhastes com que idade?
- 78) Falastes quando? Quais as primeiras palavras?
- 79) Lembras de uma palavra que era característica tua na infância?
- 80) Tiveste alguma doença natural da infância?
- 81) Tiveste outras doenças? Como reagistes?
- 82) Como teus cuidadores reagiram?
- 83) Costumavas te machucar muito fisicamente?
- 84) Como era a característica tua em relação à birra e a reclamações?
- 85) O choro era freqüente ou não? Quando?

C.4 29 frases exclamativas.

Conjunto de 29 frases exclamativas gravadas:

- 1) Que prova difícil!
- 2) É uma delícia esse bolo!
- 3) Nossa! Que legal!
- 4) Oba! Vamos ao cinema hoje!
- 5) Não quero mais ver você na minha frente!
- 6) Vá fazer sua lição de casa agora!
- 7) Que blusa linda!!!
- 8) Vai começar tudo de novo!
- 9) Saia já daqui, por favor!
- 10) Que alegria, meu Deus!
- 11) Silêncio!
- 12) Bravo! Bravo!
- 13) O final da peça foi lindo!
- 14) Seu fim foi extremamente triste!
- 15) Estou cansadíssima!
- 16) Isso nos deixou muito felizes!

- 17) Oh! que pena! O rio da minha infância está poluído!
- 18) Veja como cresceram as plantas que você plantou!
- 19) Oh! quanta maldade maltratar os animais!
- 20) Olhe as nuvens no céu! Parecem flocos de neve!
- 21) Que fascinante é essa história!
- 22) Essa história é horrível!
- 23) Esta moça é tão idiota!
- 24) Este rapaz não é nada tolo!
- 25) Já comeu a sopa toda! Que lindo!
- 26) O jantar já foi todo comido! Boa!
- 27) O bebê comeu a papinha!
- 28) Que delícia esta comida!
- 29) Como comeste!

C.5 23 Frases com onomatopéias diversas

Conjunto de 23 frases gravadas com diversas onomatopéias:

- 1) Aaai! Está doendo muito!
- 2) Ai, ai! Como estou cansado!
- 3) Ahhh! Que alívio!
- 4) Vou dar um grito! Ah!
- 5) Ha Ha Ha! Isto tudo é muito engraçado!
- 6) Atchim! Acho que estou resfriando! Coff! Coff!
- 7) Buáá! Minha mãe me bateu!
- 8) Hummmmm! Isto não vai dar certo!
- 9) Oops! Quase caí!
- 10) Arghn! Que nojo!
- 11) Shhhh! Façam silêncio!
- 12) EeeK! ic! Estou soluçando de novo!
- 13) Hã? Huh? Hein? São muitas as interrogações!
- 14) He! he! he! Essa eu ganhei!
- 15) Hmmm hum... Agora estou entendendo tudo!
- 16) Hum! Que cheiro gostoso!
- 17) Ã-hã! Foi sua única resposta.
- 18) Ihhhhhh! Isto não vai dar certo!
- 19) Uuuuuuhhh! Juiz ladrão!
- 20) Tsc, tsc, tsc, tsc! Errou de novo!

- 21) Ohhhhh! Como você é inteligente!
- 22) Aahhh! Assim não vale!
- 23) Eeeeeeehh! Ganhei de novo!

C.6 Aprenda a dizer te amo

Esta história é baseada em um fato verídico.

Não tenha vergonha de dizer Eu Te Amo!

Dois irmãozinhos brincavam em frente de casa, jogavam bolinhas de gude. Quando Júlio, o menino mais novo, disse ao irmão Ricardo:

- Meu querido irmão, eu te amo muito e nunca quero me separar de você!

Ricardo sem dar muita importância ao que Júlio disse, pergunta: - O que deu em você moleque? Que conversa besta é essa de amar? Quer calar a boca e continuar jogando?

E os dois continuaram jogando a tarde inteira até anoitecer. À noite, o senhor Jacó, pai dos garotos, chegou do trabalho, estava exausto e muito mal humorado, pois não havia conseguido fechar um negócio importante.

Ao entrar, Jacó olhou para Júlio que sorriu para o pai e disse: - Olá papai, eu te amo muito e não quero nunca me separar do senhor!

Jacó no auge de seu mal humor e stress disse: - Júlio, estou exausto e nervoso, então por favor não me venha com besteiras!

Com as palavras ásperas do pai, Júlio ficou magoado e foi chorar no cantinho do quarto. Dona Joana, mãe dos garotos, sentindo a falta do filho foi procurá-lo pela casa, até que o encontrou no cantinho do quarto com os olhinhos cheios de lágrimas.

Dona Joana, espantada, começou a enxugar as lágrimas do filho e perguntou:

- O que foi Júlio, por que choras?

Júlio olhou para a mãe, com uma expressão triste e lhe disse: - Mamãe, eu te amo muito e não quero nunca me separar da senhora!

Dona Joana sorriu para o filho e lhe disse: - Meu amado filho, ficaremos sempre juntos! Júlio sorriu, deu um beijo na mãe e foi se deitar.

No quarto do casal, ambos se preparando para se deitar, Dona Joana pergunta para seu marido Jacó:

- Jacó, o Júlio está muito estranho hoje, não acha?

Jacó muito estressado com o trabalho disse à esposa:

- Esse moleque só está querendo chamar a atenção... Deita e dorme mulher!

Então todos se recolheram e todos dormiam sossegados. Às 2 horas da manhã, Júlio se levanta, vai ao quarto de seu irmão Ricardo e fica observando o irmão dormir...

Ricardo, incomodado com a claridade, acorda e grita com Júlio:

- Seu louco, apaga essa luz e me deixa dormir!

Júlio em silêncio obedeceu o irmão, apagou a luz e se dirigiu ao quarto dos pais...

Chegando ao quarto de seus pais, acendeu a luz e ficou observando seu pai e sua mãe dormirem.

O senhor Jacó acordou e perguntou ao filho:

- O que aconteceu Júlio?

Júlio em silêncio só balançou a cabeça em sinal negativo, respondendo ao pai que nada havia ocorrido.

Daí, o senhor Jacó irritado, perguntou ao Júlio:

- Então o que foi moleque?

Júlio continuou em silêncio. Jacó já muito irritado berrou com Júlio:

- Então vai dormir seu doente!

Júlio apagou a luz do quarto, se dirigiu ao seu quarto e se deitou.

Na manhã seguinte todos se levantaram cedo, o senhor Jacó iria trabalhar, a dona Joana levaria as crianças para a escola e Ricardo e Júlio iriam à escola...

Mas Júlio não se levantou.

Então o senhor Jacó, que já estava muito irritado com Júlio, entra bufando no quarto do garoto e grita:

- Levanta seu moleque vagabundo!

Júlio nem se mexeu.

Então, Jacó avança sobre o garoto, puxa com força o cobertor do menino com o braço direito levantado, pronto para lhe dar um tapa, quando percebe que Júlio estava com os olhos fechados e que estava pálido. Jacó, assustado, colocou a mão sobre o rosto de Júlio e pôde notar que seu filho estava gelado.

Desesperado, Jacó gritou chamando a esposa e o filho Ricardo para ver o que havia acontecido com Júlio...

Infelizmente o pior. Júlio estava morto e sem qualquer motivo aparente. Dona Joana, desesperada, abraçou o filho morto e não conseguia nem respirar de tanto chorar. Ricardo, desconsolado, segurou firme a mão do irmão e só tinha forças para chorar também.

Jacó, em desespero, soluçando, e com os olhos cheios de lágrimas, percebeu que havia um papelzinho dobrado nas pequenas mãos de Júlio.

Jacó então pegou o pequeno pedaço de papel e havia algo escrito com a letra de Júlio:

“Outra noite, Deus veio falar comigo através de um sonho, disse a mim que apesar de amar minha família e de ela me amar, teríamos que nos separar.

Eu não queria isso, mas Deus me explicou que seria necessário.

Não sei o que vai acontecer, mas estou com muito medo.

Gostaria que ficasse claro apenas uma coisa:

- Ricardo, não se envergonhe de amar seu irmão.
- Mamãe, a senhora é a melhor mãe do mundo.
- Papai, o senhor de tanto trabalhar se esqueceu de viver.
- Eu amo todos vocês!”

Quantas vezes não temos tempo para parar e amar, e receber o amor que nos é ofertado?

Talvez quando acordarmos possa ser tarde demais... mas, ainda há tempo!

C.7 Sobre a gloriosa noite em que eu apitei o Fla-Flu

Amigos, a sensação é maravilhosa. Ainda no vestiário da arbitragem, eu ouvia o urro dos 120 mil fanáticos nas arquibancadas, gerais e cadeiras. De um lado, vinha um compacto e sonoro “Meeeeengooooooooo”. Do outro, vinha um constante e firme “Neeeeenseeeeeee”. Quando terminei de me vestir, olhei-me no espelho, e me vi dentro daquele bonito uniforme preto. Verifiquei meu bolso: lá estavam o apito, a caneta, o cartão amarelo e o cartão vermelho. Olhei o relógio: era a hora de subir para o gramado.

Assim que adentrei o campo, um pequeno povo na geral me direcionou palavras nada carinhosas. Em resposta, apenas sorri. Verifiquei as redes das duas balizas, e me dirigi ao centro do gramado, onde aguardaria a entrada das duas equipes. Primeiro veio o Flamengo, acompanhado de um urro colossal e de uma tempestade de bandeiras rubro-negras à esquerda das cabines. Depois veio o Fluminense, acompanhado de um canto de eterno amor, e de uma espessa nuvem de pó-de-arroz à direita das cabines. Que espetáculo bonito é o Fla-Flu, sem paralelo no mundo. Começa o jogo, e o Flamengo está melhor, os jogadores rubro-negros correm mais. O atacante mais perigoso parte para cima, tenta o drible da vaca, e é derrubado pelo zagueiro do Fluminense. Eu estava próximo do lance e, claro, não vi falta alguma. Mando o sujeito levantar, ameaçando eu mesmo erguê-lo pelas trancinhas do cabelo. Mas, no ataque seguinte, ele sofre nova falta, e de novo vai ao chão. Corro para cima dele, faço com os braços um gesto de “acabou”, e aplico-lhe o cartão amarelo pela simulação. “Se continuar reclamando, boto pra fora!”, esbravejo.

O Fluminense está distraído, e perde uma bola boba no meio-campo. Três apitadas curtas pra parar o jogo: falta a favor do Tricolor. Na cobrança, um passe errado, no pé de um defensor rubro-negro. “Hoje está difícil”, penso comigo mesmo. Graças a João de Deus, o ataque do Flamengo pára nas mãos do goleiro do Fluminense. E finalmente começa um bom ataque tricolor. Após a bela triangulação, entretanto, o

meia argentino escorrega na entrada da área. Apitei no reflexo, falta. Os tricolores pedem cartão, e eu aplico o amarelo no zagueirão, que jura não ter encostado no hermano. A falta é perigosíssima. Posiciono a barreira muito além dos 9 e 15, até porque sei que eles vão andar (todo jogador de futebol é cínico). A cobrança do argentino é irretocável: a pelota morre no ângulo. Fluminense 1 a 0, e eu corro para o centro do gramado.

Um volante rubro-negro ainda está reclamando da falta. O descarado afirma que o hermano escorregou sozinho, e insiste na idéia. Não tenho alternativa: amarelo nele. O primeiro tempo segue disputado, com lealdade. Marco apenas umas poucas faltas, quase todas para o Fluminense. Como batem os jogadores rubro-negros! Deve ser ordem daquele treinador interino. Olho o relógio e vejo 45 minutos cronometrados. Nada de acréscimo, fim da primeira etapa no Maracanã.

No intervalo, a atmosfera no vestiário é outra, silenciosa. As duas torcidas estão naquele estado de tensão máxima, parecem poupar as vozes para o segundo tempo. O bandeirinha vem falar comigo: “parabéns, poucos viriam aquela falta”. Agradeço assentindo com a cabeça. Olho meu cartão amarelo com três nomes escritos, todos do Flamengo. Penso de novo na violência exagerada do rubro-negro: desse jeito vão terminar com dez.

Tão logo apitei o início do segundo tempo, o Flamengo foi para cima do Fluminense como uma avalanche que desce rasgando uma montanha andina. É pressão total rubro-negra. Bola alçada na área, típico chuveirinho, e aí... ihhh, cortou com a mão? Juro que não vi. Os atletas do Flamengo me cercam. A torcida rubro-negra desrespeita até a minha décima-quinta geração materna. Não solto uma palavra, apenas faço o movimento de negação com a cabeça. Sobrou um cartão amarelo por reclamação. Quando vou anotar o número, vejo que é o outro zagueiro. “Isso não vai acabar bem”, penso de mim para mim mesmo.

O jogo segue ríspido, mas quatro jogadores do Fla sempre tiram o pé nas divididas. Coincidência, são os quatro que têm cartão amarelo. Será que é medo de ser expulso? A bola rola, ataque do Mengão pela direita, e opa... precisava disso? Minha nossa, que tesoura, meu lateral-esquerdo! Sou obrigado a advertir o primeiro atleta tricolor com o cartão amarelo. A massa rubro-negra vaia, queria a expulsão. Quanto exagero...

O Flamengo começa a sentir o cansaço, parece até que está jogando com um a menos. O Tricolor toca a bola inteligentemente, faz o tempo passar. Uma bola dividida sai pela linha lateral, e o atleta rubro-negro, apressado, já a pega pra fazer o arremesso. Apito e chego berrando: “Não, não, é bola do Fluminense!”. O abusado jogador atira a bola no chão com força, com raiva, balbuciando uns palavrões: amarelo nele.

O treinador interino do Flamengo resolve que é hora de partir para o tudo ou

nada, tira um dos zagueiros pendurados e põe um atacante. A área tricolor fica mais congestionada que a rua Pinheiro Machado. Os flamenguistas reclamam de dois pênaltis, mas o agarra-agarra existe, o juiz não pode marcar, senão haveria 257 penalidades por partida. Os contra-ataques tricolores são perigosos, esses moleques correm uma barbaridade. Mas como perdem gols... foram umas quatro chances claras ao longo do segundo tempo.

A segunda etapa até que passou rápido: o cronômetro já indica 44. Sinalizo para o quarto árbitro: dois minutos de acréscimo. Mais vaia da torcida do Flamengo. Penso: “poderia acrescentar mais meia-hora que esse time não vai empatar”.

O rubro-negro resolve utilizar o chuveirinho novamente, para o último ataque. O goleiro tricolor sai atrasado, ai meu Deus!, o atacante vai chegar antes na bola, que desespero!, a bola espirra, e alguém chuta para as redes. Um ensurdecidor urro de gol ecoa pelo cimento sagrado do Maracanã. Quase ninguém me ouviu apitando a falta do ataque.

Pouco depois, pedi a bola, e apitei o final do jogo no Maracanã. Jogadores de ambos os times me cumprimentam. Olho para as arquibancadas, e vejo as lindíssimas bandeiras tricolores ao vento, festejando, celebrando mais uma épica vitória do Fluminense. Não chorei por pouco, por muito pouco.

Mas gostei mesmo foi das resenhas da imprensa no dia seguinte. Pouquíssimas citaram a arbitragem, bom sinal, bom sinal! Só mesmo o Denilson, do Globo, é que esculhambou minha atuação. Mas faz parte, todo mundo sabe que ele é um torcedor fanático do Flamengo, que não consegue ser imparcial. Ele evidentemente colocaria a culpa pela derrota rubro-negra na arbitragem.

Mas voltemos à gloriosa noite em que eu apitei o Fla-Flu. Quando cheguei em casa e desabei na cama, a cabeça encostada no travesseiro, um sentimento me invadiu a alma. Tenho certeza, era a sensação do dever cumprido.