



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de São José do Rio Preto

Vanessa Marquiefável Serrani

Ambiente Web de Suporte à Transcrição Fonética Automática de Lemas
em Verbetes de Dicionários do Português do Brasil

São José do Rio Preto
2015

Vanessa Marquifável Serrani

Ambiente Web de Suporte à Transcrição Fonética Automática de Lemas
em Verbetes de Dicionários do Português do Brasil

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Orientador: Prof^a. Dr^a. Claudia Zavaglia

Coorientador: Prof. Dr. Luiz Carlos Cagliariari

São José do Rio Preto
2015

Marquifável, V. S.

Ambiente web de suporte à transcrição fonética automática de lemas em verbetes de dicionários do português do Brasil / Vanessa Marquifável Serrani. -- São José do Rio Preto, 2015
202 f. : il., tabs.

Orientador: Claudia Zavaglia

Coorientador: Luiz Carlos Cagliari

Tese (doutorado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Linguística aplicada. 2. Língua portuguesa - Brasil - Lexicografia.
3. Língua portuguesa - Fonética. 4. Linguagens e línguas - Transcrições fonéticas. 5. Linguística - Processamento de dados. I. Serrani, Vanessa Marquifável. II. Zavaglia, Claudia. III. Cagliari, Luiz Carlos.
IV. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. V. Título.

CDU – 469.0-13

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Vanessa Marquifável Serrani

Ambiente Web de Suporte à Transcrição Fonética Automática de Lemas
em Verbetes de Dicionários do Português do Brasil

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Comissão Examinadora

Prof^a. Dr^a. Cláudia Zavaglia
UNESP – São José do Rio Preto
Orientador

Prof^a. Dr^a. Vivian Regina Orsi Galdino de Souza
UNESP – São José do Rio Preto

Prof^a. Dr^a. Maria Cristina Parreira da Silva
UNESP – São José do Rio Preto

Prof. Dr. Pablo Arantes
UFSCar – São Carlos

Prof^a. Dr^a. Gladis Maria de Barcellos Almeida
UFSCar – São Carlos

São José do Rio Preto
25 de março de 2015

A Deus e à minha família, pois sem eles não somos nada.

AGRADECIMENTOS

Considerando esta tese como resultado de uma caminhada que não começou na UNESP, agradecer pode não ser tarefa fácil, nem justa. Para não correr o risco da injustiça, agradeço de antemão a todos que de alguma forma passaram pela minha vida e contribuíram para a construção de quem sou hoje.

Agradeço, particularmente, a algumas pessoas pela contribuição direta na construção deste trabalho:

À Claudia, minha orientadora, pela disponibilidade e profissionalismo e, sobretudo, pela confiança depositada em mim e em meu trabalho.

Ao prof. Cagliari pela confiança e incentivo.

Ao IBILCE, em especial, ao Programa de Pós-Graduação em Estudos Linguísticos, a todos os professores que contribuíram com a minha formação e aos funcionários, pela atenção e profissionalismo.

A todos os integrantes do NILC, por ser uma casa e uma escola; o lugar onde se estuda e aprende a trabalhar partilhando generosamente.

Aos professores que fizeram parte da minha banca de defesa, os professores Pablo, Gladis, Maria Cristina e Vivian, pelos comentários e sugestões feitos em relação ao meu trabalho.

Aos pesquisadores Fernando Perdigão e Sara Candeias pela oportunidade de estágio de pesquisa no Instituto de Tecnologia da Universidade de Coimbra.

À FAPESP pelas bolsas de doutorado e de estágio no exterior concedidas.

Aos meus pais, Wilmar e Maria José, e à minha irmã, Flávia, pelo amor, incentivo e alegrias.

Ao meu esposo Tiago, pela constante companhia, carinho e apoio.

A Deus, sem o qual nada disso teria sido possível.

Este trabalho foi financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo, através da bolsa de doutorado (Processo nº 2011/01299-4) e da bolsa estágio de pesquisa no exterior (Processo nº 2014/00613-5).

THAT'S WRONG — ISN'T IT
' γ BEFORE β EXCEPT AFTER δ '?



: δ β γ
 γ β δ

© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com

Bernie

RESUMO

A transcrição fonética automática é uma tarefa essencial para muitas aplicações de diferentes áreas. Além de sistemas de síntese de fala que necessitam de transcrições de alta qualidade para obter um bom desempenho, sistemas de transcrição fonética automática podem ser utilizados em muitas outras áreas (fonética, fonologia, ensino-aprendizagem de língua, e assim por diante) com o objetivo de se obter transcrições preliminares de grandes corpuses. Esse tipo de sistema frequentemente gera uma transcrição fonética a partir da forma ortográfica. Considerando o esforço despendido por lexicógrafos no desenvolvimento de transcrições fonéticas manuais para obras lexicográficas, este estudo teve como objetivo principal a criação de um ambiente computacional para transcrever foneticamente unidades lexicais que fazem parte da cabeça do verbete de dicionários mono ou multilíngues, do Português Brasileiro, com disponibilização gratuita, via *Web*, denominado PETRUS – *Phonetic Transcriber for User Support*. O propósito deste trabalho é contribuir, mesmo que modestamente, para que lexicógrafos, nosso principal público-alvo, possam desfrutar dos benefícios da automatização da transcrição fonética, diminuindo o tempo e as dificuldades de inserção desse tipo de informação em suas obras lexicográficas. Os resultados preliminares indicam uma taxa de acerto de 97.5% ao fone, com possibilidade de aumento.

Palavras-chave: Processo de conversão grafema-fone, Abordagem baseada em regras linguísticas, Português do Brasil, Lexicografia.

ABSTRACT

Automatic Phonetic Transcription is a crucial task for many applications of different areas. Besides Text-to-Speech systems, which need high quality transcriptions in order to have a good performance, automatic phonetic transcription tools can be useful in many areas (phonetics, phonology, language learning, and so on) in order to obtain preliminary transcriptions of large corpora. This kind of system often generate a phonetic transcription directly from the orthographic form. Considering the effort expended by lexicographers in the development of manual phonetic transcriptions to be added in their dictionaries, this study aimed the creation of a computational environment to support automatic phonetic transcription of lemmas, from mono or multilingual Brazilian dictionaries entries, to be available free of charge, via the Web, called PETRUS – PhonEtic TRanscriber for User Support. We intend to contribute, even modestly, so that lexicographers, our main target audience, can enjoy the benefits of automating the phonetic transcription, reducing the time and difficulty in integrating this kind of information in their lexicographical works. Preliminary results indicate a phone error rate accuracy of 97.5%, with the possibility of increase.

Keywords: Grapheme-phone conversion process, Rule-based approach, Brazilian Portuguese, Lexicography.

LISTA DE TABELAS

TABELA 1. AS SETE VOGAIS TÔNICAS ORAIS EXISTENTES NO PB	50
TABELA 2. AS VOGAIS PRETÔNICAS ORAIS EXISTENTES NO PB	51
TABELA 3. AS VOGAIS POSTÔNICAS NÃO-FINAIS EXISTENTES NO PB	53
TABELA 4. AS VOGAIS POSTÔNICAS ORAIS FINAIS EXISTENTES NO PB	55
TABELA 5. DITONGOS ORAIS DECRESCENTES DO PB	59
TABELA 6. DITONGOS ORAIS DECRESCENTES DERIVADOS DO PROCESSO DE VOCALIZAÇÃO DO FONEMA /L/.....	60
TABELA 7. DITONGOS NASAIS DECRESCENTES DO PB	60
TABELA 8. TRITONGOS ORAIS DO PB	61
TABELA 9. PARES DE HOMÓGRAFOS HETERÓFONOS COLETADOS	114
TABELA 10. PARES DE HOMÓGRAFOS HETERÓFONOS TRANSCRITOS FONETICAMENTE.....	115
TABELA 11. MORFEMAS LOCALIZADOS À ESQUERDA DA BASE CONSTITUÍDOS PELAS VOGAIS MÉDIAS <E> OU <O>, TRANSCRITOS PELO PETRUS	118
TABELA 12. COMPARAÇÃO DO DESEMPENHO DE SISTEMAS DE DIVISÃO SILÁBICA AUTOMÁTICA PARA O PB	129
TABELA 13. REGRAS DE TRANSCRIÇÃO FONÉTICA QUE CONSIDERAM INFORMAÇÃO DE TONICIDADE E SILÁBICA.	132
TABELA 14. REGRAS DE TRANSCRIÇÃO PARA OS GRAFEMAS DO PB	137
TABELA 15. PADRÕES DE ALTERAÇÃO DE PRONÚNCIA CONFORME O TEMPO, MODO E PESSOAS VERBAIS, RELEVANTES PARA A FLEXÃO DOS VERBOS EM PB	151
TABELA 16. VERBOS PARADIGMAS PARA O PB EM ORDEM CRESCENTE DE IRREGULARIDADE.	154
TABELA 17. EXEMPLOS DE REGRAS UTILIZADAS NO SISTEMA DE FLEXÃO E TRANSCRIÇÃO DE VERBOS PROPOSTO.	157
TABELA 18. TAXAS DE ACERTO OBTIDAS NA AVALIAÇÃO DO DESEMPENHO DO SISTEMA PETRUS	161
TABELA 19. COMPARAÇÃO ENTRE OS SISTEMAS DE CONVERSÃO BASEADOS EM REGRAS DO PB.....	162

LISTA DE FIGURAS

FIGURA 1. EXEMPLO DE VERBETE-MODELO. FONTE: ZAVAGLIA (2007, p. 213).	25
FIGURA 2. PINTURAS RUPESTRES.	33
FIGURA 3. SISTEMA IDEOGRÁFICO CHINÊS.	34
FIGURA 4. ESCRITA KATAKANA.	34
FIGURA 5. RIGVEDA MANUSCRITO EM DEVANÁGARI.	35
FIGURA 6. TRECHO DA BÍBLIA HEBRAICA DO SÉCULO 10.	35
FIGURA 7. ESQUEMA DO ESFORÇO MUSCULAR E CURVA DE FORÇA SILÁBICA.	78
FIGURA 8. MODELO LINEAR DE SÍLABA (CÂMARA JR., 1970).....	78
FIGURA 9. MODELO HIERÁRQUICO DE SÍLABA (SELKIRK, 1982).....	79
FIGURA 10. DIAGRAMA DO PROCESSO DE TRANSCRIÇÃO FONÉTICA AUTOMÁTICA DO SISTEMA PETRUS	109
FIGURA 11. PRIMEIRA ETAPA DE TRANSCRIÇÃO FONÉTICA DO SISTEMA PETRUS.	112
FIGURA 12. TELA DO PETRUS COM A TRANSCRIÇÃO FONÉTICA DA PALAVRA <GOSTO>.	115
FIGURA 13. SEGUNDA ETAPA DE TRANSCRIÇÃO FONÉTICA NO SISTEMA PETRUS.	116
FIGURA 14. ATRIBUIÇÃO DA CLASSE GRAMATICAL ÀS PALAVRAS DE ENTRADA DO SISTEMA PETRUS.....	121
FIGURA 15. TELA DO PETRUS COM A ANÁLISE DA CLASSE GRAMATICAL DA PALAVRA <CASA>.....	123
FIGURA 16. REALIZAÇÃO DA MARCAÇÃO AUTOMÁTICA DA VOGAL TÔNICA.....	124
FIGURA 17. TELA DO DICIONÁRIO CALDAS AULETE COM INDICAÇÃO DA SÍLABA TÔNICA EM ITÁLICO.	125
FIGURA 18. DIVISÃO SILÁBICA AUTOMÁTICA DAS PALAVRAS DE ENTRADA DO PETRUS.	127
FIGURA 19. ETAPA DE TRANSCRIÇÃO FONÉTICA AUTOMÁTICA DO PETRUS.	135

SÍMBOLOS

[p]	oclusiva bilabial surda
[b]	oclusiva bilabial sonora
[t]	oclusiva dental surda
[d]	oclusiva dental sonora
[k]	oclusiva velar surda
[g]	oclusiva velar sonora
[tʃ]	africada palatoalveolar surda
[dʒ]	africada palatoalveolar sonora
[f]	fricativa labiodental surda
[v]	fricativa labiodental sonora
[s]	fricativa alveolar surda
[z]	fricativa alveolar sonora
[ʃ]	fricativa palatal surda
[ʒ]	fricativa palatal sonora
[x]	fricativa velar surda
[ɣ]	fricativa velar sonora
[χ]	fricativa uvular surda
[ʁ]	fricativa uvular sonora
[h]	fricativa glotal surda
[ɦ]	fricativa glotal sonora
[m]	nasal bilabial sonora
[n]	nasal alveolar sonora
[ɲ]	nasal palatal sonora
[ŋ]	nasal velar sonora
[l]	lateral alveolar sonora
[ʎ]	lateral aproximante palatal
[r]	vibrante alveolar sonora
[ɾ]	tepe alveolar sonora
[ɹ]	aproximante alveolar sonora
[w]	aproximante velar sonora
[j]	aproximante palatal sonora
[^m]	aproximante nasal bilabial sonora
[^ɲ]	aproximante nasal palatal sonora
[^ŋ]	aproximante nasal velar sonora
[u]	vogal posterior alta fechada arredondada
[ũ]	vogal posterior alta fechada arredondada nasalizada
[o]	vogal posterior alta aberta arredondada
[i]	vogal anterior alta fechada não arredondada
[ĩ]	vogal anterior alta fechada não arredondada nasalizada
[ɪ]	vogal anterior aberta não arredondada
[o]	vogal posterior média fechada arredondada

- [õ] vogal posterior média alta arredondada nasalizada
- [ɔ] vogal posterior média aberta arredondada
- [ɛ] vogal anterior média aberta não arredondada
- [e] vogal anterior média fechada não arredondada
- [ɐ] vogal central baixa fechada não arredondada
- [ẽ] vogal central baixa fechada não arredondada nasalizada
- [a] vogal central baixa aberta não arredondada

SIGLAS

AFI	Alfabeto Fonético Internacional
ASCII	American Standard Code for Information Interchange
ALIP	Amostra Linguística do Interior Paulista
CARTs	Classification and Regression Trees
CMU	Carnegie Mellon University
Celpe-Bras	Certificado de Proficiência em Língua Portuguesa para Estrangeiros
C	Consoante
CPLP	Comunidade dos Países de Língua Portuguesa
G	Glide
G2P	Grafema-fone
HH	Homógrafos Heterófonos
IPA	International Phonetic Alphabet
LAFAPE	Laboratório de Fonética e Psicolinguística da UNICAMP
LAPS	Laboratório de Processamento de Sinais
MBL	Memory-Based Learning
NILC	Núcleo Interinstitucional de Linguística Computacional
PB	Português Brasileiro
PE	Português Europeu
PLN	Processamento de Linguagem Natural
SAMPA	Speech Assessment Methods Phonetic Alphabet
X-SAMPA	Speech Assessment Methods Phonetic Alphabet Extended
TBL	Transformation-Based Learning
UFPA	Universidade Federal do Pará
UL	Unidade Lexical
UNICAMP	Universidade Estadual de Campinas
V	Vogal

OUTRAS CONVENÇÕES

- < > representação ortográfica
- / / representação fonêmica
- [] representação fonética

SUMÁRIO

INTRODUÇÃO.....	16
Considerações Iniciais	16
Objetivos.....	26
Metodologia.....	27
Organização do Trabalho.....	30
CAPÍTULO 1 - FUNDAMENTAÇÃO TEÓRICA E ESTADO DA ARTE EM TRANSCRIÇÃO FONÉTICA	32
1.1 Fundamentação Teórica	32
1.1.1 <i>Relação grafia e som</i>	32
1.1.2 <i>Ortografia</i>	40
1.1.3 <i>Fonética</i>	43
1.1.3.1 Consoantes.....	45
1.1.3.2 Vogais	48
1.1.3.2.1 Vogais Orais	50
1.1.3.2.2 Vogais Nasais	55
1.1.3.2.3 Glides.....	57
1.1.3.2.4 Encontros Vocálicos	58
1.1.3.2.5 Casos de variação fonética não contemplados	62
1.1.4 <i>Fonologia</i>	64
1.1.4.1 Consoantes.....	66
1.1.4.2 Vogais Orais	68
1.1.4.3 Vogais Nasais	70
1.1.4.4 Encontros Vocálicos	71
1.1.4.5 Glides.....	72
1.1.4.6 Variação Alofônica.....	73
1.1.4.7 Processos Fonológicos.....	74
1.1.5 <i>Prosódia</i>	76
1.1.5.1 Sílabas	77
1.1.5.2 Acento.....	82
1.1.5.3 Clítico	85
1.2 Transcrição Fonética.....	87

1.3	Transcrição manual, automática e semiautomática.....	92
1.4	Estado da arte em transcrição fonética automática.....	93
1.4.1	<i>Sistemas baseados em dicionários.....</i>	94
1.4.2	<i>Sistemas baseados em regras.....</i>	96
1.4.3	<i>Sistemas baseados em dados.....</i>	99
1.4.4	<i>Sistemas híbridos.....</i>	104
1.4.5	<i>Por que um novo sistema de transcrição?.....</i>	104
CAPÍTULO 2 - METODOLOGIA.....		108
2.1	Diagrama do processo de transcrição do PETRUS.....	109
2.1.1	<i>Etapa 1 - Transcrição fonética de homógrafos heterófonos.....</i>	111
2.1.2	<i>Etapa 2 - Transcrição fonética de morfemas à esquerda da base.....</i>	116
2.1.3	<i>Etapa 3 - Identificação automática de categorias gramaticais.....</i>	120
2.1.4	<i>Etapa 4 - Identificação automática da sílaba tônica.....</i>	124
2.1.5	<i>Etapa 5 - Divisão silábica automática.....</i>	127
2.1.6	<i>Etapa 6 - Conversão grafema-fone.....</i>	132
2.1.6.1	Conjunto de regras linguísticas para transcrição fonética.....	136
2.1.6.2	Módulo para flexão e transcrição fonética automática de verbos do PB.....	144
2.1.6.2.1	O sistema verbal do português: características morfofonológicas.....	144
2.1.6.2.2	Metodologia.....	147
CAPÍTULO 3 - RESULTADOS.....		160
CAPÍTULO 4 - CONCLUSÕES E ENCAMINHAMENTOS FUTUROS.....		164
4.1	Considerações sobre o trabalho desenvolvido.....	164
4.2	Encaminhamentos Futuros.....	166
REFERÊNCIAS.....		170
ANEXO A – Alfabeto Fonético Internacional.....		190
ANEXO B – Exemplo de saída produzida com o sistema de flexão e transcrição verbal.....		191
ANEXO C – Erros obtidos no teste.....		193

INTRODUÇÃO

CONSIDERAÇÕES INICIAIS

O termo Lexicografia deriva das palavras gregas *lexikon* (léxico) e *graphein* (descrição) e se ocupa, portanto, da descrição do conjunto de unidades lexicais¹ que formam uma língua natural (léxico), a fim de elaborar dicionários e bases de dados lexicológicas. Em outras palavras, pode ser definida como a ciência que repertoria e organiza as unidades lexicais em uma obra lexicográfica.

Em oposição a essa Lexicografia prática temos a Lexicografia teórica, ou metalexigrafia, responsável pelo estudo dos princípios teóricos e metodológicos relativos à elaboração de dicionários. Para tanto, são realizados estudos que regulam a estrutura e o comportamento linguísticos na medida em que orientam e condicionam o trabalho do lexicógrafo.

Nesse sentido, Villar e Borba (2011, p. 17) ressaltam as habilidades tanto linguísticas quanto técnicas exigidas de um lexicógrafo. Segundo eles, além do conhecimento das propriedades inerentes aos níveis de organização linguística (fonético, fonológico, ortográfico, morfológico, sintático, semântico e pragmático), o lexicógrafo deve conhecer a técnica lexicográfica, ou seja, dominar os princípios e praxes estabelecidos ao longo do tempo pela prática lexicográfica, incluindo também, o domínio sobre as atuais técnicas computacionais de compilação e investigação de *cópus*.² Os autores apontam como uma das

¹ Unidades lexicais: segundo Biderman (1984, p. 144) é a unidade vocabular tanto com respeito à significação como com respeito à forma gramatical, que tem uso característico no discurso. Sinônimo de *lexema*, *lexia* e *palavra*, conforme o contexto. Neste trabalho, utilizamos como sinônimo de *palavra*, conscientes da problemática existente sobre sua definição.

² Neste trabalho, utilizamos a grafia “o *cópus*” e “os *cópus*” da palavra *corpus* (derivada do latim), com o intuito de fazer seu uso adaptado à morfologia portuguesa, seguindo o paradigma de “o *lápis*” e “os *lápís*”.

tarefas fundamentais desempenhadas atualmente via *córpus*, a obtenção de dados frequenciais sobre usos e acepções de itens lexicais, a fim de constituir a macro e a microestrutura de um dicionário.

Mas, o que é o dicionário? Hartmann e James (1998, p. 41) o definem como sendo um tipo de trabalho de referência que apresenta o vocabulário de uma língua em ordem alfabética, geralmente contendo explicação dos significados. Para Zavaglia (2011), o dicionário é uma obra lexicográfica composta por um repertório de unidades lexicais que se apresenta organizado e classificado conforme critérios e princípios relativos ao seu propósito de constituição e/ou uso. Dessa maneira, a autora explica que a convenção da ordem das unidades léxicas pode obedecer tanto a um critério semasiológico quanto onomasiológico e que as informações contidas nos verbetes podem ser de caráter morfossintático, valencial, fonético, semântico, pragmático, discursivo, entre outros. Em relação à dimensão da obra lexicográfica, a autora afirma com base em Biderman (1984, p. 27) que deverá ser definida, a rigor, conforme sua tipologia (tesouro - mais de 100.000 entradas, padrão - cerca de 50.000 entradas, médio ou escolar - de 10.000 a 30.000 entradas, básico - cerca de 5.000 verbetes); já no tocante a sua direção, diz que um dicionário pode ser uni, bi ou multilíngue; quanto aos exemplos relata que podem ser autênticos, abonados ou forjados; quanto à nomenclatura, assegura que poderá ser baseada em *córpus* ou não; e que seu suporte poderá ser *on-line*, eletrônico ou impresso, dentre outras possibilidades.

O dicionário é um objeto que desempenha papel importante tanto em nossa vida quanto na sociedade da qual fazemos parte. Os falantes de uma língua utilizam um dicionário, em geral, para o esclarecimento de dúvidas - mas também por mera curiosidade - sobre o uso de uma palavra, de um termo desconhecido ou já caído em desuso na língua (materna ou estrangeira), por exemplo, além de outros. O dicionário também se mostra uma ferramenta de trabalho interessante para tradutores, bem como um instrumento facilitador de comunicação entre povos, principalmente quando bilíngue.

Além do nível individual, o dicionário desempenha um papel importante na sociedade. Ao lado das gramáticas, os dicionários constituem os meios de preservar, desenvolver e disseminar qualquer língua. Quando uma língua se torna oficial, procura-se imediatamente que ela passe a dispor de um dicionário geral monolíngue que descreva o seu léxico essencial. Com efeito,

(...) busca registrar e definir os signos lexicais que se referem a conceitos elaborados e cristalizados na cultura da língua em questão. Converte-se em um instrumento de grande importância nas sociedades contemporâneas, chega a representar uma autoridade nas questões de língua e confere prestígio àqueles que o possuem, podendo até mesmo ser considerado uma obra cultural (ZAVAGLIA e

SILVA, 2006, p. 69).

O léxico se constitui enquanto conjunto aberto, em contínua expansão, de todos os itens lexicais existentes em uma língua natural, incluindo expressões, fraseologismos e itens gramaticais. Em geral, podemos observar que os dicionários podem registrar uma parcela maior ou menor do léxico (muitos tipos lexicais) ou apenas um tipo específico (verbos, homônimos, etc.); podem ser semasiológicos ou onomasiológicos; podem ser monolíngues (uma só língua), bilíngues (duas línguas), trilingues (três línguas) ou multilíngues (várias línguas), bem como descritivos (registrando como os itens lexicais são usados na realidade) ou prescritivos (determinando de que maneira palavras e expressões deveriam ser empregadas).

Para Iriarte-Sanromán (2000), conhecer uma unidade lexical (UL) implica principalmente: conhecer a sua forma oral e escrita (pronúncia e ortografia); conhecer suas variações formais (como gênero, número, grau) assim como a capacidade para formar novas unidades (derivação, composição, entre outros); conhecer as suas capacidades combinatórias, as relações possíveis com outras unidades no contexto linguístico (combinações livres, colocações, frases, por exemplo); conhecer os seus significados ou acepções, assim como as possibilidades de uso pragmático-contextual e retórico (contexto de uso, registros, domínios, usos metafóricos, variantes diacrônicas, geográficas, e assim por diante).

O verbete é a unidade estrutural de uma obra lexicográfica que transmite o conhecimento a respeito de uma UL. O verbete compreende uma palavra-entrada lematizada e um enunciado lexicográfico, que se caracteriza como um texto escrito, de caráter informativo, destinado a explicar um conceito conforme padrões descritivos sistemáticos, determinados pela obra de referência da qual faz parte: mais comumente, um dicionário ou uma enciclopédia. Essencialmente, destina-se à consulta, o que lhe impõe uma construção discursiva sucinta e de acesso fácil e imediato, embora isso não signifique necessariamente uma curta extensão. Portanto, além da definição da palavra, o verbete de um dicionário pode oferecer ainda variadas informações, de acordo com seu propósito, que podem ser compostas por um paradigma informacional que carrega informações a respeito de gênero, número, classe gramatical, transcrição fonética ou mesmo etimologia da entrada; um paradigma definicional constituído da definição da entrada; um paradigma pragmático que representa a entrada em funcionamento, em contexto de uso, também denominado abonação e outros paradigmas informacionais que complementam a cadeia de informação como a rede de

remissivas (sinônimos ou antônimos) e aspectos de natureza enciclopédica entre outras informações, que especificam a palavra-entrada. (Cf. BARBOSA, 2001).

Dessa maneira, cada um dos elementos supracitados contribui para a compreensão da palavra-entrada em sua totalidade. Entre eles, a pronúncia, que constitui um dos aspectos normativos relevantes a qualquer obra lexicográfica bilíngue, mas não somente, e nem sempre presente, é o tópico que suscita interesse especial nesta pesquisa.

Nos dicionários gerais de língua, a informação sonora das unidades lexicais pode figurar de forma sistemática em todas as entradas ou apenas naquelas que apresentam pronúncias peculiares. Para o Português Europeu (PE) foram produzidos dois dicionários que apresentam transcrição fonética para todas as suas entradas: o Casteleiro (2001) e o Teixeira, Da Costa e Da Silva (2004). De fato, a pronúncia apresenta-se sob a forma de uma transcrição fonética associada ao lema (palavra-entrada), com o uso de símbolos fonéticos do Alfabeto Fonético Internacional, apresentados entre parênteses retos, consoante as convenções adotadas e descritas na parte introdutória dessas obras.

Em contrapartida, dicionários direcionados a falantes nativos do Português Brasileiro (PB), tais como, o Ferreira, Anjos e Ferreira (2000) e o Houaiss, Villar e Franco (2004), não incluem transcrição fonética em todas as cabeças de verbete³ (CAGLIARI, 2011 p. 2). Os dicionários elaborados para o PB apresentam, por vezes, indicação, a seguir do lema, sobre a realização de algum segmento (vocálico ou consonântico) problemático (ex. *camiseta* (ê) ou *camisete* (é)) e para os casos de estrangeirismos, que mantêm inalterada a ortografia da língua fonte (ex. *e-mail*). Ou seja, nos referidos dicionários, admite-se que pelo fato de o PB ter uma correspondência, de certa forma regular, entre sua grafia e respectiva pronúncia, recorre-se a um sistema de notação figurada por meio do qual se representa a pronúncia da palavra (estrangeira) ou parte dela (não estrangeiras), transcrevendo apenas o(s) segmento(s) cuja pronúncia não é previsível ou transparente para o usuário. De maneira geral, em PB, tais opacidades entre grafia e som dizem respeito ao timbre das vogais, a tonicidade das sílabas e a relação arbitrária entre letra consonantal/som consonantal.

Entretanto, tais transcrições nem sempre são marcadas da mesma maneira nas diferentes obras lexicográficas do PB e, por vezes, um mesmo autor pode alterar a maneira como realiza a representação fonética de uma UL (parcial ou total) assistematicamente, conforme comprovou Pontes (2007), por exemplo, ao analisar dicionários escolares brasileiros. Nadin

³ Cabeça de verbete: De acordo com Welker (2004, p. 110), compreende o lema (que é a entrada de verbete, palavra-entrada, ou simplesmente entrada) e as informações anteriores à definição ou às definições (ou equivalentes, nos dicionários bilíngues), a saber: as variantes ortográficas, a pronúncia, a categoria gramatical, as informações flexionais e/ou sintáticas, a etimologia e as marcas de uso.

(2009) analisou dicionários escolares bilíngues de língua espanhola e constatou que tais obras não apresentam uma orientação “fonético-fonológica que guie as transcrições fonéticas presentes nos verbetes” (NADIN, 2009, p. 162); e que um deles apresenta, inclusive, transcrições equivocadas quanto à marcação da sílaba tônica. Cagliari (2011) analisou as transcrições fonéticas e a variante linguística representada em dicionários monolíngues (PB) e bilíngues (português-inglês) de grande circulação no Brasil (tais como, WEISFLOG, 2000; PROCTER, 1995; LUFT, 2001; FERREIRA, 2000; HOUAISS, VILLAR e FRANCO, 2004) e constatou uma falta de sistematicidade na inclusão de informações fonéticas nessas obras. Esse mesmo autor diz ainda, em relação aos dicionários bilíngues analisados, que embora realizem transcrição fonética em todas as cabeças de verbete, a variedade linguística que foi adotada para essa tarefa não é apresentada de maneira clara porque, além de não aparecer indicada na parte introdutória do dicionário, as transcrições fonéticas contidas ao longo da obra não seguem uma variedade linguística específica do PB.

Contudo, sob a ótica pedagógica, acreditamos que seja de fundamental importância a apresentação de transcrição fonética em dicionários para aprendizes de língua (materna ou estrangeira) que têm conhecimento sobre os símbolos do Alfabeto Fonético Internacional (AFI, mais conhecido por sua sigla em inglês IPA – *International Phonetic Alphabet*). Com isso, concordamos com Welker (2004, p. 113) quando afirma que o IPA deveria ser usado em todos os dicionários e ser ensinado nos cursos de língua estrangeira. O IPA não é apenas um alfabeto, mas também um conjunto de princípios de transcrição, os quais se diferenciam conforme as necessidades da transcrição a ser realizada. Assim, a finalidade da transcrição fonética e do alfabeto fonético é justamente a transcrição e a leitura de um som em qualquer idioma por uma pessoa treinada. Assim, por meio do uso do IPA, professores poderiam promover a utilização da notação fonética em escolas com o objetivo de que as crianças possam adquirir uma pronúncia mais realista das línguas estrangeiras que aprendem, além de ser uma maneira de ensinar as crianças mais novas a ler.

Em consonância com Welker (2004, p. 113), acreditamos que o IPA não se adapte apenas aos dicionários bilíngues, mas também aos monolíngues do PB, posto que a crescente proeminência do Brasil no cenário mundial está fazendo com que cada vez mais pessoas desejem aprender a vertente brasileira do português. A esse respeito, de acordo com o Censo Demográfico de 2013, o Brasil possui cerca de 150 a 180 línguas indígenas. Esses dados permitem antever um grande número de nascidos no país, que não tem o português como língua materna e que seriam potenciais aprendizes desse idioma. Outro grupo de falantes não

nativos do português são os estrangeiros que vêm trabalhar ou estudar no Brasil. Esse grupo tem mostrado um crescimento expressivo, dados os convênios acadêmicos com instituições do exterior e da crescente globalização do mercado. Por esse motivo, centros de línguas de universidades públicas e escolas privadas de idiomas, atentos para a demanda vertiginosa desse mercado, têm criado cursos de língua portuguesa para estrangeiros, o que amplia o mercado de trabalho para especialistas da área. Além desse ensino-aprendizado em contexto de imersão, há um crescimento do número de cursos de português no exterior. O Mercosul, por exemplo, cujo tratado foi assinado em 1991, fortaleceu o ensino de língua portuguesa na América Latina. Tais cursos também são oferecidos no programa de várias universidades nos Estados Unidos da América (Yale, Princeton) e em vários países da Europa (Oxford e Cambridge (Inglaterra), Estocolmo (Suécia), Aarhus (Dinamarca), Oslo (Noruega). O português também é ensinado nos centros de línguas de demais universidades estrangeiras, em escolas privadas de idiomas, nas embaixadas brasileiras e em centros culturais luso-brasileiros espalhados pelo mundo.

A ampliação do ensino de língua portuguesa no exterior, bem como o surgimento de novos institutos (Certificado de Proficiência em Língua Portuguesa para Estrangeiros – Celpe-Bras)⁴ e de associações (Comunidade dos Países de Língua Portuguesa – CPLP)⁵, cada qual com seus propósitos específicos, contribuíram para a urgência da confecção de dicionário(s) do PB que forneça(m) a pronúncia dos vocábulos nele(s) registrados, e não apenas de palavras estrangeiras.

Segundo Zavaglia (2010, p. 81), “A indicação da pronúncia em um dicionário monolíngue de português pode ser uma ferramenta importante para estrangeiros que estejam

⁴ O Celpe-Bras é o Certificado de Proficiência em Língua Portuguesa para Estrangeiros, desenvolvido e outorgado pelo Ministério da Educação (MEC) do Brasil, aplicado no Brasil e em outros países pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) com o apoio do Ministério das Relações Exteriores (MRE). O Celpe-Bras é o único certificado de proficiência em português como língua estrangeira reconhecido oficialmente pelo governo brasileiro. É aceito internacionalmente, em empresas e instituições de ensino como comprovação de competência na língua portuguesa. No Brasil, é exigido pelas universidades para ingresso em cursos de graduação e em programas de pós-graduação, bem como para validação de diplomas de profissionais estrangeiros que pretendem trabalhar no país. O Celpe-Bras é requisito para inscrição profissional em algumas entidades de classe, a exemplo do Conselho Regional de Medicina (CRM). Entretanto o Celpe-Bras não é um diploma para interessados em dar aulas de português para falantes de outras línguas. Essa função pressupõe uma formação específica, que envolve habilidades e competências não avaliadas no Exame. Fonte: <http://download.inep.gov.br/outrasacoes/celpebras/manual/2012/manualexaminandocelpebras.pdf>.

⁵ A Comunidade dos Países de Língua Portuguesa (CPLP) é o foro multilateral privilegiado para o aprofundamento da amizade mútua e da cooperação entre os seus membros. Criada em 17 de Julho de 1996, a CPLP goza de personalidade jurídica e é dotada de autonomia financeira. A Organização tem como objectivos gerais: A concertação político-diplomática entre seus estados membros, nomeadamente para o reforço da sua presença no cenário internacional; A cooperação em todos os domínios, inclusive os da educação, saúde, ciência e tecnologia, defesa, agricultura, administração pública, comunicações, justiça, segurança pública, cultura, desporto e comunicação social; A materialização de projectos de promoção e difusão da língua portuguesa. Fonte: <http://www.cplp.org/id-2595.aspx>.

aprendendo nossa língua”. Isso porque tal indicação é imprescindível para o aprendizado da pronúncia de palavras que não fazem parte da competência do aprendiz-falante. Portanto, utilizar o AFI e suas convenções universalmente utilizadas, facilitaria o ensino-aprendizado da pronúncia do PB por aprendizes nativos ou não, uma vez que até mesmo falantes nativos do PB podem ter dúvidas sobre a pronúncia de uma determinada palavra.

Considerando, por exemplo, a forma verbal <velejo> contida na declaração do velejador Scheidt feita para a Folha de São Paulo em 23/01/2013: “Estou feliz por começar o ano com esta vitória e saber que ainda velejo dentro do mesmo nível alto de quando deixei a classe”. Como devemos pronúciá-la? Uma primeira hipótese seria pronúciá-la [ve'leʒu], com [ɛ] timbre tônico semiaberto, analogamente à pronúncia da forma verbal flexionada da primeira pessoa do singular, do Presente do Indicativo, do verbo <pregar>, por exemplo, que é [ˈpɾɛɡu]. A segunda possibilidade é a de pronúciá-la como [ve'leʒu], considerando que se trataria de um verbo regular do PB, com pronúncia semifechada de sua vogal tônica [e], assim como acontece com a forma flexionada na 1ª pessoa do singular do Presente do Indicativo do verbo <almejar>: [aw'meʒu]. Entretanto, para se ter certeza de que uma hipótese é melhor do que a outra, não contamos com a existência de um acento gráfico (agudo ou circunflexo) sobre a vogal que indique a sua pronúncia. Nem mesmo podemos encontrar a transcrição desse verbo flexionado em uma gramática ou dicionário do PB, de acordo com as obras consultadas.

De acordo com Cagliari (2011), uma possível explicação para a ausência de informação fonética em dicionários monolíngues do PB seria a existência de uma opinião generalizada de que a interpretação oral de cada letra em PB é relativamente clara, devido à correlação relativamente transparente presente entre os caracteres ortográficos e o sistema fonológico desse idioma.

Entretanto, de acordo com Braga et al. (2003), apesar de a escrita da língua portuguesa ter natureza alfabética, isto é, baseada numa relação unívoca entre um som e um grafema,⁶ na verdade, essa relação não se verifica atualmente, devido à tensão existente entre o dinamismo da evolução fonética da língua e o caráter conservador, próprio de todos os sistemas de escrita, surgindo, conseqüentemente, dificuldades de conversão de alguns grafemas, e em especial dentro do vocalismo do português. Dentre os problemas relacionados às

⁶ Morais (1995, p.76) define grafema como sendo “todos os grupos de letras que podem ser lidas como um único fonema”. Scliar-Cabral (2003, p.27) afirma que “deve-se entender o grafema como uma ou mais letras que representam um fonema (no sistema alfabético do PB, não mais que duas letras). Por exemplo, em ‘nasce’ temos cinco letras e quatro grafemas para representar /ˈnasi/. No caso, o grafema ‘sc’ é um dígrafo.” Neste trabalho, ao fazermos referência a grafemas, usamos os sinais de menor (<) e maior (>) conforme Crystal (1997, p.257).

idiossincrasias da língua (DUTOIT, 2001) podemos elencar: o fato de (i) uma sequência de caracteres poder corresponder a um único fone, como ocorre, por exemplo, com os dígrafos em ‘banha’, onde o <nh> é transcrito como apenas um fone; (ii) um caractere não corresponder a nenhum fone, como o <h> de ‘haja’; e (iii) o mesmo caractere poder ser transcrito de maneiras diferentes, de acordo com seus caracteres precedentes e subsequentes. Nas palavras ‘taxi’ e ‘exemplo’, ‘excluir’ e ‘caixa’, por exemplo, o <x> deve ser transcrito com som de [ks], [z], [s] e [ʃ] respectivamente, entre outros. Há, ainda, o caso das palavras ‘s[e]de’ e ‘s[ɛ]de’, no qual o contexto linguístico não auxilia na determinação de sua pronúncia correta.

Ademais, a transcrição fonética de cópuz, tradicionalmente realizada de maneira manual, está diretamente associada a importantes inconvenientes: (i) processo moroso que exige atenção minuciosa durante sua execução; (ii) necessidade de procedimentos padronizados e explícitos de anotação para que o formalismo utilizado possa ser replicado de forma sistemática; (iii) sujeição a erros humanos por distração ou cansaço do anotador; (iv) falta de coerência, muitas vezes, entre os anotadores que podem deixar sua subjetividade marcada durante a anotação, quando não se tem um protocolo de anotação bem definido e claro; (v) o fato de ser uma atividade que consome demasiado tempo para ser executada, até mesmo anos, dependendo do tamanho do cópuz almejado; e, por fim, (vi) o fato de ser custosa financeiramente. Somando-se a esses problemas as dimensões dos cópuz pretendidos serem cada vez maiores, tanto linguistas quanto informatas sentiram a necessidade de recorrer a sistemas computacionais que executassem o processo de transcrição fonética de maneira automática, conforme veremos posteriormente neste trabalho.

Além dessas dificuldades na realização de transcrição fonética manual, temos os desafios naturalmente impostos ao lexicógrafo.

A esse propósito, Duran e Xatara (2007, p. 1) dizem:

O lexicógrafo não é apenas um técnico, um ‘fazedor’ de dicionários, um dicionarista. O lexicógrafo elabora, sim, dicionários, seja de que tipo for, mas o produto que apresenta ao público é substancialmente embasado em estudos lexicológicos e metalexográficos, ou seja, o lexicógrafo só chegará à elaboração de um dicionário, após ter refletido e analisado, com critérios claramente científicos, o tipo de unidade lexical que ele escolherá para compor a nomenclatura⁷ de sua obra, e após ter estabelecido com rigor como será a macro e a microestrutura⁸ dessa obra (DURAN e XATARA, 2007, p. 1).

Entretanto, essa não é a realidade da maioria das produções lexicográficas, pois o que se observa na prática é que, apesar do avanço dos estudos sobre o fazer, são poucas as obras que estão totalmente baseadas em uma teoria que dê sustento a sua elaboração. Algumas obras estabelecem critérios para a construção das definições, outras para a seleção macroestrutural, mas poucas estabelecem critérios para todos os aspectos que constituem um dicionário. Dessa forma, as obras lexicográficas perdem em consistência e coerência, e, portanto, em qualidade.

De fato são muitos os problemas enfrentados por lexicógrafos na elaboração de suas obras: (i) seleção do tipo de usuário em função dos objetivos editoriais e público-alvo; (ii) definição das unidades lexicográficas que devem constituir as entradas dos dicionários, tais como a inclusão, ou não, de sufixos, prefixos, formas flexionadas e derivadas, lexias simples, compostas, expressões fixas, idiomáticas, siglas, nomes próprios, estrangeirismos, etc.; (iii) tratamento de expressões fixas e idiomáticas para evitar redundância e economizar espaço na obra; (iv) o tipo de informação que deve constituir um verbete do dicionário, tais como transcrição fonética, informações morfológicas, sintáticas, marcas de usos, entre outros.; (v) especificamente sobre a informação gramatical, decidir pela inclusão de informações gramaticais extensivas, reduzidas ou apenas demonstrar o uso com exemplos; (vi) o tipo de

⁷ A nomenclatura ou macroestrutura de um dicionário é a sequência vertical das unidades lexicais lematizadas (as palavras-entrada de um dicionário), que podem ser ordenadas semasiológica ou onomasiologicamente. Denomina-se lematização a transformação de unidades de discurso em lema, “unidade léxica ideal que representa um paradigma de formas flexionadas. Essa unidade constitui a típica entrada de dicionário e representa todas as demais formas do paradigma” (BIDERMAN, 1984, p.139).

⁸ Segundo Baldinger (1960, p. 524) apud Welker (2004, p.107) “a microestrutura responde à pergunta sobre as diversas acepções da palavra.” A microestrutura de um verbete pode conter informações sobre: i) grafia, pronúncia, acentuação, classe gramatical, flexão, etimologia, marcas de uso; ii) informações explicativas, ou seja, a definição do lema; iii) uso do lema, ou seja, a sua contextualização ou ilustração, construção e colocação, expressões idiomáticas, provérbios; iv) sinônimos, antônimos, parônimos; v) informações semânticas sobre metáforas; (vi) informações sobre remissivas; vii) ilustrações, gráficos, símbolos, dependendo do objetivo do dicionário. Algumas dessas informações possuem um caráter quase obrigatório, enquanto que outras parecem mais facultativas. Assim, se considerarmos um dicionário de cunho histórico, a presença de informação etimológica é uma informação imprescindível, enquanto que em um dicionário escolar, essa informação não é relevante.

critério a ser adotado para a inclusão das entradas que irão compor o dicionário, tais como a inserção ou não de palavras com base na frequência de uso; (vii) a problemática de como deve ser feita a definição, sobre o modo como deve ser explicada a palavra-entrada.⁹

Acreditamos, portanto, que os desafios para a confecção de uma obra lexicográfica aliados às dificuldades para a realização de transcrição fonética manual contribuem para que a informação sonora das palavras seja colocada de forma assistemática (em uma mesma obra ou comparando dicionários diferentes), de forma não adequada (como comprovado por Cagliari (2011)) ou mesmo contribuindo para que o lexicógrafo opte pela inserção a posteriori desse tipo de informação. Em relação a esse terceiro caso, podemos exemplificá-lo com base no trabalho de Zavaglia (2007), que objetivou discutir o processo de elaboração de um dicionário multilíngue nas direções italiano-português-inglês e português-italiano-inglês para os subdomínios cromáticos “branco”, “negro” e “vermelho”. Tal discussão convergiu para a geração do seguinte verbete-modelo:

CROMÔNIMO em italiano ou português [transcrição fonética], (categoria gramatical) (plural irregular).
*Equivalente em português ou italiano: xxxxxx Exemplo em português ou italiano (de cópula) (autores ou fontes) // Exemplo em italiano ou português (de cópula) (autores ou fontes) *Equivalente em inglês: xxxxxx Exemplo em inglês (de cópula) (autores ou fontes) # Definição em português ou italiano SIN. V.*
REMISSÃO A OUTRO VERBETE

Figura 1. Exemplo de verbete-modelo. Fonte: Zavaglia (2007, p. 213).

No final, a autora afirma que foram elaborados no total 300 verbetes para o referido dicionário, os quais, entretanto, não apresentam transcrição fonética. Podemos deduzir com base nos percalços que a inserção desse tipo de informação promove, que Zavaglia (2007) decidiu que uma das últimas etapas da elaboração de sua obra lexicográfica fosse, justamente, a inserção da transcrição fonética. Devido a esse fato, todos os verbetes possuem marcas para indicar o local no qual será adicionada a transcrição fonética da entrada cromática.

Há ainda outras questões, além da relação grafia-som de uma língua, relacionadas com a informação sonora/transcrição fonética contida em uma obra lexicográfica. Uma delas diz respeito ao modelo de pronúncia adotado, que pode corresponder a uma variedade de prestígio, geralmente correlacionada com a fala de uma determinada classe social ou

⁹ Nesse sentido, Villar e Borba (2011, p. 17) ressaltam as habilidades tanto linguísticas quanto técnicas exigidas por um lexicógrafo. Segundo eles, além do conhecimento das propriedades inerentes aos níveis de organização linguístico (fonético, fonológico, ortográfico, morfológico, sintático, semântico e pragmático), o lexicógrafo deve conhecer a técnica lexicográfica, ou seja, dominar os princípios e praxes estabelecidos ao longo do tempo pela prática lexicográfica, incluindo também o domínio sobre as atuais técnicas computacionais de compilação e investigação de cópula. Os autores apontam como uma das tarefas fundamentais desempenhadas via cópula atualmente, a obtenção de dados frequenciais sobre usos e aceções de palavras, a fim de constituir a macro e a microestrutura de um dicionário.

localização geográfica.

Diante desses desafios, podemos dizer que a disponibilização de um recurso computacional para anotação fonética automática, como o que aqui propomos, poderá contribuir consideravelmente para pesquisas e tarefas nas quais a anotação fonética é um item de vital importância. Assim, acreditamos que a natureza híbrida deste trabalho ilustra como a intersecção entre as áreas das Ciências do Léxico, do Processamento de Linguagem Natural (PLN) e da Fonética e Fonologia pode atuar em prol dos lexicógrafos, reduzindo o tempo e as dificuldades em adicionar transcrição fonética aos dicionários do PB que poderão ser produzidos futuramente.

OBJETIVOS

O objetivo principal desta pesquisa visou a construção do PETRUS (*Phonetic TRanscriber for User Support*), um ambiente *Web* gratuito de suporte à transcrição fonética automática das unidades lexicais que fazem parte da cabeça do verbete de dicionários mono ou multilíngues do PB.

De fato, a automatização desse tipo de procedimento torna possível a anotação fonética de pequenos ou grandes corpúscos por meio da aplicação de um conjunto fixo e objetivo de critérios de anotação. Sem que haja intervenção humana, é produzida transcrição sem possíveis subjetividades de anotadores humanos e que, quando realizada com o mesmo sistema computacional de transcrição, garante que os resultados alcançados (qualidade de transcrição) sejam idênticos. Esse é um fator importante quando se cria um corpúscos de grandes dimensões. Além disso, as transcrições automáticas podem ser criadas com uma fração dos custos e do tempo com que são geradas transcrições manuais.

Uma vez disponível gratuitamente na *Web*, demais interessados em obter transcrição fonética de modo automático (professores de português, foneticistas, especialistas da área de síntese de fala, por exemplo) poderão também utilizá-lo, aumentando assim o perfil do usuário que poderá ter acesso ao sistema desenvolvido.

Devido à interdisciplinaridade aqui envolvida, em alguns tópicos foi necessária a incursão em áreas da Computação que, a princípio, não são comuns com a Linguística, o que tornou os objetivos deste trabalho desafiadores e motivantes.

Quanto aos objetivos específicos almejamos:

- 1) Propor um conjunto de regras linguísticas para tratar a problemática da conversão

grafema-fone de unidades lexicais simples do PB, segundo uma variante padrão desse idioma, e avaliar seu desempenho;

- 2) Disponibilizar livre e gratuitamente a ferramenta proposta, a partir de um Portal de recursos linguísticos, como o mantido pelo Núcleo Interinstitucional de Linguística Computacional - NILC,¹⁰ com uma interface gráfica de fácil interação, de modo a facilitar o acesso de potenciais usuários ao referido sistema;
- 3) Disponibilizar livre e gratuitamente os algoritmos para conversão grafema-fone utilizados no sistema PETRUS, de modo que possam ser aproveitados em outros projetos, diminuindo o tempo e as dificuldades envolvidas na construção de um sistema como este.
- 4) Disponibilizar os recursos linguísticos produzidos durante a confecção do sistema, a saber: lista com homógrafos heterófonos, lista com prefixos, cópulas de teste.

METODOLOGIA

Apesar de alguns autores considerarem a conversão grafema-fone em sentido estrito, ou seja, apenas a tarefa de transformação de caracteres ortográficos em seus respectivos símbolos sonoros (no caso, fonéticos), entendemos tal processo de conversão em um sentido mais amplo, assim como Braga (2008), no qual consideramos igualmente importantes a existência de um módulo de divisão silábica, um de marcação de sílaba tônica, um de identificação da classe gramatical e, por fim, um de transcrição fonética, visto que as informações dos três primeiros têm influência na construção e desempenho deste último.

Diferentes técnicas têm sido utilizadas para o desenvolvimento desse tipo de sistema de transcrição, tanto para o português europeu quanto para o brasileiro, cada uma com suas características negativas e positivas.

A escolha pelo melhor método de conversão depende, geralmente, dos recursos disponíveis (tais como memória e capacidade de processamento da máquina que irá realizar o processo de conversão) e, o mais importante, a língua na qual o sistema de transcrição fonética automática será construído. Uma descrição mais detalhada das técnicas existentes para a realização da tarefa de conversão de grafemas em fone(mas) podem ser consultadas no Capítulo 1 deste trabalho. Em suma, podemos dizer que as principais abordagens utilizadas

¹⁰ Grupo interdisciplinar de linguistas e cientistas da computação, criado em 1993 para desenvolver pesquisas e projetos relacionados com Linguística Computacional e Processamento de Língua Natural, tais como desenvolvimento de léxicos e cópulas, sumarização automática, tradução automática e ferramentas de suporte a escrita. Pode ser acessado em: <http://www.nilc.icmc.usp.br/nilc/>.

para a construção do conhecimento linguístico necessário em um sistema como tal são: (i) a baseada em conhecimento (*knowledge based*), que é geralmente amparada por um linguista e faz uso de regras pré-definidas para a conversão de forma ortográfica em seu respectivo símbolo sonoro e a (ii) baseada em dados (*data-driven*), a qual se vale do aprendizado de máquina para a geração automática de regras, a partir de um dicionário de treino. No tocante ao nosso sistema em particular, seguiremos a primeira metodologia mencionada, na qual um conjunto de regras linguísticas dependentes de contexto (silábico, tônico e gramatical) indica como transcrever os grafemas do PB em suas respectivas unidades fonéticas, aliada ao uso de um dicionário fonético, ou seja, uma lista com palavras cujas transcrições fonéticas não seguem as regras de transcrição propostas, no caso, homógrafos heterófonos.

Essas regras linguísticas que foram implementadas no sistema PETRUS operam apenas no domínio da palavra (UL) simples, ou seja, somente na sequência de caracteres constitutivos de uma palavra isolada, ou seja, separada por espaços em branco. Não estão previstas, portanto, regras relacionadas aos processos fonológicos que ocorrem na fronteira de duas ou mais palavras, responsáveis por fenômenos de coarticulação.

A arquitetura do sistema automático de transcrição ora descrito é composta por seis módulos, cujo conhecimento linguístico necessário ao seu desenvolvimento será apresentado com mais detalhes no Capítulo 2.

Em relação ao acesso, o PETRUS está disponível livre e gratuitamente para uso, no portal do NILC: <http://www.nilc.icmc.usp.br/petrus/> e constitui-se num sistema multiplataforma, ou seja, pode ser executado em mais do que uma plataforma (ambiente), podendo, portanto, ser acessado via computador, *tablet* ou *smartphone*.

No tocante à variedade linguística adotada para o sistema desenvolvido, segundo Corrêa et al. (2010) todo lexicógrafo deve optar por um padrão de pronúncia na transcrição fonética das entradas vocabulares de um dicionário bilíngue. Na “lexicografia bilíngue do português e do alemão [...] as obras mais novas surgiram apenas na forma de dicionários de bolso, que, aliás, em sua totalidade, refletem a pronúncia europeia” (NOLL, 2008, p. 28 apud CORRÊA et al., 2010). A língua portuguesa europeia tem como norma padrão, a partir do século XIV, os dialetos¹¹ falados entre Coimbra e Lisboa, com especial relevo para este último.

No Brasil, a norma padrão evoluiu do dialeto de Lisboa para o do Rio de Janeiro (com a vinda da corte para o Brasil em 1808). E, desde então, para uma influência partilhada pelas variedades cultas em uso nas maiores cidades do país (TEYSSIER, 1997), em especial as duas

¹¹ Dialeto: as propriedades linguísticas compartilhadas por um grupo específico de falantes.

maiores metrópoles globais na costa da região sudeste: São Paulo e Rio de Janeiro. São os dialetos que apresentam maior exposição, devido à condição de centro econômico e midiático das duas cidades. Qual pronúncia adotar para este estudo?

Para responder a essa pergunta, apoiamo-nos em uma investigação desenvolvida pelo Núcleo de Pesquisa e Ensino de Português/Língua Estrangeira, no Instituto de Pesquisas Linguísticas “Sedes Sapientiae” da PUC/SP, sob coordenação da Profa. Dra. Regina Silveira. No referido estudo, Silveira (2008, p. 21) relata que foram realizadas uma série de gravações com diferentes variedades da língua (padrão gramatical normativo, padrão real e padrão nativo) e com diferentes variedades/variações geográficas e sociais, como, por exemplo, a pronúncia de contadores de casos, cantores sertanejos e de outras músicas populares brasileiras, apresentadores de noticiários locais e nacionais, políticos, apresentadores de anúncios publicitários, professores universitários e participantes de mesas-redondas. Em seguida, os informantes deveriam, ao ouvir essas gravações, indicar qual pronúncia “consideravam a “melhor” e que teriam como expectativa adquirirem” (Silveira, 2008, p. 21). Os informantes eram falantes nativos de português e falantes de outras línguas (tanto de distanciamento quanto de interface com a língua portuguesa), sendo que estes últimos, conforme a autora detalha, estavam em estado de interlíngua e encontravam-se em exposição ao português brasileiro, bem como com o ensino formal em São Paulo. De acordo com Silveira (2008, p. 21), após várias seções, de maneira geral, tanto nativos quanto estrangeiros atribuíram o grau ótimo à pronúncia dos apresentadores do “Jornal Nacional” da TV Globo, o noticiário da televisão brasileira com os maiores índices de audiência no Brasil.

Os resultados e considerações de Silveira (2008) indicam que o “globês” foi construído

(...) com a neutralização de traços articulatorios específicos de nossas variedades/variações linguísticas orais, apresentando-se como uma variável mais neutra, com o objetivo de ser amplamente aceita por falantes/ouvintes do português brasileiro, de forma a conseguir um grande público de telespectadores, ainda que estes apresentem variações linguísticas diferentes. O fato de ter amplo acesso geográfico diário, nacional e internacional, propiciou que essa arquinorma, progressivamente, devido ao “prestígio” que a Globo tem com os telespectadores, fosse instaurada como a unidade na diversidade de pronúncias dos diferentes grupos linguísticos sociais e geográficos. Essa arquinorma passa a ser reconhecida por nativos e estrangeiros como a pronúncia mais representativa do brasileiro e, ideologicamente, avaliada com o grau ótimo de aceitabilidade. Nesse sentido, passa a ser, política e ideologicamente, reconhecida, neste momento histórico, como uma pronúncia padrão identitária (SILVEIRA, 2008, p. 33).

Embora o globês seja uma pronúncia estandardizada inventada pela mídia para não

deixar em evidência traços dialetais na pronúncia dos locutores, pois tais traços não seriam muito agradáveis para algumas comunidades de falantes, acreditamos que pelo fato dessa variedade ser reconhecida por diferentes falantes (nativos e estrangeiros) como representativo de seu dialeto, pode ser uma opção interessante de variedade linguística a ser representada foneticamente em dicionários do PB.¹²

Feita essa escolha, apoiamo-nos em estudos fonéticos e fonológicos que descreveram a pronúncia da variedade em questão. A variedade descrita e adotada por Cagliari (2012), por exemplo, se assemelha à utilizada pela Rede Globo. Apesar disso, a transcrição fonética realizada de modo automático pelo sistema PETRUS exigiu uma certa generalização comprometida tanto com uma variedade de língua do PB que fosse bem aceita pela comunidade linguística, quanto com o seu processamento automático, conforme poderá ser observado com mais detalhes na seção 1.4. Tal generalização se deu justamente em direção à determinação do conjunto de regras de transcrição implementado, visto que, apesar de ter sido considerada uma pronúncia de prestígio, há variação na pronúncia de algumas palavras e, assim como Cagliari (2012), tivemos que privilegiar algumas características fonéticas. Os símbolos utilizados para a transcrição fonética são os do Alfabeto Fonético Internacional (AFI). Concebido com caráter científico e universal, apresenta a vantagem de ser utilizado, de modo geral, em lexicografia, tanto em dicionários bilíngues quanto monolíngues. Assim, as transcrições são apresentadas entre parênteses retos e o acento tônico primário marcado sempre antes da sílaba tônica, indicado através do diacrítico (').

A partir de então, propusemos, com base na variedade padrão aqui adotada, um sistema computacional para realizar uma transcrição fonética que representasse um discurso silabado, ou seja, com velocidade de produção um pouco mais lenta que a fala normal, fato que refletiu, como veremos mais adiante, na transcrição fonética dos ditongos crescentes.

ORGANIZAÇÃO DO TRABALHO

A concretização dos objetivos apontados anteriormente implicou na realização de um conjunto de tarefas acerca da construção de recursos linguísticos necessários ao funcionamento pleno da ferramenta de conversão grafema-fone aqui proposta, que passamos a

¹² Importante dizer que o glôbês foi apenas a primeira variedade escolhida para constituir o sistema PETRUS, devido aos motivos supra mencionados. Entretanto, pretendemos, como trabalho futuro, implementar regras de transcrição que reflitam também outras variedades do PB. Dada a metodologia de sua implementação, poderá, inclusive, comportar regras de transcrição fonética para outros idiomas.

apresentar. No Capítulo 1 são tecidas considerações sobre o tema “transcrição fonética”, que compreendem desde a escolha do alfabeto fonético a ser utilizado até questões envolvidas no processo manual, semiautomático e automático para a obtenção de transcrição fonética. Há também, neste mesmo capítulo, uma revisão dos fundamentos teóricos que auxiliaram na execução deste projeto, entre eles, questões ortográficas, fonéticas, fonológicas e prosódicas que permearam a escolha dos símbolos fonéticos e a posterior construção manual de regras para realizar a conversão automática de formas grafadas em seus respectivos correlatos sonoros (fones), além de questões sobre a variante linguística adotada neste estudo – o padrão utilizado pela Rede Globo. No Capítulo 2, apoiando-nos nas reflexões realizadas sobre os trabalhos correlatos e aportes teóricos trazidos para este trabalho, temos a elaboração da proposta de trabalho deste projeto: a construção do PETRUS. No Capítulo 3, apresentamos os resultados preliminares das avaliações feitas sobre o PETRUS e, no último, o Capítulo 4, temos as considerações finais e a indicação das próximas atividades para a pesquisa apresentada.

CAPÍTULO 1

FUNDAMENTAÇÃO TEÓRICA E ESTADO DA ARTE EM TRANSCRIÇÃO FONÉTICA

Neste capítulo é apresentada uma revisão bibliográfica relacionada ao estado da arte sobre transcrição automática e do aporte teórico (relacionado com a fonética, fonologia e prosódia da língua-alvo e, conseqüentemente, da variante linguística escolhida) necessários ao desenvolvimento do sistema de conversão grafema-fone PETRUS. Subjacente ao desenvolvimento de sistemas conversores como o que aqui propomos deve haver uma minuciosa compreensão do sistema linguístico da língua-alvo, em relação aos aspectos intrínsecos dessa relação grafia-som. Isso porque a falta de isomorfismo existente entre grafemas e fones costuma causar dificuldades na realização desse tipo de conversão de forma automática. Portanto, para que possamos atingir nosso objetivo principal, foi necessário que conhecêssemos os padrões sonoros específicos da variedade particular do português brasileiro adotada para este trabalho: a variedade linguística adotada pela Rede Globo. Dessa maneira, nas próximas seções, serão apresentados os fundamentos dos aspectos supracitados, que serviram de base para o desenvolvimento deste trabalho.

1.1 FUNDAMENTAÇÃO TEÓRICA

1.1.1 Relação grafia e som

É certo que toda comunidade possui uma língua falada, condição essa considerada por muitos pesquisadores uma característica inata e definidora da espécie humana. Contudo, nem todas as línguas têm sistemas que as representem na forma escrita.

Um sistema de escrita é um tipo de comunicação realizada por meio de símbolos, que são utilizados para registrar visualmente uma língua falada, objetivando uma comunicação. O conjunto ordenado das letras constitui um alfabeto. Em português, utilizam-se também os acentos (agudo e circunflexo) para indicar a pronúncia de algumas vogais em contexto tônico. O conjunto de normas que estabelece o uso dos sinais gráficos constitui a ortografia de uma língua.

Os sistemas de escrita distinguem-se de outros sistemas simbólicos utilizados para a comunicação pelo fato de que devemos predispor de conhecimento a respeito da língua em questão para que possamos ler e compreender as letras. Contrariamente às letras, existem símbolos que não exigem necessariamente algum tipo de conhecimento prévio de língua para permitir extrair deles o significado a eles associado, como ocorre por exemplo com fórmulas matemáticas, mapas, pinturas, sinais de trânsito, entre outros.

Há diferentes sistemas para representação da escrita nas diferentes línguas humanas:

- *Sistema Pictográfico*: sistema mais primitivo de escrita em que objetos são representados por desenhos que visam retratar o mundo da forma o mais fiel possível. Foi a base da escrita cuneiforme e dos hieróglifos. Na Figura 2 temos exemplos de pinturas rupestres encontradas no sítio arqueológico de Tassil n'Ajjer, na Argélia, produzidos entre 7 mil a 13 mil anos atrás.

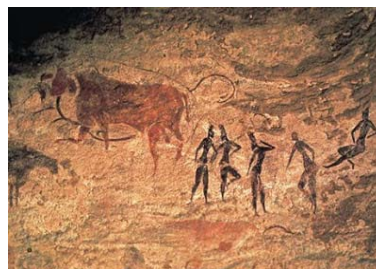


Figura 2. Pinturas rupestres.
Fonte: <http://en.wikipedia.org>.

- *Sistema Ideográfico*: são utilizados ideogramas para representar palavras ou conceitos, como por exemplo, os caracteres chineses. Na Figura 3 temos representado à esquerda o ideograma chinês para *pinheiro*, formado a partir dos elementos semânticos madeira (ideograma central) e beleza (ideograma da direita)



Figura 3. Sistema ideográfico chinês.
Fonte: Morais (1995, p. 53)

- *Sistemas silábicos*: comportam informação sobre o modo como se deve pronunciar uma palavra, cujos símbolos representam uma sílaba. A escrita suméria e a japonesa (que utiliza dois silabários, o hiragana e o katakana, ambos desenvolvidos por volta do ano 700) constituem exemplos desse tipo de sistema de escrita. Nesses sistemas, pode haver cinco símbolos para sílabas iniciadas por /m/: um símbolo para *ma*, outro para *me*, e assim por diante. A palavra <mame> seria escrita utilizando apenas dois símbolos, por exemplo. Na Figura 4 temos a escrita katakana que é utilizada para escrever palavras estrangeiras, dar ênfase a alguns termos e reproduzir sons (onomatopéias).

A	ア a	カ ka	サ sa	タ ta	ナ na	ハ ha	マ ma	ヤ ya	ラ ra	ワ wa	ン (n)
I	イ i	キ ki	シ shi	チ chi	ニ ni	ヒ hi	ミ mi		リ ri		
U	ウ u	ク ku	ス su	ツ tsu	ヌ nu	フ fu	ム mu	ユ yu	ル ru		
E	エ e	ケ ke	セ se	テ te	ネ ne	ヘ he	メ me		レ re		
O	オ o	コ ko	ソ so	ト to	ノ no	ホ ho	モ mo	ヨ yo	ロ ro	ヲ wo	

Figura 4. Escrita katakana.
Fonte: <http://estudejapones.com/lessons/3lesson3.html>.

- *Sistemas abugidas* (ou alfasilabário), é um conjunto de sinais que visa representar os fonemas de uma língua. Situa-se entre um sistema silabário e um alfabético. É constituído por sinais que representam as sílabas dotadas de vogais, e de outros sinais anexos que modificam, substituem ou suprimem certas vogais. O devanágari é um exemplo de sistema de escrita alfabeto-silábico, da família brâmica, do sul da Ásia, utilizado desde o século XII. Muitas línguas da Índia, como o sânscrito e o híndi fazem uso do devanágari. A Figura 5 é um exemplo de manuscrito em devanágari anterior ao século 19.

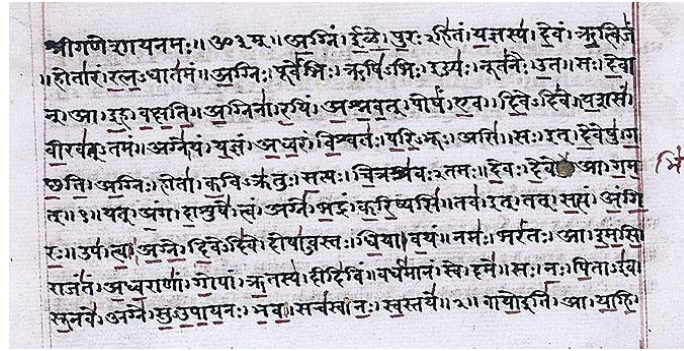


Figura 5. Rigveda manuscrito em Devanágari
Fonte: <http://en.wikipedia.org>.

- *Sistemas abjads* (ou consonantários): os símbolos das letras que representam as consoantes podem ser lidos como sendo acompanhados por uma ou mais sílabas. Em muitos abjads podem ser opcionalmente adicionadas indicações de vogais, utilizando-se para isso sinais diacríticos, que são pequenos traços ou pontos colocados abaixo ou acima das consoantes, no sentido de conduzir a leitura ao assunto abordado. A título de exemplo de sistemas de escrita abjads, podem ser citadas as escritas aramaica, siríaca, árabe e hebraica. Na Figura 6, trecho da bíblia hebraica do século 10.

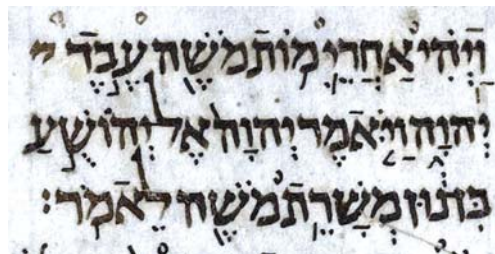


Figura 6. Trecho da bíblia hebraica do século 10.
Fonte: <http://en.wikipedia.org>

- *Sistemas alfabéticos*: é uma forma de escrita classificada como segmental, pois é constituída por grafemas que representam fonemas de uma língua. Como exemplo podemos citar o alfabeto latino, o grego, o cirílico e o glagolítico.

Portanto, uma condição *sine qua non* para um indivíduo dominar o sistema alfabético é a capacidade de segmentar a fala em fonemas para poder representá-los por meio de grafemas (SCLIAR-CABRAL, 2003, p. 50). A esse respeito, Morais (1995) afirma que aprender a utilizar um código alfabético é, ao mesmo tempo, aprender a encontrar os correspondentes

fonêmicos das letras e a fundir os fonemas sucessivos. Sobre a questão da fusão de fonemas sucessivos, o autor está se referindo ao fenômeno da coarticulação.¹³

Sistemas de escrita alfabéticos são baseados na noção de que a forma ortográfica é uma representação convencional da pronúncia de uma palavra. A ortografia foi inventada para evitar os efeitos da variação linguística, sendo responsável também por sistematizar o modo como cada palavra deve ser grafada em cada língua. Assim, podemos explicar o porquê das palavras apresentarem variadas formas de serem expressas oralmente, mas apenas uma maneira de serem produzidas graficamente.

A respeito da natureza da ortografia, Kato (2002) afirma que, embora a intenção inicial tenha sido a de criar um alfabeto de natureza fonética, o fato de toda língua mudar, ter diferenças dialetais e variações estilísticas que afetam sua pronúncia, impediu que a escrita alfabética pudesse ter uma natureza estritamente fonética. De acordo com a autora, a natureza ortográfica do português é essencialmente fonêmica, ou seja, procura representar na escrita o que é funcional no sistema de sons da língua, isto é, aquilo que possui valor contrastivo, no caso, os fonemas. Por exemplo, as palavras do português “bato”, “fato”, “gato”, “mato”, “pato”, “rato”, “tato” ilustram claramente o valor contrastivo das consoantes , <f>, <g>, <m>, <p>, <r>, <t>, que ao serem revezadas diante da sequência “ato”, geraram novas unidades lexicais. Assim, comutando essas consoantes do português, podemos deduzir que /b/, /f/, /g/, /m/, /p/, /r/, /t/ além de unidades da língua, também representam fonemas.

Dessa maneira, a unidade do sistema de escrita é garantida, de acordo com Kato (2002), pela natureza fonêmica de nossa ortografia portuguesa. Se nossa escrita fosse de base estritamente fonética, ou seja, representasse exatamente todos os sons possíveis da fala, teríamos uma diversidade tamanha de possibilidades de símbolos para representação sonora, que a unidade da língua poderia ficar comprometida. Isso porque deveria ser reproduzida na escrita toda a variedade dialetal verificada nos falares dos diferentes falantes do português e, conseqüentemente, a escrita perderia seu caráter unificador, dado que deveríamos considerar as diferenças regionais, socioculturais e até mesmo individuais de um país de dimensões continentais como o nosso ou mesmo das dimensões globais de um idioma como o português, adotado como língua oficial em oito países.

¹³ Em decorrência do *continuum* sonoro da fala e das modificações do trato vocal para a produção de um segmento, em muitos casos, tem-se o processo de coarticulação, tradicionalmente definido como a interferência de um segmento na realização de outro.

Na verdade, o que acontece é que as formas ortográficas não representam a fala de ninguém – sua função é, basicamente, anular a variação linguística no nível da palavra. Portanto, eu posso falar “BAUDGI”, meu vizinho “BAUDI”, um conhecido “BARDGI”, e um amigo do Sul, “BALDE”, mas todos escrevemos “BALDE”, não porque esta forma representa uma “pronúncia correta” (o que não é verdade), mas porque houve um consenso, na sociedade, de que todos devem escrever esta palavra assim. Não adianta ficar se perguntando o porquê de se ter escolhido esta forma e não outra “por exemplo, “BAUDE”), já que esta escolha esbarra em questões históricas, que nos escapam com o passar dos anos (MASSINI-CAGLIARI, 2001, p. 124).

Assim, a ortografia neutraliza esse tipo de variação fonética e cada palavra poderá ter, então, um aspecto reconhecido imediatamente por todos os falantes alfabetizados da língua, mas que não impede que cada palavra escrita seja pronunciada de modo diferente em cada região.

Ainda a respeito da relação língua oral e língua escrita, Kato (2002, p. 19) afirma que a “língua oral muda e a escrita é conservadora, o que acarreta um afastamento gradativo entre as duas, e o que resulta é um misto de relações motivadas e arbitrárias”. Nas palavras de Cagliari (1999a, p. 65-66), devido ao fato de nosso sistema ortográfico visar representar por meio de sinais gráficos o que é essencialmente significativo na fala oral, neutralizando, portanto, a variação linguística existente, surge uma série de “incongruências” entre os sons e as letras.

Devido à natureza fonológica de nosso sistema alfabético, podemos ser induzidos, por exemplo, à conclusão de que há mesmo uma relação unívoca entre a letra e o som que ela representa, por conta do princípio acrofônico, que define que no nome das letras podemos encontrar o som que ela representa, por exemplo, letras “be”, “de”, “o”, etc. Dessa maneira, em uma situação ideal, teríamos uma correspondência biunívoca entre grafema e fonema, ou seja, uma grafema para cada fonema e vice-versa. Na realidade, isso não acontece na língua portuguesa. Como veremos a seguir, há muitos casos que fogem a essa relação um-para-um.

A esse respeito, Braga et al. (2003) afirmam que apesar de a escrita da língua portuguesa ter natureza alfabética, isto é, baseada numa relação unívoca entre um som e um grafema, na verdade, essa relação não se verifica atualmente, devido à tensão existente entre o dinamismo da evolução fonética da língua e o carácter conservador, próprio de todos os sistemas de escrita. Surgindo, conseqüentemente, dificuldades na elaboração de regras dedicadas à conversão automática de alguns grafemas do PB (ex.: consoante <x>) e, em especial, dentro do vocalismo do português (vogais <e>, <o>).

Em PB, são biunívocos os grafemas <b, f, j, k, p, v, q, y>. Isso quer dizer que não existe nenhum caso em que o fonema /b/ não seja representado pelo grafema , o fonema /f/ representado por <f>, e assim por diante. Há, entretanto, casos em que a relação grafema-som do PB não é perfeita, como podemos observar a seguir, considerando a variedade linguística aqui adotada:

- <m> → /m/ ou [m̃], o grafema <m> pode ser representado fonemicamente como a consoante nasal bilabial /m/ (*maca* ~ /maka/), mas também pode ser utilizado como símbolo para representar um elemento nasal homorgânico (ou seja, que possui o mesmo lugar de articulação) imediatamente após a vogal nasal, como em *bomba* [ˈbõ^mba] (Cf. seção 1.1.2.2.2).
- <n> → /n/ ou [ñ], assim como no caso anterior, o grafema <n> pode tanto ser representado fonemicamente como uma consoante nasal alveolar /n/ (ex. /nata/) quanto como elemento nasal, utilizado anteriormente às consoantes [ʃ, ʒ, tʃ, dʒ].
- <l> → /l/, como em *lata* ~ /lata/, *placa* /plaka/; entretanto, realiza-se foneticamente como [ʊ] em final de sílaba (*filme* [ˈfiõmɪ]) ou final de palavra (*formol* [forˈmõʊ]).
- <t> → /t/ ou [tʃ], como em /tatu/, mas diante de [i] ou [ɪ], torna-se uma consoante africada alveopalatal, como em *titia* [tʃiˈtʃia].
- <d> → /d/ ou [dʒ], como em “/d/ado”, “/d/edo” e “/d/uro”, mas diante de [i] ou [ɪ], torna-se uma consoante africada alveopalatal: *ditado* [dʒiˈtadõ].
- <c> → /k/ ou /s/, como em “/k/abeça” e “/s/inema”.
- <g> → /g/ ou /ʒ/, como em “/g/ato” e “/ʒ/irafa”.
- <h> → letra que não corresponde a nenhum som: *hiato* /iato/.
- <s> → /s/ ou /z/, como em “/s/aca” e “ca/z/a”.
- <z> → /s/ ou /z/, como em “arro/s/” e “/z/ebra”.
- <r> → [x], [r] ou [ʁ] como em “ca[x]o”; “ca[r]o”; “p[r]ato”; “ve[r]de”; “[x]ua”; “ma[x]”; “go[ʁ]do”.
- <x> → /ʃ/, /z/, /s/ ou /ks/, como em “/ʃ/ícara”; “e/z/ame”; “pró/s/imo” e “tá/ks/i”.
- <w> → /v/ ou /u/, como na palavra alemã *Volkswagen* ou na palavra chinesa *kiwi*.
- <a> → /a/, /ẽ/ ou [a], como em “/a/mor”; “c/ẽ/ma” e “[a]uto”.¹⁴
- <e> → /e/, /ɛ/, /i/ ou [ɪ], como em “m/e/do”; “l/ɛ/va”; “p[i]rigo” e “[i]scola”.
- <i> → /i/ ou [ɪ], como em “/i/greja”; “dent[ɪ]”.

¹⁴ Para mais informações sobre o uso de [a] nos ditongos <au, al>, vide Cagliari (2012, p. 480).

- <o> → /o/, /ɔ/, /u/ ou [ʊ], como em: “/o/vo”; “/ɔ/vos”; “m[u]leque” e “pov[ʊ]”.
- <u> → /u/ ou [ʊ], como em “[u]va” e “suc[ʊ]”.

Há casos ainda em PB nos quais pode haver a associação de grafemas para a representação de um único som, formando os casos de dígrafos, presentes em nossa língua:

- <qu> → /k/, como em “/k/eijo”.
- <gu> → /g/, como em “/g/erra” e “/g/inada”.
- <ch> → /ʃ/, como em “/ʃ/ulé”.
- <nh> → /ɲ/, como em “ma/ɲ/ã”.
- <lh> → /ʎ/, como em “fo/ʎ/a”.
- <ss> → /s/, como em “no/s/a”.
- <sç> → /s/, como em “de/s/a”.
- <sc> → /s/, como em “na/s/er”.
- <xc> → /s/, como em “e/s/eção”.
- <rr> → [x], como em “to[x]e”.

Além disso, há grafemas no PB aos quais podem ser adicionados diacríticos que lhes modificam o valor, como a cedilha (<ç> → [s], como na palavra *coça*); o til para marcar nasalidade (como na palavra *maçã*) ou os acentos agudo e circunflexo, utilizados para modificar o valor das vogais (como em *vovô* e *vovó*).

Ainda que breve, esse panorama sobre as relações que se estabelecem entre o sistema gráfico e o sonoro de nossa língua nos possibilita antever a complexidade existente na tarefa de se criar um conversor fonético que realize a relação grafia-som de maneira rápida, automática e o mais correta possível.

Os erros reportados na literatura sobre conversão automática grafema-fone(ma) para o português, decorrem especialmente desse não isomorfismo existente entre as letras e seus respectivos sons, como podemos deduzir a partir das taxas de erros publicadas na literatura referentes aos conversores ortografia-fone(ma) produzidos. Mesmo no caso do português, caracterizado por uma elevada regularidade, foi possível citar nesta seção alguns exemplos nos quais a correlação som-grafia não é nada linear.

1.1.2 Ortografia¹⁵

Após o Decreto presidencial no 6.583, de 29/9/2008, houve a promulgação do Acordo Ortográfico de 1990 (AO90). A partir desse momento, passou a vigorar no Brasil o período de transição para as regras ortográficas que se finaliza em 31 de dezembro de 2015. Sua entrada em vigor está prevista para 1/1/2016, conforme o Decreto n° 7.875, de 27/12/2012. Até lá, coexistirão a norma ortográfica atual e a norma estabelecida pelo acordo.

Algumas modificações foram feitas no sentido de promover a união da ortografia de língua portuguesa dos seguintes países que têm o português como língua oficial: Angola, Moçambique, Cabo Verde, Guiné-Bissau, São Tomé e Príncipe, Timor Leste, Brasil e Portugal. Embora a grafia passe a ser unificada em tais países, a pronúncia e o vocabulário não mudarão.

As modificações ortográficas ratificadas no referido acordo dificultaram ainda mais alguns casos de identidade fonética de certos vocábulos da língua portuguesa. O AO90 propõe mudanças variadas na língua portuguesa escrita, porém tem influência apenas em quatro momentos da tarefa de transcrição automática para o PB:

1. Inclusão das letras “k”, “w” e “y” no alfabeto oficial

Para tais letras, nosso sistema utiliza os respectivos fones [k], [u], [i] para realizar a transcrição fonética.

2. Abolição do trema

Até a entrada em vigor do AO90, o trema era usado no PB para indicar que a letra <u> nas combinações *güe, güi, qüe, qüi* deve ser pronunciada. A partir do AO90, o trema deixou de ser usado, exceto em nomes próprios e derivados. Assim, palavras como *ágiue, sangüíneo, consequência* e *argüir* deixam de ter trema. No entanto, continua a ser usado em palavras estrangeiras e derivadas como *Günter, Bündchen, Müller e mülleriano*.

Para manter a pronúncia natural das palavras que perderam o trema, foi necessário incluir em nosso transcritor regras específicas para esses casos, pois continuam a ser pronunciadas da mesma forma apesar da alteração ortográfica. As regras criadas para esses casos são:

- Se a letra <q> for sucedida pela vogal <u>, a vogal <e> mais a consoante <n, s> e a letra <q> não é a primeira letra da palavra, então o sistema PETRUS transcreverá a letra <u>

¹⁵ As regras do Acordo Ortográfico de 1990 aqui apresentadas são adaptações das regras apresentadas por Bechara (2009) e Tufano (2008).

como [ʊ] (Cf. <cinquenta> [sĩ'koẽĩta], <equestre> [e'kœstri]). Caso contrário, <u> não será transcrito (Cf. <aquela> [a'kela]).

- Se a letra <g> for sucedida pela vogal <u>, a vogal <e> em final de palavra ou se a sequência <gue> é sucedida pelas consoantes <m, n, t, r, l>, então <u> é transcrito como [ʊ]. Exemplos: <enxag[ʊ]e> [ẽĩ'ʃagʊ], <ag[ʊ]emos> [a'goemos], <ag[ʊ]entar> [agʊẽĩ'tax], <alcag[ʊ]eta> [aũka'gœta], <Anhang[ʊ]era> [aĩã'gœera], <cirig[ʊ]ela> [siri'gœela].
- Se a letra <g> for sucedida pela vogal <u>, mais a vogal <i> em final de palavra ou se a sequência <gui> é sucedida pelas consoantes <d, ç, r, s, m>, então <u> é transcrito como [ʊ]. Por exemplo, <ambig[ʊ]idade> [ãbigʊi'dadʒi], <arg[ʊ]içã> [axgʊi'šã], <arg[ʊ]ir> [ax'gʊix], <Birig[ʊ]i> [biri'gʊi], <ling[ʊ]ística> [lĩ'gʊistĩka] e <ping[ʊ]im> [pĩ'gʊĩ].
- Se a letra <q> for sucedida pela vogal <u>, mais a vogal <i> e uma das consoantes <s, c, r, v, d, q, l>, então <u> é transcrito como [ʊ]. Por exemplo, <antiq[ʊ]íssimo> [ãtĩ'kʊisĩmʊ], <aq[ʊ]icultura> [akʊikuʊ'tura], <delinq[ʊ]ir> [delĩ'kʊix], <exeq[ʊ]ível> [eze'kʊiveʊ], <liq[ʊ]idificador> [likʊidʒifika'dox], <q[ʊ]inq[ʊ]agésimo> [kʊĩkoa'ʒezĩmʊ], <tranq[ʊ]ilidade> [trãkʊili'dadʒi].

Vale dizer que essas regras implementadas segundo o AO90 foram criadas de forma arbitrária, baseadas em uma análise por amostragem, na qual foi possível identificar uma tendência ou um padrão de escrita. É possível que existam palavras que não se encaixem nas regras implementadas, o que resultará numa transcrição inadequada das mesmas.

3. Exclusão do acento agudo nos ditongos 'éi' e 'ói' em palavras paroxítonas

A partir do AO90, não se utiliza acento agudo nos ditongos abertos 'éi' e 'ói' de palavras paroxítonas. Assim, *bóia*, *idéia*, *platéia* são escritas, após o acordo, como *boia*, *ideia*, *plateia*. Entretanto, Tufano (2008) ressalva que continuam a serem grafados os acentos agudos das palavras oxítonas e dos monossílabos tônicos terminados em 'éis' e 'óis'. Por exemplo, *papéis*, *dói*, *herói*, *heróis*.

Para manter a pronúncia natural dos referidos ditongos em palavras paroxítonas, elaboramos uma lista com palavras contendo o ditongo 'ói', cuja grafia estava adequada aos critérios estabelecidos antes do acordo ortográfico, ou seja, as palavras ainda estavam acentuadas. Assim, conseguimos analisar o contexto de ocorrência desse ditongo de modo a encontrar padrões que indiquem quando se trata de um ditongo aberto (i.e., *androide*) ou de um ditongo fechado (i.e., *coisa*). Dessa forma, foi possível obter a seguinte regra de transcrição para o sistema PETRUS:

- A sequência gráfica <oi> será transcrita foneticamente como [ɔi] quando sucedida por: <a> (*jiboia*), <am> (*apoiam*), <co> (*heroico*), <ca> (*Perestroica*), <de> (*asteroide*), <e> (*apoie*), (*apoiem*).

Dada a oscilação existente entre a pronúncia aberta (*ideia*, *boleia*) e a fechada (*aldeia*, *baleia*) será necessário, futuramente, um estudo de caso mais aprofundado (por exemplo, considerar a sequência de caracteres anterior ao ditongo) para especificar no formato de regras de transcrição os casos em que a sequência gráfica <ei> deve ser transcrita foneticamente como [ɛi] ou como [ei].

4. Emprego (ou ausência) do hífen em algumas palavras compostas.

O hífen é um sinal gráfico que serve tanto à ligação de palavras compostas (*guarda-chuva*, *ex-presidiário*) quanto para unir pronomes átonos a verbos (*dar-lhe-ei*, *dar-me*). Com o AO90, houve algumas mudanças em relação ao seu uso. Algumas palavras compostas do PB perderam o emprego do hífen (*antirrevolucionário* e não mais *anti-revolucionário*, *contrassenha* e não mais *contra-senha*), enquanto que em outras palavras seu uso foi mantido, tais como em palavras prefixadas iniciadas por h (*anti-higiênico*) ou quando a última letra do prefixo é igual à primeira letra da palavra prefixada (*contra-atacar*).

Devido ao fato de o emprego do hífen ter sido abolido em palavras prefixadas (por exemplo, *intermunicipal*, *coeducar*) tornou-se ainda mais complexa a criação de regras de transcrição específicas para determinar o valor sonoro de vogais médias (e, o) em prefixos, uma vez que os prefixos possuem processos fonológicos específicos.¹⁶ Por exemplo, antes da exclusão do hífen, poderíamos estabelecer a seguinte regra para o prefixo auto-: a sequência <auto->, quando em início de palavra, deve ser transcrita como [ˈautɔ]. Entretanto, após o AO 90, essa regra se torna inválida, uma vez que o PETRUS poderá se deparar com a sequência gráfica <auto> da palavra <automóvel> e transcrevê-la erroneamente. No Capítulo 2 são dados mais detalhes sobre os casos de prefixos constituídos por vogais médias que são tratados pelo sistema PETRUS.

¹⁶ A esse respeito, Marcato (2010, p. 45), após analisar os prefixos com vogal média na variedade rio-pretense, chegou a conclusão de que “os prefixos apresentam comportamentos morfofonológicos distintos a depender, em termos dos domínios prosódicos, sobretudo, da informação sobre forma livre versus forma presa e da classificação dos prefixos quanto ao acento em átonos e tônicos. Com base nesses critérios, verificamos, em nossa análise, que os prefixos classificados como palavras fonológicas independentes, quando exercem a função de forma livre, estão sujeitos à aplicação dos fenômenos fonológicos de *sândi externo*, *tapping* e *alçamento vocálico* da postônica final. Quando os prefixos funcionam como forma presa, sílabas pretônicas, estão sujeitos aos fenômenos de *neutralização vocálica*, *alçamento vocálico da pretônica* e *harmonização vocálica*. Por fim, verificamos, quanto aos prefixos inacentuados, portanto, somente considerados como formas presas, que ocorrem os fenômenos fonológicos de *vozeamento intervocálico da fricativa*, *epêntese*, *assimilação da nasal*, *harmonização vocálica* e *alçamento vocálico*”.

1.1.3 Fonética

Segundo Cagliari (2001, p.42), a Linguística é o estudo científico da linguagem. Está voltada para a explicação de como a linguagem humana funciona e de como são as línguas em particular, quer fazendo o trabalho descritivo previsto pelas teorias, quer usando os conhecimentos adquiridos para beneficiar outras ciências e artes que usam, de algum modo, a linguagem falada ou escrita. Os linguistas organizam o estudo da linguagem em certo número de áreas, que são estudadas mais ou menos independentemente. As divisões mais comuns entre as áreas são as seguintes: fonética, fonologia, morfologia, sintaxe, semântica, lexicologia, terminologia, estilística, pragmática e filologia. Nem todos os linguistas concordam que todas essas divisões tenham grande significado ou que estejam corretas, pois essas divisões se sobrepõem consideravelmente. Por exemplo, a divisão gramática usualmente cobre fonologia, morfologia e sintaxe. Ainda existem campos como os da linguística teórica e da linguística histórica. Há duas disciplinas, em específico, dedicadas ao estudo do nível sonoro da linguagem: a Fonética e a Fonologia. São ciências relacionadas, pois apresentam o mesmo objeto de estudo, embora considerados sob perspectivas diferentes.

A Fonética é uma disciplina antiga. Estudos indicam que a Fonética tenha surgido por volta do século IV a.C., na Índia, com Panini (o qual estima-se ter vivido entre os anos de 520 a 460 a.C.), quando apresentou reflexões sobre estudos gramaticais do sânscrito (antiga língua sagrada indiana e descoberta somente no século XVIII). O estudo de Panini é considerado a primeira gramática de uma língua da civilização humana. No Brasil, os estudos fonéticos foram implementados pela primeira vez em 1915, no Maranhão, por Franco de Sá.

O *Dicionário de Linguística*, de Dubois (1993, p. 282-284), define Fonética como o estudo dos sons da língua em sua realização concreta, “independentemente de sua função linguística”; enquanto que a Fonologia seria o estudo dos sons do ponto de vista funcional, ou seja, “sua função no sistema de comunicação linguística”.

Segundo Cagliari (2002, p. 17), a Fonética preocupa-se principalmente com a descrição dos fatos físicos que caracterizam linguisticamente os sons da fala, descrevendo quais mecanismos e processos de sua produção constituem um determinado segmento da cadeia sonora. A respeito da Fonologia, esse mesmo autor diz que esta visa interpretar os resultados apresentados pela Fonética, em função dos sistemas de sons das línguas e do modelos teóricos que existem para descrevê-los.

Assim, a Fonética descreve os aspectos articulatórios, acústicos e as propriedades físicas de todos os sons que ocorrem na produção linguística, enquanto que a Fonologia analisa as

diferenças fônicas intencionais, distintivas, ou seja, as que se vinculam a diferenças de significado, estabelecendo as relações entre os elementos de diferenciação, bem como as condições em que se combinam para formar morfemas, palavras e frases (CALLOU e LEITE, 2003, p. 11).

A unidade de estudo da Fonética é o *fone*, enquanto que da Fonologia é o *fonema*. O termo “fone” é utilizado em Fonética para designar o menor segmento discreto perceptível de som em um fluxo da fala (*continuum* fônico ou substância fônica). Do ponto de vista da Fonologia segmental, os fones são a realização física dos fonemas (CRYSTAL, 2008, p. 361).

Os ramos de investigação da Fonética subdividem-se em:

- *Fonética Acústica*: analisa as características físicas dos sons da fala, ou seja, seu foco de interesse está nos “efeitos acústicos da onda sonora produzida pela corrente de ar em sua passagem pelo aparelho fonador” (CALLOU e LEITE, 2003, p. 15). Em síntese, estuda a transmissão dos sons. Para as análises dessa área, normalmente realizadas em laboratórios especializados, são utilizadas várias técnicas e aparelhos.
- *Fonética Auditiva*: para Callou e Leite (2003, p. 15), a fonética auditiva se ocupa do “estudo das impressões acústicas e de suas interpretações no processo de decodificação”, isto é, estuda os efeitos físicos que os sons da fala provocam no ouvido do destinatário. Fazem parte de seu domínio o modo pelo qual o aparelho auditivo capta os sons e os mecanismos cerebrais para a sua decodificação.
- *Fonética Instrumental*: compreende o estudo das propriedades físicas da fala, levando em consideração o apoio de instrumentos laboratoriais (CRYSTAL, 2008, p. 363).
- *Fonética Articulatória*: também chamada de fisiológica, estuda como os sons são produzidos, isto é, faz parte de seu campo de atuação o estudo da produção dos sons pelo aparelho fonador do falante.¹⁷ A caracterização que os sons recebem está relacionada a certos pontos de nossa anatomia (pontos de articulação) e à maneira como o ar passa por esses pontos (modos de articulação).

Os diferentes modos pelos quais o fluxo de ar é modificado permitem estabelecer duas grandes classes de sons: as consoantes e as vogais. As vogais são produzidas com o fluxo de

¹⁷ O aparelho fonador é formado por um conjunto de órgãos, que podem ser agrupados em três conjuntos: o primeiro composto pelos pulmões, brônquios e traqueia (órgãos respiratórios responsáveis por produzir a corrente de ar necessária à fonação), o segundo pela laringe e cordas vocais (que produzem a vibração utilizada na fala) e, por fim, o terceiro conjunto de órgãos, composto pela faringe, boca e fossas nasais (que funcionam como caixas de ressonância, pois são responsáveis pelos diversos sons da linguagem). Considerando uma perspectiva anatômica e biológica, não é adequado dizer que existe um aparelho fonador. O equívoco se dá pelo fato de não existir órgãos única e exclusivamente destinados à produção do som. O que existe é uma anatomia integrada para essa finalidade.

ar passando livremente, sem obstruções, pelo trato vocal. Já as consoantes são articuladas com alguma obstrução no trato oral, seja ela parcial ou total. Do ponto de vista fonético, as vogais são analisadas segundo os parâmetros de altura, avanço ou recuo da língua e arredondamento dos lábios. Em relação as consoantes, utilizam-se as características de ponto (lugar) e modo (maneira) de articulação, bem como de sonoridade.

Vejamos, a seguir, a classificação das consoantes em PB, conforme os critérios citados.

1.1.3.1 Consoantes

Denominamos *consoantes* os sons da fala produzidos com obstrução parcial ou total à passagem de ar da cavidade bucal. O seu *Modo (ou Maneira) de Articulação* está relacionado ao tipo e ao grau de obstrução da corrente de ar causada pelos articuladores, durante a produção de um segmento consonantal. De acordo com esse critério, podemos definir os seguintes tipos consonantais:

- *Oclusivas (ou Plosivas)*: consoantes produzidas com bloqueio total da passagem de ar, desde as cordas vocais até o exterior da boca. Quando essa barreira se rompe bruscamente, dá-se um ruído semelhante a uma explosão. São oclusivas os pares [p] e [b], nos quais a obstrução é feita com os lábios; [t] e [d], que é feita na língua colocada nos alvéolos; [k] e [g], com o dorso da língua colocado no final do palato mole.¹⁸

- *Fricativas (ou Constrictivas)*: consoantes produzidas com obstrução parcial da corrente de ar, que ao passar pelos articuladores produz uma fricção. Os fones [f], [s], [ʃ] são classificados como consoantes fricativas surdas, enquanto que [v], [z], [ʒ] são denominadas consoantes fricativas sonoras.

- *Africadas*: são aquelas que na fase inicial de sua produção seus articuladores promovem uma obstrução completa na passagem de corrente de ar através da boca, ficando o véu palatino levantado, como nas oclusivas, e na fase final dessa obstrução (na soltura da oclusão) ocorre fricção decorrente da passagem central da corrente de ar. (CRISTÓFARO-

¹⁸ Cada um desses pares possui o mesmo modo de articulação e os mesmos articuladores. O que os diferencia, portanto, é o papel desempenhado pelas pregas vocais. Na produção do primeiro som de cada par ([p], [t], [k]), as pregas vocais estão abertas, e não vibram (consoantes surdas), enquanto que para formar os segundos ([b], [d], [g]), há vibração das pregas quando o ar que sai dos pulmões as atravessa (consoantes sonoras). As oclusivas nunca ocorrem em posição de coda silábica, apesar de a ortografia nos deixar a impressão de que poderia ser possível ocorrer uma oclusiva em final de sílaba. Entretanto, no PB, a regra geral para esses casos é a de inserir uma vogal epentética [i] para não permitir que uma vogal oclusiva fique em posição de coda (cf. *objeto* [obr'ʒeto], *pacto* ['pakto], *submarinho* [subima'riɲo], *apto* ['apto], *advogado* [adʒivo'gado]).

SILVA, 2000, p. 33). Em PB, as africadas não têm valor distintivo, mas aparecem como variação de dois fonemas em contextos específicos, como é o caso dos fonemas /t/ e /d/, os quais, diante de [i], [ɪ] e [ĩ] podem se realizar como [tʃ] [dʒ], em algumas variedades linguísticas do PB, como a paulista e a carioca.

- *Laterais*: são produzidas quando o articulador ativo (língua) toca o articulador passivo (alvéolo no caso do [l] e palato no caso do [ʎ]), obstruindo a corrente de ar na linha central do trato vocal, sendo o ar expelido por ambos os lados dessa obstrução.

- *Nasais*: são produzidas com o abaixamento do véu palatino, permitindo que a corrente de ar, ou parte desta, escape pelo nariz. Esse abaixamento do véu palatino altera a área de circulação da corrente de ar no trato vocal (nas cavidades oral e nasal) e a qualidade vocálica. As consoantes nasais do português são [m], [n], [ɲ], presentes, respectivamente, em: *mama* ~ [mẽma], *mana* ~ [mẽna] e *manha* ~ [mẽɲa].

- *Róticas*: segundo Ladefoged e Maddieson (1996, p. 215), são os sons representados ortograficamente por “r” (sendo uma ocorrência comum na maior parte das línguas do mundo), e podem ser produzidos por meio de diferentes modos de articulação, tais como: vibrantes (*trills*), tepes (*taps* ou *flaps*), fricativas e aproximantes; são articulados em vários pontos do aparelho fonador, desde a glote até a ponta da língua. No PB podemos observar a ocorrência dos seguintes tipos de róticos: tepe simples [r], vibrante [r], fricativas [x, ɣ, h, ɦ] e a aproximante retroflexa [ɽ].¹⁹

¹⁹ Toda essa variedade de róticos surge, primeiro, relacionada à variedade do português, visto que no Brasil, há enorme variedade dialetal; em segundo, relaciona-se ao contexto linguístico em que ocorre. No PB observa-se a ocorrência de róticos nos seguintes contextos: se manifesta foneticamente como um tepe [r] ou vibrante simples em todos os dialetos do português (CRISTÓFARO-SILVA, 2000, p. 160; MONARETTO et al., 1996, p. 218) seguindo consoante tautossilábica, ou seja, na mesma sílaba (Cf. *prato*) ou posição intervocálica (Cf. *arara*). Alternativamente, pode ser produzido em início de palavra (Cf. *rato*), em posição intervocálica (Cf. *carro*) ou seguindo consoante em sílaba distinta (Cf. *Israel*), variando consideravelmente conforme a região, como fricativa velar [x] no Rio de Janeiro, como fricativa glotal [h] em Belo Horizonte e como a retroflexa alveolar [ɽ] que é típica apenas em final de sílaba do dialeto caipira falado no interior de São Paulo, sul de Minas Gerais, (CRISTÓFARO-SILVA, 1998, p.38-39; MONARETTO et al., 1996, p. 215) leste do Mato Grosso do Sul, sul de Goiás e norte do Paraná. Sua delimitação e caracterização datam de 1920, com a obra “O Dialeto Caipira”, de Amadeu Amaral (BAGNO, 2011, p. 245). Em posição de coda silábica no PB, podem ocorrer as seguintes possíveis realizações das vibrantes: fricativa velar surda e sonora ([x, ɣ]), fricativa uvular surda e sonora ([χ, ʁ]), glotal surda e sonora ([h, ɦ]), vibrante alveolar sonora ([r]), vibrante uvular sonora ([ʀ]), tepe alveolar sonoro ([r]), aproximante alveolar sonora ([ɹ]) e aproximante retroflexa sonora ([ɽ]). Conforme descrito na literatura (CALLOU et al, 1997; SILVA, 2002; BARBOSA e ALBANO, 2004), há indicação da produção da variante fricativa em falantes de Rio de Janeiro (RJ), Belo Horizonte (MG) e cidades da região Nordeste do país, como Salvador (BA) e Recife (PE). Já a variante tepe é encontrada em São Paulo (SP) e Porto Alegre (RS); a variante vibrante pode ser observada em algumas regiões de São Paulo (SP). A variante retroflexa é verificada na fala de sujeitos de cidades do interior, caracterizando o falar “caipira”. O dialeto caipira poderia ainda ser dividido em cinco subdialetos: (i) o falado na região sul do estado de São Paulo (regiões do Vale do Ribeira, Sorocaba, Itapetininga e Itapeva); (ii) o subdialeto das regiões do Médio Tietê (Campinas, Piracicaba, Capivari, Porto Feliz, Itu, Santa Bárbara d'Oeste, Americana, Limeira, Rio Claro, São Carlos, Araraquara e Jaú); (iii) o subdialeto que compreende as regiões norte, nordeste e noroeste (Ribeirão Preto, Franca, São José do Rio Preto)

Na variedade linguística aqui adotada, admitimos o tepe alveolar sonoro [r] para representar foneticamente o ‘r’ ortográfico em posição intervocálica (*caro* ['karu]) e seguindo consoante na mesma sílaba (*praga* ['praga]) e a fricativa velar surda [x] foi eleita para os seguintes contextos linguísticos de ‘r’ e ‘rr’: início de palavras (*rato* ['xatu]), em posição intervocálica (*carro* ['kaxu]), seguindo consoante em sílaba subsequente (*Israel* [isxa'ɛw]), finais de palavras (*estudar* [istu'dax]), finais de sílaba que antecedem uma consoante surda (*carta* ['kaxta]). Já a fricativa velar sonora [ɣ] foi utilizada nos finais de sílaba que antecedem uma consoante sonora, como em *verde* ['veydʒɪ]. Futuramente, com base em estudos específicos sobre a produção dos róticos na variante linguística adotada, decidiremos se as fricativas velares [x, ɣ] ou as fricativas glotais [h, h̃] são as mais adequadas para serem adotadas em nosso sistema.

O *Lugar (ou Ponto) de Articulação* é o ponto da cavidade bucal no qual se localiza o obstáculo para a passagem de ar. Conforme o ponto de articulação, as consoantes podem ser (CRISTÓFARO-SILVA, 2000):

- *Bilabiais*: há contato entre os lábios superior e inferior durante a formação de uma consoante. São bilabiais as consoantes [p], [b], [m].
- *Labiodentais*: quando o lábio inferior tem contato com os dentes incisivos superiores. É o caso das consoantes fricativas [f] e [v].
- *Dentais*: sons produzidos quando a ponta da língua é colocada atrás dos dentes incisivos, antes do alvéolos. Os fones [t], [d], [n], [l] são dentais, mas também podem ser realizados alveolarmente.
- *Alveolares*: sons produzidos com a língua colocada na parte que fica atrás dos dentes, os alvéolos. É o caso dos fones [s], [z], [r].
- *Alveolopalatais*: são representadas pelas africadas (consoante oclusiva seguida de uma fricativa) [tʃ], [dʒ], que não são fonemas em português, mas sim alofones²⁰ dos fonemas [t] e [d], respectivamente.

e oeste (Araçatuba, Presidente Prudente, Bauru, Marília, Lins) do estado de São Paulo, região sudoeste de Minas Gerais (São Sebastião do Paraíso, Cássia (Minas Gerais), Passos (Minas Gerais)), além de parte do estado do Paraná e de Mato Grosso do Sul; (iv) o subdialeto da Mesorregião do Vale do Paraíba Paulista (Taubaté, Guaratinguetá) e Litoral Norte (São Paulo); e, por último, o subdialeto falado no Triângulo Mineiro e a maior parte do estado de Goiás. Fonte: Wikipedia. Disponível em: <<http://pt.wikipedia.org/wiki/Dialetocaipira#endnoteBagn01>>. Acesso em: 25 de nov. 2014.

²⁰ Os alofones, que correspondem às realizações de um fonema, não distinguem significado. Quando as variantes de fonema dependem do ambiente fonético em que o som vocal se encontra, assimilam-se os traços dos sons contíguos. Em Fonologia, o que distingue o fonema do alofone é a capacidade do fonema de distinguir os significados da língua (CÂMARA JR, 1986).

- *Palatais*: são fonemas produzidos com o dorso da língua encostado no palato duro (céu da boca), como em: [ʃ], [ʒ], [ɲ] e [ʎ].
- *Velares*: são consoantes produzidas com o dorso da língua em direção ao palato mole, como em: [k], [g].
- *Uvulares*: são produzidas na parte final da boca, o dorso da língua vai em direção à úvula, no palato mole: [x].

A *Função das Cordas Vocais* permite observar se há ou não vibração no momento de pronúncia das consoantes. Ao ocorrer vibração das cordas, a consoante é classificada como *vozeada* (sonora), caso contrário, será denominada *desvozeada* (surda). Os pares de consoantes [p, b]; [t, d]; [k, g]; [s, z]; [f, v]; [ʃ, ʒ]; [ʎ, ʒ] são homorgânicos, isto é, possuem o mesmo modo e ponto de articulação, sendo a única diferença entre eles, o vozeamento ou não das cordas vocais. Vale dizer que as consoantes nasais e laterais também são vozeadas, assim como as vibrantes, apesar de não constarem nos pares acima.

1.1.3.2 Vogais

As vogais não são classificadas em função de modos e pontos de articulação como acontece com as consoantes, pois as vogais possuem articulação livre, isto é, não há obstáculo à passagem de corrente de ar quando de sua produção vocal. Articulatoriamente, as vogais se distinguem das consoantes devido aos seguintes fatores: i) a vogal não tem ponto de articulação fixo, já que é classificada segundo a posição da língua na boca (no sentido horizontal e no vertical); ii) a forma dos lábios distingue vogais arredondadas e não-arredondadas; iii) são sonoras e fonologicamente ocupam sempre o núcleo da sílaba; iv) a passagem do ar é livre: não há nenhuma obstrução à corrente de ar; e v) são pronunciadas com maior duração e tonicidade que as consoantes (WEISS, 1988, p. 24).

Ainda que o método das *Vogais Cardeais* seja um recurso descritivo preciso para a caracterização de sistemas vocálicos,²¹ é bastante comum realizar na literatura o agrupamento de vogais segundo uma classificação articulatória na qual estão envolvidos o corpo da língua e os lábios, tais como: arredondamento ou não dos lábios; anterioridade, posterioridade ou

²¹ Método das Vogais Cardeais. Trata-se de um sistema descritivo e classificatório para as vogais. Para mais informações, vide Cristófar-Silva (1999) e Cagliari (1981, 2009).

centralidade da posição da língua e altura da língua, que pode ser organizada em três ou quatro grupos segundo o sistema vocálico em questão (CRISTÓFARO-SILVA, 1999).

Com relação à altura da língua, no PB, as vogais podem ser classificadas em:

- *Alta*. Aquela em que o dorso da língua se eleva ao máximo, estreitando o trato vocal, mas sem produzir fricção: [i, u].
- *Média-alta*. Aquela em que o dorso da língua encontra-se em uma posição intermediária, entretanto, em uma posição mais próxima da posição *alta*: [e, o].
- *Média-baixa*. Aquela em que o dorso da língua encontra-se em uma posição intermediária, no entanto, em uma posição mais próxima da posição baixa: [ɛ, ɔ].
- *Baixa*. Aquela em que a língua se encontra na posição mais baixa no trato oral: [a].

Conforme o avanço ou recuo da língua (a posição do corpo da língua em relação à abóbada palatal), as vogais podem ser classificadas em:

- *Anterior*. Aquela em que a língua avança para frente, no palato: [ɛ, e, i].
- *Posterior*. Aquela em que o dorso da língua se inclina para trás, em direção ao palato mole: [ɔ, o, u].
- *Central*. Aquela em que a língua não se movimenta, fica no centro da boca: [a]. Na pronúncia da vogal [a], a língua está abaixada e um pouco mais avançada do que para a produção da vogal [ɐ].

Segundo a posição dos lábios, as vogais podem ser classificadas em:

- *Arredondadas*. Vogais produzidas com os lábios arredondados: [u], [o] e [ɔ]. As vogais arredondadas também são denominadas de labializadas.
- *Não-arredondadas*. Vogais produzidas com os lábios distendidos: [ɛ, e, i, a].

Há autores, como Malmberg (1954), que têm classificado as vogais também em função da abertura/fechamento do trato oral. Assim, as vogais são classificadas em:

- *Fechada*. O lábio inferior sobe: [i, u];
- *Meio-fechada*. O lábio desce um pouco: [e, o];
- *Meio-aberta*. O lábio desce mais um pouco: [ɛ, ɔ];
- *Aberta*. O lábio fica baixo: [a].

1.1.3.2.1 Vogais Orais

De acordo com Massini-Cagliari e Cagliari (2001), as vogais orais são caracterizadas por serem produzidas pelo ar que, expelido dos pulmões, faz vibrar as cordas vocais, sem encontrar qualquer tipo de obstáculo durante sua passagem pelo aparelho fonador.

Existem sete fonemas vocálicos que podem estar presentes em sílaba tônica (CALLOU; LEITE, 2003, p. 79), os quais possuem distribuição homogênea em todas as variedades do PB (CRISTÓFARO-SILVA, 2000, p. 79).

A Tabela 1 traz uma lista das vogais orais tônicas presentes no sistema vocálico brasileiro, representadas pelo sistema de notação fonético IPA, seguida de uma palavra-exemplo e sua respectiva transcrição fonética:

Tabela 1. As sete vogais tônicas orais existentes no PB

IPA	Exemplo	Transcrição fonética
[a]	<s <u>a</u> la>	[ˈsala]
[e]	<se <u>l</u> o>	[ˈselʊ]
[ɛ]	<be <u>l</u> o>	[ˈbɛlʊ]
[i]	<si <u>l</u> o>	[ˈsilʊ]
[o]	<bo <u>l</u> o>	[ˈbolʊ]
[ɔ]	<so <u>l</u> o>	[ˈsɔlʊ]
[u]	<bu <u>l</u> a>	[ˈbula]

Esse sistema vocálico completo só existe em sílaba tônica, pois em sílabas átonas ocorre o que se convencionou chamar, dentro da linguística estrutural europeia, de *neutralização*.²² Em decorrência da neutralização, o sistema de vogais é reduzido para cinco em posição pretônica: [a, e, i, o, u], sendo a ocorrência das vogais [ɛ] e [ɔ], nessa posição, marca de variação dialetal geográfica característica das regiões norte e nordeste do Brasil, ou mesmo de idioleto, conforme afirmam Cristófaros-Silva (2000, p. 81), Antenor Nascentes (1953) e Câmara Jr. (1970).

²² *Neutralização*: quando o contraste entre fonemas é perdido em certas posições de uma palavra. Com efeito, o fonema só é entendido como entidade autônoma significativa por que se distingue de outro dentro do mesmo sistema fonológico. Assim, o /d/ é um fonema sonoro do português por que a ele opõe-se o correspondente surdo /t/. Há situações, entretanto, em que a distinção entre fonemas deixa de ser um fato relevante para o falante da língua. Por exemplo, a palavra *paz*. A pronúncia dessa palavra nas diferentes regiões geolinguísticas brasileiras varia devido à consoante final: [paz], [pas], [paʃ]. Verifica-se, pois, que os fones [z], [s] e [ʃ], em final de sílaba, perdem distinção entre si; esse fato é denominado de *neutralização de traços distintivos*.

Na Tabela 2 consta uma lista das vogais orais pretônicas presentes no sistema vocálico do PB, representadas pelo sistema de notação fonético IPA, seguida de uma palavra-exemplo e sua respectiva transcrição fonética:

Tabela 2. As vogais pretônicas orais existentes no PB.

<i>IPA</i>	<i>Exemplo</i>	<i>Transcrição fonética</i>
[a]	<amor>	[a'mox]
[e]	<segredo>	[se'gredɔ]
[i]	<igreja>	[i'greʒa]
[o]	<bolacha>	[bo'laʃa]
[u]	<cueca>	[ku'eka]

Sem dúvida, uma das riquezas do PB é a gama de diferentes realizações do sistema vocálico em pauta pretônica, com predominância das realizações abertas ou fechadas de acordo com cada região de nosso país.

A variação dialetal²³ relacionada à realização das vogais pretônicas estabelece uma linha divisória entre os diferentes falares de nosso país, tema já abordado por Nascentes (1953). A maioria dos estudos realizados sobre o PB indica que nas regiões Sul/Sudeste prevalece a opção pela pronúncia fechada [e] e [o], e no Norte/Nordeste, a realização mais aberta [ɛ] e [ɔ] (por exemplo, as palavras <coração> e <elevado> seriam pronunciadas como [kora'sẽõ] ou [kɔra'sẽõ] e [ele'vado] ou [ɛle'vado], respectivamente, conforme a região geográfica considerada esteja mais ao sul/sudeste ou mais ao norte/nordeste do país).²⁴ Essa divisão dialetal, em duas grandes regiões, admitiria ainda, segundo Nascentes (1981), quatro subdivisões:

Talvez possamos admitir quatro subdialetos: o nortista (Amazonas, Pará, litoral dos Estados desde o Maranhão até a Bahia), o fluminense (Espírito Santo, Rio de Janeiro, sul de Minas e zona da mata, Distrito Federal), o sertanejo (Mato Grosso, Goiás, norte de Minas, sertão dos Estados litorâneos desde o Maranhão à Bahia) e o sulista (São Paulo, Paraná, Santa Catarina, Rio Grande do Sul e triângulo mineiro). (NASCENTES, 1981, p.117)

²³ Variação dialetal: As “línguas variam em razão de condicionamentos situacionais que afetam os falantes, tais como o momento histórico em que se acham, o espaço geográfico, social e temático em que se movem. O conjunto dessas circunstâncias interage sobre os fatores da comunicação e daqui fluem variantes linguísticas” (CASTILHO, 1978, p. 13). Mais detalhes sobre variação dialetal na seção 1.1.5 desta tese.

²⁴ Por outro lado, tem-se a harmonia vocálica, já analisada por muitos linguistas que a observaram na fala de vários estados brasileiros, que eleva de forma variável a vogal média diante de uma vogal alta [i,u], como em qu[e]rido ~ qu[i]rido, c[o]ruja ~ c[u]ruja. Conjuntamente a esse processo, temos outros casos de elevações, denominadas por Bisol (2009) como “sem motivação aparente” e que se apresentam em um pequeno grupo de palavras, como em c[o]légio ~ c[u]légio, b[o]neca ~ b[u]neca e p[e]queno ~ p[i]queno.

Todos os falantes do PB também apresentam as vogais médias abertas [ɛ, ɔ] em posição pretônica em formas derivadas com os sufixos -mente, -inh, -zinh, -íssim, quando o radical do substantivo/adjetivo apresenta [ɔ, ɛ] em posição tônica (CRISTÓFARO-SILVA, 2000, p. 81-82). Por exemplo: *mole* ['mɔli], *molezinho* [mɔli'ziɲɔ].

A pronúncia das vogais médias pretônicas no PB é tema recorrente em trabalhos, como os de: Bisol (1981), sobre a regra de harmonia vocálica no falar gaúcho; Callou; Leite (1986), sobre a ação da regra de harmonia vocálica na norma culta do Rio de Janeiro; Maia (1986), sobre o comportamento das vogais pretônicas médias na fala de Natal; Viegas (1987; 2001), sobre o alçamento das vogais médias pretônicas na região de Belo Horizonte, em Minas Gerais; Silva (1989), sobre as vogais pretônicas na variedade culta de Salvador, na Bahia; Bortoni e Malvar (1992), sobre os condicionamentos das regras de elevação e abaixamento no dialeto de Brasília; Battisti (1993), sobre a elevação das médias pretônicas em sílaba inicial de vocábulo na fala gaúcha; Freitas (2001), sobre as vogais médias pretônicas na cidade de Bragança, no Pará; Schwindt (2002), sobre a regra de harmonização vocálica no Rio Grande do Sul, Celia (2004) estudou as vogais médias pretônicas de Nova Venécia-ES e Marques (2006), sobre as vogais médias pretônicas em situação de contato dialetal entre paraibanos e cariocas no Rio de Janeiro e entre brasileiros e portugueses na cidade de Lisboa.

Além de serem consideradas um fator que diferencia os dialetos brasileiros, as realizações do sistema vocálico em pauta pretônica também caracterizam diferenças entre o PB e o PE;²⁵ sendo, talvez, uma das características que mais diferenciam as duas variedades, conforme ressalta Teyssier (1997).

Sendo assim, para o estabelecimento de regras de transcrição para o sistema PETRUS, referentes ao timbre das vogais pretônicas, manteve-se aqui a qualidade meio-fechada [e] e [o] sempre que a ortografia das palavras apresentasse as letras <e> e <o> em posição pretônica (Cf. *chacoalhar* [ʃakoa'kax], *enseada* [ẽise'ada]), assim como realizado em Cagliari (2011; 2012). Como trabalho futuro, pretendemos implementar regras para transcrição fonética das vogais médias pretônicas em contextos nos quais o alçamento dessas vogais é, de acordo com Bisol (1981), categórico. São eles: o contexto de hiato (“b[u]emia”), prefixo (“d[i]sgoverno”) e da vogal <e> em início de palavra sucedida por <s> ou <x> (“[i]screve”; “[i]xclui”).

²⁵ Conforme aponta Bisol (1981), nossa herança tomou rumos diferentes daqueles tomados pelas pretônicas em Portugal em cujo dialeto as posteriores tornam-se altas enquanto as anteriores são produzidas com /e/ mudo. Nesse sentido, segundo Teyssier, (1997 apud VIEGAS, 2001, p. 75) a pronúncia brasileira perpetua a pronúncia do PE antes de mutações que nele ocorreram no século XVIII.

Por fim, temos as *vogais postônicas orais* que podem ser agrupadas em *vogais postônicas finais* e *vogais postônicas não-finais*.

É sabido que tanto [e] quanto [o] postônico medial podem sofrer redução de [e] para [ɪ] e de [o] para [ʊ], respectivamente, em algumas palavras. Sobre esse assunto, Cagliari (2012) afirma que não se trata de uma regra geral presente em todas as palavras pronunciadas conforme o dialeto por ele descrito. Devido ao fato de as decisões aqui tomadas dependerem sobremaneira da possibilidade de sistematização, no formato de regras linguísticas, decidimos, assim como Cagliari (2012), realizar a transcrição fonética das vogais postônicas não finais <a, e, i, o, u> com o uso dos mesmos símbolos vocálicos utilizados para representar foneticamente as pretônicas [a, e, i, o, u] (ex.: *último* ['uotʃimo], *ímpeto* ['ĩpeto], *pérola* ['perola], *óculos* ['okulos]), mesmo nos casos em que poderia haver uma variação com um ditongo (cf. *demônio* [de'mõnio], *mútua* ['mutua]). Desse modo, diminuimos possíveis erros que poderiam ocorrer na transcrição de certas palavras, como por exemplo, <mísera>, que poderia ter o <e> transcrito inadequadamente como [ɪ] (mís[ɪ]ra) se adotássemos aqui um conjunto reduzido de vogais, no qual <e> se tornaria [ɪ] e <o> se tornaria [ʊ] sempre que presentes em posição postônica. Como trabalho futuro, pretendemos investigar os contextos linguísticos em que <e> e <o> não sofrem redução vocálica e sistematizar no formato de regras linguísticas para serem implementadas no sistema PETRUS.

Na Tabela 3 são apresentadas as vogais postônicas orais não-finais presentes no sistema vocálico aqui adotado, representadas pelo sistema de notação fonético IPA, seguida de uma palavra-exemplo e sua respectiva transcrição fonética:

Tabela 3. As vogais postônicas não-finais existentes no PB.

IPA	Exemplo	Transcrição fonética
[a]	<sílaba>	['silaba]
[e]	<mísera>	['mizera]
[i]	<tráfíco>	['trafikʊ]
[o]	<pérola>	['perola]
[u]	<cédula>	['sedula]

De acordo com Cagliari (2012, p. 481), as vogais finais postônicas são [ɪ, a, ʊ] para a maioria dos falantes do PB, devido à perda do traço que possibilita distinguir as vogais médias entre si. Assim, as vogais postônicas finais <e, i> e <o, u> se realizam foneticamente como [ɪ] e [ʊ], como podemos observar nas palavras *dela* ['dɛla], *dele* ['dɛlɪ], *dedo* ['dɛdʊ]. Entretanto, em algumas variantes de nosso idioma, como acontece, por exemplo, na pronúncia

tradicional de Curitiba ou no interior do Rio Grande do Sul, podemos observar um conjunto um pouco maior de vogais na posição final: [a, e, i, o].

Em PB, há divergência também quanto à representação fonética da vogal <a> postônica final. Para Cagliari (2009, p. 59; 2011, p. 4), a opção pelo uso de [ə, ɐ] em vez de [a] para representar a vogal <a> em posição átona final acontece apenas com o português europeu. A respeito da representação da vogal <a> em posição átona, Cagliari (2009, p. 113-114) chama a atenção para o fato de que não se pode confundir variação de tonicidade com variação da qualidade vocálica. Para ilustrar essa questão, toma como exemplo a palavra <casa> ['kaza], que possui duas vogais iguais quanto a qualidade, porém, ressalva, que a primeira é tônica e a segunda é átona. Acrescenta que a primeira é mais longa que a segunda do ponto de vista perceptivo, mas a segunda é fisicamente mais longa que a primeira. Diz ainda que a primeira tem uma intensidade acústica maior e relativamente constante, enquanto que a segunda começa com uma intensidade acústica que vai se reduzindo progressivamente. Por fim, o autor afirma que a primeira vogal é falada com uma altura melódica em geral diferente da altura melódica da segunda. Ao final, Cagliari conclui que todos esses fatos contribuem para que a primeira vogal seja diferente foneticamente da segunda, mas nos chama a atenção para o fato de que, devido à qualidade vocálica (timbre vocálico) ser a mesma em ambas as vogais, deve-se transcrever as duas com o mesmo símbolo fonético: [a].

Por outro lado, Barbosa (2012) realizou um estudo que objetivou caracterizar acusticamente a vogal /a/ postônica final no PB. Para tanto, utilizou dados de oito falantes oriundos de cinco estados brasileiros: 8 falantes universitários (4 homens e 4 mulheres) entre 25 e 35 anos na época da gravação, assim distribuídos: um homem e uma mulher do interior de Minas Gerais; um homem e uma mulher de São Paulo, sendo a mulher da capital e o homem do interior; um homem e uma mulher do Ceará, ambos de Fortaleza, e um homem do Piauí e uma mulher do Pará. Ao final, concluiu que o símbolo mais indicado para transcrever foneticamente a vogal <a> em posição postônica final é [ɐ], que representa o som de uma vogal entre a articulação semiaberta e aberta, em posição central. O [ɐ] denota um som que é um pouco menos aberto do que [a] e um pouco mais aberto que [ə].

Para o desenvolvimento do PETRUS, devido ao fato de não haver uma investigação acústica sobre o <a> postônico final na variedade linguística crida pela Rede Globo, decidimos que a opção pela transcrição de <a> em posição átona como [a] pode ser válida para o nosso sistema, por ora, por se tratar de transcrição fonética de palavras lidas isoladamente. Futuramente, como todos os dialetos do português parecem possuir um [ɐ]

postônico final, e, portanto, o “globês” não poderá fugir a esse padrão, poderão ser realizados estudos acústicos com base em *córpus* de falantes da referida variedade para averiguar se a qualidade dessa vogal é mantida independentemente da posição (pretônica, tônica, átona) que ela ocupa em uma palavra, e poderemos escolher definitivamente o símbolo fonético que melhor representa o <a> postônico final na variedade linguística criada pela Rede Globo.

Na Tabela 4 constam as vogais postônicas orais finais presentes no sistema vocálico do PB, mais especificamente do dialeto globês aqui adotado, representadas pelo sistema de notação fonético IPA, seguida de uma palavra-exemplo e sua respectiva transcrição fonética:

Tabela 4. As vogais postônicas orais finais existentes no PB.

<i>IPA</i>	<i>Exemplo</i>	<i>Transcrição fonética</i>
[a]	<deda>	[ˈdɛda]
[ɪ]	<dele>	[ˈdɛlɪ]
[ʊ]	<dedo>	[ˈdɛdʊ]

O símbolo fonético [ɪ] é utilizado tanto para representar as vogais <e> em posição átona final de palavras (ex.: “tard[ɪ]”), quanto nos casos em que ocorre epêntese vocálica, processo fonológico caracterizado pela inserção de um segmento vocálico em sequências consonantais (consoante oclusiva, nasal bilabial ou fricativa alveolar surda sucedida por uma consoante) do PB (CAGLIARI, 1981; LEE, 1993; COLLISCHONN, 1996; CAGLIARI e MASSINI-CAGLIARI, 2000). Por exemplo: “p[ɪ]neu”, “ab[ɪ]dicar”, “pac[ɪ]to”.

1.1.3.2.2 Vogais Nasais

As vogais nasais são produzidas com o véu palatino abaixado, fazendo com que a corrente de ar passe tanto pela cavidade oral quanto pela nasal. Esse tipo de articulação traz modificações mais acentuadas para algumas vogais do que para outras. Por exemplo, as vogais altas [i] e [u], articuladas com a língua em posição elevada, apresentam um pequeno abaixamento do véu palatino. Dessa maneira, a configuração do trato oral para produção dessas vogais, tanto orais ([i], [u]) quanto nasais ([ĩ], [ũ]) são muito semelhantes.

Segundo Seara (2000), a articulação de vogais nasais que são produzidas com a língua na posição mais baixa necessita de um maior abaixamento do véu palatino para elas soarem como nasais. Nesse caso, há uma diferença bastante grande entre a articulação de uma vogal baixa oral e uma nasal. É o que ocorre, por exemplo, com a vogal baixa central [a]. Devido a

essa diferença articulatória (e portanto acústica), a autora aconselha que a representação de sua contraparte nasal seria mais adequada com o uso do símbolo fonético [ẽ].

Além das propriedades anteriormente apresentadas para a caracterização das vogais (seção 1.1.2.2), é possível ainda caracterizá-las a partir de algumas propriedades articulatórias secundárias, como *duração*, *desvozeamento*, *tensão* e *nasalização*.

Em relação a essa última propriedade articulatória secundária, são denominadas vogais *nasalizadas*, as vogais cujo som se nasaliza em decorrência de seus contextos adjacentes. É o que ocorre em palavras como <cana>, <venho>, <pinho>, <sonho> e <punho>, que são transcritas foneticamente como [ˈkɛ̃na], [ˈvɛ̃ɲo], [ˈpĩɲo], [ˈsõɲo], [ˈpũɲo]. Nessas vogais, o abaixamento do véu palatino para a articulação da consoante nasal adjacente é realizado antes da completa articulação da vogal que antecede esse segmento nasal. Em decorrência disso, tais vogais são percebidas como sendo nasalizadas.

Em contexto tônico, a nasalização²⁶ é mais perceptível do que em contexto átono. Por esse motivo, optamos por transcrever, neste trabalho, apenas os casos de nasalização que ocorrem em contexto tônico.

Ainda quanto à nasalização, Cagliari (1977, 2009, 2011, 2012) afirma que ocorre um elemento nasal imediatamente após a vogal nasal. “O elemento nasal é geralmente homorgânico à consoante seguinte, ou seja, deve apresentar o mesmo lugar de articulação” (CRISTÓFARO-SILVA, 2000, p. 92). Assim, o elemento nasal [ᵐ] precede [p, b], [ⁿ] precede [t, d], [ᵑ] precede [ʃ, ʒ, tʃ, dʒ] e [ᵑ] precede [k, g]. A diferença entre um segmento nasal – por exemplo [ᵐ] – e o elemento nasal a ele correspondente - [ᵐ] – é, sobretudo, o tempo de articulação. O segmento nasal requer mais tempo de articulação do que o elemento nasal homorgânico. Assim, [ˈtoᵐbu], [ˈtɛᵑta] e [ˈkɛᵑga] apresentam uma breve articulação nasal entre a vogal nasal e a consoante seguinte. Em Cagliari (2011; 2012), o elemento nasal não é grafado no canto superior direito da vogal nasal, sendo notados da seguinte forma: [ˈtombu], [ˈtɛnta] e [ˈkɛnga].

Como nosso objetivo é o de que o sistema PETRUS produza uma transcrição fonética larga, com vistas ao seu processamento automático, optou-se por não marcar com esse tipo de detalhamento fonético a produção das vogais nasais. A descrição proposta por Cagliari (2009)

²⁶ Diz-se nasalização, quando a não articulação da vogal nasal causa diferença de significado, como em “lá” e “lã”. É obrigatoriamente um som nasalizado em qualquer dialeto do português. A nasalidade, ao contrário, ocorre quando a não articulação da vogal marca a variação dialetal e não causa diferença de significado, como em “j[a] nela ou j[ã] nela – janela”. Do ponto de vista fonológico, a nasalização é um tipo de processo bastante comum no português. Quase toda vogal tônica que precede uma consoante nasal se nasaliza (SOUZA e SANTOS, 2004, p. 49).

é essencialmente estreita, uma perspectiva de tratamento diferente da nossa, o que realmente traduz uma pronúncia mais detalhada, como se observa no uso dos elementos nasais citados.

Assim, nos algoritmos implementados no PETRUS, ele ([m]) vale como fone em palavras como <amar>, e a nasalização é sempre marcada nesses mesmos algoritmos com o diacrítico [̃] (como em *canto* ['kẽ̃to], que Cagliari, entretanto, transcreve como ['kɐ̃ⁿto]).

Além disso, “a maioria dos autores que trabalha com português adota os símbolos das vogais com til sobreposto para representar vogais nasalizadas” (CRISTÓFARO-SILVA, 2000, p. 91). Adotamos o símbolo [ẽ̃] tanto para vogais nasais quanto para as vogais nasalizadas em contextos fonológicos, como em <amo> ~ ['ẽ̃mɔ] e <canto> ~ [k'ẽ̃tɔ].

Há, ainda, um grupo de palavras em que a não articulação da vogal nasal marca variação dialetal, mas não causa variação de significado, denominada *nasalidade*. É o caso, por exemplo, da palavra <caneta>, que pode ser transcrita como [ka'neta] ou [kẽ̃'neta]. A *nasalidade* de uma vogal acontece quando se tem uma consoante nasal (<m>, <n> ou <nh>) em posição subsequente à de uma vogal oral. A par dos casos de vogais nasais que ocorrem em todos os dialetos do português, as vogais nasalizadas²⁷ variam consideravelmente *de dialeto para dialeto* (sendo a frequência de aplicação da regra decrescente no sentido norte-sul do país), conforme a *posição do acento na palavra* (as vogais tônicas são mais facilmente nasalizáveis que as átonas) e a *natureza da consoante nasal* (a consoante nasal palatal [ɲ] nasaliza a vogal precedente com mais frequência que as não palatais (Cf. Moraes; Wetzels, 1992). Frente a isso, decidimos não criar regras de transcrição para os casos de nasalidade pretônica.

1.1.3.2.3 Glides

Os ditongos, como veremos na seção 1.1.2.2.4 constituem uma sequência de segmentos vocálicos, na qual um é interpretado como vogal e o outro como *semivocóide*, *semicontóide*, *semivogal*, *vogal assilábica* ou *glide* (CRISTÓFARO-SILVA, 2000). Assim como essa autora, adotaremos o termo glide para designar as vogais de menor proeminência acentual em ditongos, uma vez que esse termo tem sido empregado nos estudos mais recentes da área de Fonética, com o objetivo de uniformizar a nomenclatura específica dessa área no nível internacional.

²⁷ A nasalidade é determinada pela presença de consoante nasal em início de sílaba seguinte a da vogal nasalizada, de onde o traço [+nas] se espalharia regressivamente para a vogal anterior.

Como vimos, do ponto de vista fonético, segmentos vocálicos apresentam passagem livre de corrente de ar quando são produzidos, enquanto que segmentos consonantais apresentam algum tipo de fricção e/ou obstrução. Já os glides podem apresentar tanto características articulatórias vocálicas quanto consonantais, ou seja, podem ser articulados com passagem livre de corrente de ar ou não. Essa é uma questão controvertida na análise fonológica dos glides em PB: considerá-los como parte do inventário fonológico das vogais ou das consoantes, conforme será discutido na seção 1.1.3.5.

Na literatura, as vogais [i] e [u] em posição assilábica podem ser indicadas pelos símbolos [j], [y], [ɪ] e [w], [ʊ], respectivamente, conforme a interpretação dos glides como vogais ou consoantes. O diacrítico [̩] é usado sob o glide para representar a vogal assilábica quando se faz uso dos símbolos [ɪ] e [ʊ]: [ɪ̩], [ʊ̩]. Em uma transcrição fonética larga, pode-se simplesmente omitir esse diacrítico e utilizar apenas um símbolo fonético para indicar uma vogal assilábica em um ditongo. Essa estratégia é utilizada em transcrições do inglês contidas no Dicionário Cambridge²⁸ (ex.: <say> ['seɪ]) e foi a mesma estratégia utilizada por Cagliari (2011, 2012) na confecção manual de transcrições fonéticas para dicionário de aprendizes.²⁹ Assim também o será para a representação dos glides na transcrição fonética gerada pelo sistema PETRUS. Utilizaremos apenas os símbolos [ɪ] e [ʊ] nas transcrições de ditongos decrescentes, sem o uso do diacrítico [̩] na posição inferior de tais glides, para indicar sua posição não nuclear na sílaba da qual fazem parte, porque tal informação pode ser obtida a partir do *output* do módulo de divisão silábica do sistema PETRUS. Além disso, o emprego de diacríticos e de marcas suprasegmentais é evitado ao máximo em transcrições fonéticas do tipo larga, como a que adotamos. Portanto, utilizamos apenas o diacrítico de acento principal nas palavras transcritas pelo PETRUS, a fim de mantermos uma facilidade/praticidade de entendimento das transcrições por parte de não especialistas em fonética ou linguística.

1.1.3.2.4 Encontros Vocálicos

Os encontros vocálicos constituem sequências de dois segmentos vocálicos adjacentes que podem tanto pertencer à mesma sílaba (ditongo) quanto fazerem parte de sílabas distintas (hiato).

²⁸ *Cambridge Dictionaries Online*: <http://dictionary.cambridge.org/help/phonetics.html>.

²⁹ Nas transcrições fonéticas realizadas para o Dicionário *Kenermann*, Cagliari (2011, 2012) não assinalou os ditongos com o diacrítico tradicional, como em *vou* ['vʊ̩], *sei* ['seɪ̩], nem se fez uso dos símbolos [j] e [w] (Cf. *vou* ['vow], *sei* ['sej]) para representação dos glides.

Conforme afirma Peixoto (2011, p. 49), sob o ponto de vista articulatório, os encontros vocálicos podem ser produzidos pela decomposição progressiva de um movimento articulatório complexo (ditongo) ou pela modificação da forma da cavidade orofaríngea em decorrência da produção de movimentos articulatórios distintos (hiato). Ainda segundo a autora, a distinção fonético-articulatória entre hiatos e ditongos encontra-se apoiada no fato de os primeiros serem produzidos por diferentes emissões de voz (sílabas distintas), isto é, integram núcleos silábicos diferentes e os segundos serem produzidos por uma mesma emissão de voz (mesma sílaba), ou seja, não estão separados por fronteira silábica.

No PB, a consoante <l>, quando localizada na coda silábica pode sofrer o processo de vocalização e realizar-se foneticamente como [ʊ], formando um ditongo com a vogal que a precede, como acontece, por exemplo, com a palavra *natal* [na'taʊ]. Segundo Leite (1974), a consoante <l> também pode vocalizar para [i] no processo de formação de plural em PB, como ocorre com as palavras <hotel>/<hotéis> e <anzol>/<anzóis>, por exemplo.

Ainda com relação aos ditongos do PB, existem duas possibilidades de ocorrência em uma sílaba: <vogal> + <glide> ou <glide> + <vogal>. As sequências terminadas com glide, necessariamente nessa ordem, são denominadas *ditongos decrescentes*, pois terminam com uma vogal com menor proeminência acentual.

Na Tabela 5 são apresentados os ditongos decrescentes orais possíveis no sistema vocálico do PB, representados, na primeira coluna, pelo sistema de notação fonético IPA, enquanto que na segunda coluna são apresentadas palavras-exemplo contendo ditongos:

Tabela 5. Ditongos orais decrescentes do PB.

<i>Transcrição fonética</i>	<i>Exemplos</i>
[ai]	pa <u>i</u> , ba <u>ix</u> o
[ei]	le <u>i</u> te, de <u>i</u>
[ɛi]	pa <u>péi</u> s, pa <u>stéi</u> s
[oi]	fo <u>ic</u> e, bo <u>i</u>
[ɔi]	int <u>ro</u> ito, dó <u>i</u>
[ui]	circu <u>i</u> to, fu <u>i</u>
[aʊ]	ca <u>us</u> a, pa <u>u</u>
[eʊ]	de <u>us</u> a, me <u>u</u>
[ɛʊ]	me <u>l</u> ro, me <u>l</u>
[iʊ]	fil <u>tr</u> o, ri <u>u</u>
[oʊ]	go <u>l</u> , so <u>l</u> to
[ɔʊ]	vo <u>l</u> to, pro <u>l</u>
[uʊ]	cu <u>l</u> to, az <u>u</u> l

Em muitos dialetos do Brasil, devido à vocalização do fonema /l/ em final de sílaba, também são considerados ditongos decrescentes os casos de <vogal> + <l> em mesma sílaba.

Apesar de Cagliari (2009, 2011; 2012) indicar o uso da vogal [ɑ] para a sequência <al> ou <au>, como em “m[ɑ]ometano” e “m[ɑ]l”, adotamos [a] devido ao fato de almejarmos uma representação fonética larga para o sistema PETRUS e, portanto, avaliarmos [ɑ] como sendo um símbolo fonético que sugere uma conotação mais estreita para representar foneticamente os ditongos formados pelas sequências gráficas <al> e <au>.

Na Tabela 6 são apresentados os ditongos decrescentes orais possíveis no sistema vocálico do PB, em decorrência da vocalização de /l/ em final de sílaba, representados, na primeira coluna, pelo sistema de notação fonético IPA, enquanto que na segunda coluna são apresentadas palavras-exemplo:

Tabela 6. Ditongos orais decrescentes derivados do processo de vocalização do fonema /l/.

<i>Transcrição fonética</i>	<i>Exemplos</i>
[aʊ]	sal
[eʊ]	f <u>el</u> tro
[ɛʊ]	me <u>l</u>
[iʊ]	funi <u>l</u>
[oʊ]	fol <u>g</u> ado
[ɔʊ]	so <u>l</u>
[uʊ]	azu <u>l</u>

O ditongo é nasal quando suas vogais são nasais, ou seja, os sons passam também pelo nariz ou sofrem nasalização quando são produzidos oralmente. “São sempre decrescentes e constituem portanto uma sequência de [vogal nasal - glide]” (CRISTÓFARO-SILVA, 2000, p. 99).

Na Tabela 7 encontramos os ditongos decrescentes nasais do sistema vocálico brasileiro, representados, na primeira coluna, pelo sistema de notação fonético IPA, enquanto que na segunda coluna são apresentadas palavras-exemplo:

Tabela 7. Ditongos nasais decrescentes do PB.

<i>Transcrição fonética</i>	<i>Exemplos</i>
[ẽĩ]	mãe
[ẽõ]	sã <u>o</u>
[õĩ]	aç <u>õ</u> es
[ũĩ]	ruim

A adoção do diacrítico [~] para representar os ditongos nasais se deu com base no estudo de Cagliari (2011, 2012), segundo o qual, na realização fonética dos ditongos nasais, o

glide sofre assimilação da nasalidade da vogal. Portanto, palavras como *mãe*, *bem*, *sanções* e *muito* terão seus ditongos nasais transcritos como [ãĩ], [ẽĩ], [õĩ] e [ũĩ], respectivamente. Adotar esse diacrítico é uma maneira de informar ao usuário do sistema PETRUS que a nasalização recai sobre a glide em ditongos nasais.

Por fim, em relação aos ditongos nasais, Cagliari (2012) afirma que a formação dos ditongos nasais do tipo [ẽĩ] e [õõ] é mais comum do que a ocorrência de vogais nasais simples. Portanto, transcrevemos como [ẽĩ] e [õõ] todos os casos em que ocorreriam as vogais nasais simples [ẽ] e [õ]. Assim, do ponto de vista de uma fonética larga e tendo em vista o seu processamento automático, todas as sequências grafemáticas <em, en> são aqui transcritas como [ẽĩ], enquanto que as sequências gráficas <om, on> são transcritas como [õõ].

Os *ditongos crescentes* são sequências constituídas de <glide> + <vogal>, finalizadas pelo segmento de maior proeminência: a vogal. Na escrita, são ditongos crescentes da língua portuguesa: ea (área), eo (vídeo), ia (história), ie (série), io (notório), oa (mágoa), ua (precípua), uo (árduo).³⁰ Entretanto, na fala há a possibilidade de esses dois segmentos serem pronunciados como se constituíssem sílabas separadas. Mais detalhes sobre esse tema são encontrados na seção 1.1.3.4.

Ainda em relação às possibilidades de encontros vocálicos no PB, podemos encontrar três segmentos dessa natureza (uma glide, uma vogal e outra glide, necessariamente nessa ordem) ocupando uma mesma sílaba, na qual apenas a vogal ocupa a posição de pico silábico. Tais encontros são denominados *tritongos*.

Na Tabela 8 são mostrados os tritongos orais do sistema vocálico brasileiro, enquanto que na Tabela 9, os nasais.

Tabela 8. Tritongos orais do PB.

<i>Transcrição fonética</i>	<i>Exemplos</i>
[ɔai]	Urugu <u>ai</u>
[ɔei]	enxagu <u>ei</u>
[ɔiu]	delinqu <u>iu</u>
[ɔou]	apazigu <u>ou</u>

Tabela 9. Tritongos nasais do PB.

<i>Transcrição fonética</i>	<i>Exemplos</i>
[ɔẽõ]	sagu <u>ão</u>
[ɔẽĩ]	enxagu <u>em</u>
[ɔõĩ]	sagu <u>ões</u>

Embora haja no dialeto aqui adotado a possibilidade de redução de ditongos em determinados contextos (Cf. “pei.xe”/“pe.xe”, “cai.xa”/“ca.xa”), decidimos não implementar regras linguísticas para transcrição desse tipo de redução. Vislumbramos, portanto, como

³⁰ Para Câmara Jr. (1970, p. 54), há uma tendência em transformar ditongos em hiatos, como observados na pronúncia corrente de <suador>, que podemos ter duas realizações: <sua.dor> ou <su.a.dor> - ou em <vaidade> - <vai.da.de> ou <va.i.da.de>. Nesses casos os ditongos podem ser realizados como hiato.

trabalho futuro, um estudo específico sobre esse assunto na variedade adotada para podermos implementar no sistema PETRUS regras relacionadas com essa temática.³¹

Por fim, temos no português, a possibilidade de haver uma sequência de duas vogais pertencentes a sílabas distintas. Tal ocorrência é denominada na literatura de *hiato*. De acordo com Cavaliere (2005), há hiato nos seguintes casos:

1. Entre vogais átonas iguais: <ca.a.tin.ga>, <co.or.de.na.ção>;
2. Entre vogais iguais, na qual a segunda é tônica: <al.co.ó.li.co>, <xi.i.ta>;
3. Entre vogais iguais, sendo a primeira tônica: <vo.o>, <ve.em>;
4. Entre vogais átonas diferentes: <do.a.ção>, <es.te.re.o.ti.pa.do>;
5. Entre vogais diferentes, sendo a primeira tônica: <fa.ri.a>, <pa.vi.o>;
6. Entre vogais diferentes, sendo a segunda tônica: <hi.a.to>, <fre.a.da>.

1.1.3.2.5 Casos de variação fonética não contemplados

Segundo Castilho (2012, p. 197), variação e mudança são propriedades das línguas decorrentes das situações sociais em que nos envolvemos em nosso dia a dia, bem como da adaptação da língua às novas situações históricas. O autor aponta os seguintes eixos que organizam a heterogeneidade linguística existente no PB:

(1) *variação geográfica*. Conforme afirma o autor, existe uma correlação entre a região de origem dos falantes e as marcas específicas de sua produção linguística, por isso que brasileiros e portugueses não falam do mesmo jeito, nem tão pouco brasileiros oriundos de diferentes regiões do Brasil. Para Castilho, dentre todas as variedades do português, a variedade geográfica é a mais perceptível.

(2) *variação sociocultural*. Ainda que se considere falantes provenientes de uma mesma região geográfica, Castilho (2012, p. 204-211) afirma que, ainda assim, há marcas específicas em sua produção linguística, pois cada pessoa pertence a um seguimento social diferente. O autor afirma que é padrão sistematizar as variedades socioculturais considerando-se as variáveis *falante não escolarizado* e *falante escolarizado*. Na primeira variável são considerados os falares de analfabetos e pessoas não escolarizadas, que utilizam um português popular, não culto, não padrão para poderem se comunicar. Em contrapartida, pessoas

³¹ Para mais informações a respeito do fenômeno de supressão da semivogal nos ditongos [ou] e [ei] vide Callou; Leite (2003, p. 94).

escolarizadas utilizam um português culto, padrão, que adquiriu maior prestígio social por ser a variante vigente na classe social com maior prestígio no país. Segundo Correia (2001), trata-se da variedade linguística ensinada na escola, utilizada na maioria dos livros, revistas, textos científicos e em alguns programas de televisão.

(3) *variação individual*. Ao referirmos à fala específica de um indivíduo adotamos o termo *idioleto*. As propriedades particulares da fala de um indivíduo caracterizam seu idioleto (CRISTÓFARO-SILVA, 2000). Conforme Castilho (2012, p. 211-212), os parâmetros registro,³² idade³³ e sexo³⁴ podem ser usados, assim como o nível socioeconômico, para observar a variação individual dentro do PB.

(4) *variação de canal*. Segundo Castilho (2012, p. 212), a língua falada e a língua escrita são variações de canal utilizados conforme a presença ou ausência de um interlocutor, que de certa forma controla os enunciados produzidos.

(5) *variação temática*. Outra característica da fala que Castilho (2012, p. 213) cita como fator que leva à variação linguística é o modo como tratamos o enunciado produzido. Utilizamos, portanto, o *português corrente* para nos referirmos a assuntos de nosso cotidiano em contraposição ao *português técnico*, utilizado quando nos referimos a assuntos especializados.

De acordo com Castilho (2012, p. 198), cada uma dessas variações é organizada, por sua vez, por um conjunto de variantes (i.e., variantes padrão, não padrão, de sexo, etária, formal, informal). Afirma ainda, que o conjunto de variantes (ou dialetos, ou seja, as propriedades linguísticas compartilhadas por um grupo específico de falantes) configura uma variação, e o conjunto de variações configura a variedade linguística.

Portanto, se não fosse o ideal conservador da ortografia, as palavras poderiam ser escritas de tantas maneiras, que acabaria dificultando a leitura nas diferentes variedades linguísticas existentes.

Há variantes fonéticas, que não foram contempladas em nosso trabalho, pois este não comporta atualmente uma indicação aos usuários das possibilidades de pronúncia que uma mesma palavra pode ter em nosso idioma. Assim, decidimos optar por registrar uma única variante, no caso, a de maior prestígio, mais comum, ou mesmo da escolha do grupo de pesquisadores envolvidos (orientadores e doutoranda). Futuramente, quando da

³² Formal (ou refletido) vs. *informal* (ou coloquial).

³³ Pessoas mais idosas tendem a ser mais conservadoras quanto as mudanças que ocorrem nas línguas do que pessoas mais jovens, como por exemplo, o uso de gírias não é comum.

³⁴ Feminino vs. masculino, expressam-se conforme Cristóforo-Silva (2000), em termos de frequência de uso, pois não há no PB marcas gramaticais, palavras específicas ou padrões de entoação utilizados por falantes de um único sexo.

implementação de um sintetizador de voz ao sistema PETRUS, será possível aos usuários vislumbrarem tanto os diferentes modos de transcrição fonética de uma palavra em PB, quanto ouvir as diferentes pronúncias que pode possuir o PB.

Assim, as variações fonéticas não contempladas, por ora, pelo sistema PETRUS são:³⁵

1. Nasalidade de vogais átonas seguidas por consoante nasal: *menino* [mĩ'nĩnɔ].
2. Redução de ditongos: *pedreiro* [pe'dreĩɾɔ] ~ [pe'drerɔ]; *couraça* [koʊ'rasa] ~ [ko'rasa].
3. Ditongação das sílabas tônicas finais de palavras terminadas em fricativa alveolar surda (/s/): *vocês* [vo'seis], *rapaz* [xa'pais], *voz* ['vois].
4. Variações de [õõ], como: *bomba* ['bõõmba] ~ ['bõmba] ~ ['bomba] ~ ['boʊmba].
5. Variações de [êĩ], como: *pente* ['pẽĩtʃĩ] ~ ['pẽntʃĩ] ~ ['pẽĩtʃĩ] ~ ['peĩtʃĩ].
6. Uso enfático, depois de ditongo ou ditongo reduzido de [ŋ], [ɲ]: *homem* ['õmẽĩ] ~ ['õmẽĩŋ] ~ ['õmẽɲ]; *campo* ['kẽŋpɔ] ~ ['kẽmpɔ] ~ ['kẽpɔ]; *lã* ['lẽ] ~ ['lẽŋ]; *bom* ['bõõ] ~ ['bõŋ]

1.1.4 Fonologia

Enquanto a Fonética estuda a natureza física da produção e da percepção dos sons da fala (fones), a Fonologia se preocupa com a maneira como os sons se organizam dentro de uma língua, classificando-os em unidades capazes de distinguir significados, denominadas fonemas.

Desde a criação do termo fonema, vários foram os conceitos a eles atribuídos. Callou e Leite (2003, p. 35) listam alguns: “de início, igualado ao som da linguagem; depois conhecido sob um prisma essencialmente psíquico, como intenção de significado; mais tarde, sob um prisma físico, funcional e abstrato”. O fonema é uma unidade mínima de som que, dentro de um sistema fônico determinado, é capaz de estabelecer a distinção de significado entre dois vocábulos: a diferença de significado entre *mata* e *nata* existe, pois os fonemas /n/ e /m/ são unidades mínimas de som possíveis de promover distinção de significado entre tais palavras, por exemplo. Assim, o fonema não possui significado por si só, mas pode, sozinho, distinguir significados em palavras. Aliás, são identificados justamente por essa função que exercem na língua.

Interessante também notar a distinção que se deve fazer entre fonema e letra. O primeiro é um elemento abstrato, funcional, enquanto que o segundo é um sinal gráfico, que representa o fonema, segundo a convenção de uma dada língua. Essa distinção se faz importante pois,

³⁵ Exemplos de variação fonética não contemplados apresentados com base nos exemplos de Cagliari (2012).

como veremos mais a frente (na seção 2.3.5), nem sempre pode haver correspondência unívoca entre letra e som no PB.

Um dos objetivos da análise fonológica é definir quais sons têm valor distintivo dentro do sistema de uma língua. Para tanto, a Fonologia se vale do fonema, unidade sem significado, mas com função distintiva, para determinar a diferença de significado de uma palavra em relação à outra.

Para Borba (1970, p. 186), “compete à fonologia identificar os fonemas, determinar os traços pertinentes, as oposições e correlações e seus tipos.” Sendo os fonemas entidades abstratas e funcionais, parte-se das realizações concretas da fala para chegar ao sistema fonológico da língua, daí a razão pela qual se diz que a Fonética é básica para a Fonologia. Isso equivale a dizer que faz parte do domínio da Fonologia as questões relacionadas à seleção e à combinação dos sons para formar palavras. No primeiro caso, temos implícita a questão da oposição paradigmática, ou seja, selecionar dentre os sons possíveis da língua, qual será utilizado em determinada posição da palavra. Assim diferenciamos, por exemplo, *pato* de *mato*, porque /p/ e /m/ são fonemas diferentes. No segundo caso, trata-se de combinações sintagmáticas, que determinam como organizar um conjunto de sons para que se forme determinada palavra. Assim diferenciamos *amor*, *Roma*, *ramo* e *mora*, por exemplo, já que em cada um dos quatro casos, que correspondem às organizações possíveis para um mesmo conjunto de sons, temos palavras, que possuem significado diferente das demais.

O sistema sonoro de uma língua pode ser analisado fonologicamente em nível segmental, ou seja, procede-se à análise dos segmentos consonantais e vocálicos. No entanto, várias outras características fonológicas podem afetar os segmentos de uma língua, tais como sílaba, palavras, frases e sentenças. A análise de traços fonológicos de unidades maiores que o segmento é uma preocupação de vários modelos fonológicos. Padrões de ritmo, tempo, volume, tom de voz, são geralmente estudados pela fonologia suprasegmental, na escola americana, e na prosódia, na escola europeia (CRYSTAL, 1997).

Com o objetivo de tentar estabelecer relações entre o nível fonético e o fonológico, descrevendo e analisando como as línguas organizam os sons da fala, surgem as diferentes teorias fonológicas (BISOL, 1996). Os modelos teóricos surgidos na evolução dos estudos fonológicos podem ser organizados em dois grandes grupos: os modelos lineares e os não-lineares.

Os modelos lineares, que analisavam a fala a partir de uma sequência de segmentos, contribuíram significativamente para a evolução dos pensamentos fonológicos ao desenvolver as noções de traços distintivos, classes naturais e regras fonológicas. No entanto, tais modelos

apresentavam limitações ao tentar explicar fenômenos fonológicos suprasegmentais e prosódicos. Entre os modelos lineares de análise fonológica, podemos citar o estudo fonológico do Círculo Linguístico de Praga, a fonêmica estruturalista norte-americana, e o modelo gerativista de Chomsky e Halle (1968).

Os modelos não-lineares, portanto, buscaram analisar a fala não como uma combinação unidimensionalmente ordenada de segmentos, pois propunha-se que os segmentos se organizavam hierarquicamente. Assim, segmentos formam sílabas, que, por sua vez, formam pés métricos, que formam palavras fonológicas, que formam grupo clítico, que formam frase fonológica, que formam frase entonacional e que por fim formam o enunciado. Para mais detalhes sobre essas categorias, vide Bisol (2005).

O grande mérito do modelo autosegmental foi a incorporação da sílaba à teoria. No modelo gerativo, a sílaba era tratada como um traço [+ silábico] que era atribuído a um segmento. Esse tipo de análise era insuficiente e não conseguia dar conta de diversos fenômenos relativos à sílaba. No modelo autosegmental, portanto, a sílaba adquiriu status fonológico. Os segmentos passam, então, a não ser somente um conjunto desorganizado de traços, mas um conjunto que possui uma estrutura interna organizada hierarquicamente. Dentro desses modelos não-lineares encontram-se as fonologias autosegmental, métrica, lexical, da sílaba e prosódica.

Cada língua tem os seus próprios fonemas, que são elementos fônicos dotados de função representativa no sistema. Na língua portuguesa temos 26 fonemas segmentais (7 vogais e 19 consoantes) e um fonema suprasegmental, o acento, que não é um segmento, mas sim uma qualidade que se sobrepõe a certos segmentos (CALLOU; LEITE, 2003, p. 42).

1.1.4.1 Consoantes

As consoantes não constituem centro de sílaba e aparecem sempre ao lado de uma vogal (CÂMARA JR, 1953). Dessa maneira, dependendo da posição que ocupam na sílaba, podem ser caracterizadas em consoantes pré-vocálicas ou pós-vocálicas. Na posição anterior a vogal, temos o seguinte conjunto de consoantes possíveis, de acordo com seu *Modo de Articulação* e as particularidades dialetais da variedade aqui adotada:

oclusivas	/p/, /b/; /t/, /d/; /k/, /g/
nasais	/m/, /n/, /ɲ/
vibrante	/r/
fricativas	/f/, /v/; /s/, /z/; /ʃ/, /ʒ/; /x/, /χ/
laterais	/l/, /ʎ/

Com exceção de /χ/, é plenamente possível que as demais consoantes acima listadas ocorram em posição de *onset* inicial (início da sílaba) ou medial de palavra. Apenas em algumas poucas palavras de origem indígena é que a consoante nasal palatal [ɲ] aparece em posição inicial de palavra: (Cf. *nhambu* [ɲẽ'bu] ou às palavras *nhô* ['ɲo], *nhá* ['ɲa], *nhoque* ['ɲokɪ]. Já a lateral palatal [ʎ] ocorre em posição de *onset* medial: palha ['paʎa], e sua ocorrência em início de palavras é muito rara: *lhama* ['ʎẽma], *lhe* [ʎe].

Além da divisão pelo *Modo de Articulação*, estabelece-se a seguinte classificação fonológica consonantal em relação ao *Lugar de Articulação*:

bilabial	/p/, /b/; /m/
labiodental	/f/, /v/
alveolar	/t/, /d/; /s/, /z/; /n/; /l/ /r/
pós-alveolar	/ʃ/, /ʒ/
palatal	/ɲ/; /ʎ/
velar	/k/, /g/; /x/, /χ/

Segundo Crystal (2008, p. 33), *arquifonema* é um termo utilizado em Fonologia que se refere à maneira encontrada para lidar com o problema da neutralização (quando o contraste entre fonemas é perdido em certas posições de uma palavra). A solução sugerida por Nicolai Trubetskoy foi estabelecer uma nova categoria para tais casos, a qual ele chamou de arquifonema, e transcrevê-los com um símbolo diferente, como por exemplo, utilizando letras maiúsculas para representá-los fonologicamente. No PB temos os seguintes arquifonemas:

/S/, resultante da neutralização de /s/ e /ʃ/
 ortografia: <pasta>, <asma>
 transcrição fonética: ['pasta] e ['azma]
 transcrição fonológica: /'paSta/ e /'aSma/

/R/ resultante da neutralização dos fonemas /R/ e /r/
 ortografia: <porta>
 transcrição fonética: ['pɔhta], ['pɔxta], ['pɔɪta], ['pɔɾta], ['pɔɾta],], ['pɔɾta]

transcrição fonológica: /'pɔRta/:

/L/, resultante da neutralização de /l/ e /u/

ortografia: <salto>

a) transcrição fonética: ['saltu] ou ['sɔtu]

b) transcrição fonológica: /'saLto/

/N/, resultante da nasalização da vogal

ortografia: <bomba>

a) transcrição fonética: ['bõ^mba]

b) transcrição fonológica: /'boNba/

1.1.4.2 Vogais Orais

De acordo com Câmara Jr. (1970, p. 41-43), as vogais orais do PB são representadas por um sistema triangular, em cujo vértice mais baixo está a vogal /a/. A elevação gradual da língua, seja na parte anterior, seja na parte posterior, dá origem à seguinte classificação articulatória: vogal baixa, vogais médias de primeiro grau, vogais médias de segundo grau e vogais altas, esquematizadas pelo autor da seguinte maneira:

Tabela 10. Vogais tônicas segundo modelo de Câmara Jr. (1970, p. 41-43).

altas	/u/		/i/	
médias	/o/		/e/	(2º grau)
médias	/ɔ/	/ɛ/		(1º grau)
baixa		/a/		
	posteriores	central	anteriores	

Podemos observar na Tabela 10, que no contexto de sílaba tônica, os fonemas vocálicos /e/, /ɛ/ e /o/, /ɔ/ estão em oposição, ou seja, a alternância dessas vogais médias em palavras do PB pode causar diferença de significado, como em: 's/e/de' – 's/ɛ/de' e 'av/o/' – 'av/ɔ/'. Entretanto, em posição átona, esse contraste observado entre as vogais médias desaparece.

Assim, as sete unidades fonológicas tônicas são reduzidas a cinco em posição átona (CÂMARA JR, 1970, p. 41-43).

Conforme veremos na Tabela 11, a existência de cinco vogais é observada tanto em posição pretônica quanto em postônica não-final.

Tabela 11. Vogais átonas não-finais segundo modelo de Câmara Jr. (1970, p. 41-43).

altas	/u/		/i/
médias		/o/	/e/
baixa		/a/	
	posteriores	central	anteriores

Entretanto, o sistema vocálico em posição átona final se reduz a três realizações, conforme mostrado na Tabela 12:

Tabela 12. Vogais em posição final átona segundo modelo de Câmara Jr. (1970, p. 41-43).

altas	/u/		/i/
baixa		/a/	
	posteriores	central	anteriores

Em suma, podemos afirmar que, do ponto de vista fonológico, o PB possui um sistema vocálico composto por sete fonemas orais /a/, /e/, /ɛ/ e /o/, /ɔ/, /u/, que, dependendo da posição que ocupam na sílaba, seja ela tônica ou átona, podem ser reduzidos a diferentes realizações, diminuindo o contraste entre a altura das vogais posteriores e as anteriores.

1.1.4.3 Vogais Nasais

A questão da interpretação das vogais nasais constitui um ponto controverso da fonologia portuguesa, despertando, portanto, interesse de especialistas e promovendo interpretações variadas, consoante diferentes pontos de vista teóricos.

A questão que se coloca na análise das vogais nasais são as hipóteses de haver doze fonemas vocálicos distintos (sete orais e cinco nasais) ou sete fonemas vocálicos apenas e as vogais nasais consistirem na combinação de uma vogal oral com o arquifonema nasal /N/.

Dentre os autores que defendem a primeira hipótese estão Moraes e Wetzels (1992), Leite (1974), Head (1964), Back (1973), Abaurre-Gnerre (1983). Para esses autores, o contraste existente em certos pares mínimos observados na língua, tais como, “lá”/“lã”, “lida”/ “linda”, “boba”/“bomba” evidenciam a oposição oral/nasal existente no português. Em outras palavras, o fato de encontrarmos esses tipos de pares mínimos na língua dá suporte para a afirmação de que as vogais nasais em PB apresentam estatuto fonêmico, isto é, podem ser representadas como vogais nasais.

Por outro lado, na segunda hipótese, considera-se a não existência de vogais nasais, mas sim, de uma nasalização decorrente de uma combinação de uma vogal oral seguida de arquifonema nasal /N/. De acordo com essa proposta, as vogais nasais devem ser representadas fonemicamente como /aN, eN, iN, oN, uN/. Entre seus defensores estão Mateus (1975), Almeida (1976), Pardal (1977), Wetzels (1991), Câmara Jr. (1953, 1970), bem como o foneticista Cagliari (1977; 1981), com seu estudo experimental sobre as vogais nasais na variante paulista.³⁶

Câmara Jr. (1953, 1970) apresenta alguns argumentos em favor de sua interpretação fonológica das vogais nasais. O primeiro deles diz respeito à sílaba com vogal nasal. De acordo com esse autor, as vogais nasais comportam-se de maneira similar às vogais que ocorrem em sílaba travada por consoante, pois quando uma palavra terminada em vogal nasal é sucedida por uma palavra iniciada por vogal oral, não há crase. Exemplos de palavras que ilustram esse primeiro argumento de Câmara Jr. são “lã amarela” e “jovem amor”. Outro argumento do estudioso em favor de sua hipótese bifonêmica baseia-se na distribuição do <r> na estrutura silábica do português. Devido à impossibilidade de as vogais nasais virem

³⁶ Segundo Rothe-Neves e Reis (2012, p. 301) seguiram-se, a partir do trabalho de Cagliari (1977), as primeiras análises acústicas da nasalidade vocálica (SOUZA, 1994; JESUS, 1999; SEARA, 2000; MEDEIROS, 2009), bem como a utilização de diferentes técnicas para análise do fenômeno, dentre elas, as autoras citam a ressonância magnética (GREGIO, 2006), síntese de fala (TEIXEIRA, 2000), nasometria (GALVÃO, 1998) e fibroscopia ótica (LOVATTO et al., 2007).

seguidas de /r/ simples (*flap*), contrariamente ao que acontece com as vogais orais que, depois delas, pode haver os dois tipos de vibrantes (ex.: *carro, caro*), configurando um travamento consonântico após a vogal nasal (por exemplo: *tenro*). Câmara Jr. (1970) argumenta também que temos hiatos em português (*pi.a.da*), mas não temos hiatos com vogal nasal. Nos casos em que poderia ser possível a ocorrência de hiato com a vogal nasal, ou a nasalidade desaparece (por exemplo, *bo.a*) ou o segmento correspondente ao segmento nasal passa a ocupar uma posição consonantal na sílaba seguinte (*va.len.to.na*). Por fim, o autor argumenta que não podemos considerar que vogais nasais tenham *status* de fonema em línguas como o português que não apresentam pares mínimos do tipo francês, como /bõ/ e /bon/, ou seja, vogal nasal que comute com vogal oral seguida de consoante nasal. (CRISTOFARO-SILVA, 2000, p. 166)

Do ponto de vista teórico, ambas as análises são possíveis: a que interpreta as vogais nasais como sendo a combinação de uma vogal oral seguida de um arquifonema nasal /VN/ que ocorre em posição pós-vocálica, possibilitando afirmar a existência de um conjunto de sete fonemas vocálicos do português associado ao arquifonema nasal /N/; bem como a que assume a existência do contraste fonológico vogal oral/vogal nasal e, conseqüentemente, a presença de doze fonemas vocálicos (sete orais e cinco nasais). Isso é possível devido ao fato de os dados da língua oferecerem argumentos razoáveis para as duas análises.

1.1.4.4 Encontros Vocálicos

Os *ditongos* e *hiatos* estão relacionados com sequência de vogais. Se há uma sequência de duas vogais em uma mesma sílaba, temos um ditongo (ex.: *pais*). No entanto, se a sequência de duas vogais se encontra em sílaba distinta, temos um hiato (ex.: *pa.ís*). Deste modo, em um ditongo, temos um núcleo que representa a parte central da sílaba e a parte periférica ocorre juntamente com a vogal.

Trabalhos brasileiros que versam sobre a ditongação concordam sobre a existência de dois tipos possíveis de ditongos: os fonológicos e os fonéticos (BISOL, 1989). Os primeiros são considerados os verdadeiros ditongos da língua, pois criam oposições no léxico (ex.: *mau/má*) e não estão sujeitos à variação livre com o hiato (ex.: *pau.la.da/pa.u.la.da**). No entanto, os ditongos fonéticos podem variar livremente com o hiato (ex.: *nó.do.a/nó.doa*) e estão sujeitos a reduções (ex.: *pei.xe/pe.xe*).

Devido ao fato de os ditongos crescentes poderem variar livremente com o hiato (Cf.

<su.a.dor>/<sua.dor>), decidimos transcrever apenas os ditongos decrescentes contidos na Tabela 5 apresentada na seção 1.1.2.2.4. Vale ressaltar, entretanto, que nos casos em que as vogais <i> e <u> são precedidas por vogais e seguidas por consoantes nasais <m, n, nh> ou por <ir>, tais encontros vocálicos deverão ser transcritos como hiatos. Exemplos: <ra[i]nha>, <conclu[i]r>, <sa[i]rmos>.

1.1.4.5 Glides

Há opiniões conflitantes a respeito dos glides no inventário fonêmico do português: considerá-los como parte do inventário fonológico das vogais ou das consoantes.

A esse respeito, Câmara Jr. (1953), argumenta que os glides em português devem ser interpretados como fonemas consonantais constitutivos da parte pós-vocálica de sílabas travadas (CVC). Nessa abordagem, incluir os glides aos demais fonemas consonantais significa aumentar o inventário fonêmico das consoantes de dezenove para vinte e uma; por outro lado, significa diminuir os tipos silábicos a descrever, pois nessa primeira interpretação, mantemos uma estrutura silábica já prevista em português: a sílaba travada por um segmento consonantal pós-vocálico ((C)VC). Portanto, embora tenhamos um sistema fonotático mais simples (que exclui sílabas CVV), temos um sistema fonêmico mais complexo (que inclui os fonemas /y, w/).³⁷

Para Câmara Jr. (1953), os glides de ditongos crescentes (f/y/el, m/y/olo, s/w/ar) não são considerados fonemas, pois alternam com vogais altas, sem qualquer valor distintivo, exceto /kw/ e /gw/. Entretanto, os glides de ditongos decrescentes são considerados fonemas que constituem centro de sílaba polifonemático com vogais silábicas, ou seja, o glide é um fonema localizado no núcleo silábico, junto à vogal (CÂMARA JR, 1953, p. 57). Para sustentar sua hipótese, o autor considerou como pares mínimos as palavras <vou> /vow/ e <voo> /vou/, verbos com clíticos <viu> /viw/ e <viu-o> /viu/ e verbos do presente do indicativo e da primeira pessoa do presente do subjuntivo, nos quais o glide /y/ sucede a vogal /u/.

Uma proposta alternativa a essa é interpretar os glides como parte do sistema vocálico. Em 1969 e 1970, Câmara Jr. revê a proposta que defendeu em 1953 e demonstra que os glides

³⁷ Cristóforo-Silva (2000) afirma que, do ponto de vista fonológico, a utilização de /y/ e /w/ sugere que os glides comportam-se de maneira análoga aos segmentos consonantais na estrutura silábica. Ela explica que a escolha por tratar as semivogais como segmentos vocálicos ou consonantais tem a ver com a proposta de se manter o padrão silábico do português CVC, ou incorporar um novo padrão silábico do tipo CVV, respectivamente.

em português devem ser analisados como segmentos vocálicos (não havendo necessidade de incluir os fonemas /y, w/), mantendo-se, assim, um sistema fonológico com dezenove fonemas consonantais. Nessa nova abordagem, Câmara Jr. (1970) propõe que o glide se localiza na coda silábica (CVV), paralela às demais consoantes, constituindo ditongo decrescente. Com essa segunda hipótese, o autor reduz os tipos silábicos do PB para cinco: i) sílaba simples constituída por V (é); ii) sílaba complexa, aberta ou livre, constituída por CV (só); iii) sílaba fechada ou travada, constituída por VC (arco); (iv) sílaba completa, constituída por CVC (for) e v) sílaba leve, terminada em vogal CVV (pauta).

Em suma, na primeira hipótese, se considerar os glides do PB como consoantes, tem-se um sistema fonotático mais simples (exclui sílabas CVV), entretanto, um sistema fonêmico mais complexo (com a inclusão dos fonemas /y/ e /w/). Na segunda hipótese, assume-se o padrão silábico CVV para interpretar os glides como vogais e excluímos os fonemas /y/ e /w/ do inventário fonêmico, embora se obtenha um sistema fonotático mais complexo (inclusão de sílabas CVV). Para o desenvolvimento do sistema PETRUS, adotamos a segunda hipótese. Futuramente,³⁸ pensamos em comparar as duas hipóteses para verificar se há alteração no desempenho tanto das transcrições de ditongos do PB quanto da divisão silábica realizadas pelo sistema.

1.1.4.6 Variação Alofônica

Como vimos anteriormente (seção 1.1.3.1), os fonemas são capazes de distinguir significado entre unidades lexicais de uma língua, já os alofones são as realizações fonéticas de um mesmo fonema, que não alteram significado de palavras. Tais variações podem ser de diferentes tipos: *posicionais* (inerentes à vizinhança fonética ou coarticulação), *regionais* (em função da variedade geográfica, como o /S/ ([ʃ]) carioca e o /s/ ([s]) paulista), *estilísticas* (decorrentes de intenção expressiva, por resultarem de opções individuais, como o alongamento de [e] em <eeeeuuu?!> e a vibração alongada de [r] em <rrrrrrronaldo!!>) ou *livres* (decorrente de características individuais do falante, independentemente de qualquer fator condicionante).

De acordo com Callou e Leite (2003, p. 42) as variantes *posicionais* são as que mais interessam aos foneticistas, pois decorrem do próprio contexto fônico em que o alofone é realizado. São exemplos de alofones condicionados pelo contexto, as realizações dos fonemas

³⁸ Como um possível tema de pesquisa, por exemplo, para um pós-doutorado.

/t/ e /d/, os quais apresentam no dialeto paulista e carioca, por exemplo, uma realização palatal diante de /i/ e [ɪ] (ex.: tomate [to'matʃɪ], time ['tʃimɪ], ditado [dʒi'tadʊ], idade [i'dadʒɪ]) e uma realização alveolar ou dental diante de outras vogais (ex.: tudo ['tudo], tecla ['tɛkla], tara ['tara], tola ['tola], doação [doa'sẽõ], dúvida ['duvida], data ['data].

Para o desenvolvimento das regras de transcrição implementadas no PETRUS, nos foi de grande interesse as variantes posicionais - pois sua realização é determinada pelo contexto fônico - e regionais, visto que consideramos o falar da Rede Glodo para ser reproduzido nas transcrições geradas por nosso sistema.

1.1.4.7 Processos Fonológicos

Conforme afirma Cagliari (2002, p. 99), os processos fonológicos constituem “alterações sonoras que ocorrem nas formas básicas dos morfemas, ao se realizarem foneticamente”, quando se combinam para formar palavras ou na junção delas.

É um fenômeno natural das línguas, ocasião em que alguns sons exercem influência sobre outros, provocando, assim, alterações na articulação dos sons que foram influenciados. Tais alterações podem ser estudadas em um determinado momento (sincronicamente), ou seja, podem ser estudadas as modificações pelas quais passam as palavras por influência de contextos vizinhos – como é o caso desta pesquisa – ou podem ser investigadas comparando-se palavras em diferentes momentos da história de nosso idioma (diacronicamente). Essas mudanças podem alterar ou acrescentar traços articulatorios, bem como eliminar ou inserir segmentos. Os processos fonológicos são classificados em função dessas alterações.

Para a nossa pesquisa, consideramos apenas os processos fonológicos aparentes em estruturas internas às palavras. Assim, incorporamos à construção das regras de transcrição fonética do sistema PETRUS, os processos que:

A) Acrescentam traços ou alteram sua especificação

Assimilação: quando um som torna-se mais semelhante a outro, que lhe está próximo, adquirindo uma propriedade fonética que ele não tinha (CAGLIARI, 2002, p. 99), tais como a vocalização da lateral em final de sílaba ou palavra: sal ['saw]; maldade [maʊ'dadʒɪ]; o vozeamento da fricativa /s/ em coda silábica seguida por consoante vozeada: mesmo ~ ['mezmo]); a nasalização diante de consoante nasal, principalmente, em sílabas tônicas: cama

['kẽma] e a palatização das consoantes oclusivas alveolares diante de vogal anterior palatal [i]: dia ['dʒia]).

Os processos fonológicos de *harmonização vocálica* (assimilação do traço de altura das vogais tônicas /i/ e /u/, em que as vogais médias se tornam altas pela influência de uma vogal alta presente na sílaba seguinte à da pretônica-alvo, como em m[i]nino, f[u]rmiga) e de *metafonia* (alternância relacionada ao timbre da vogal que quando no singular, se apresenta como um som fechado, e no plural, com som aberto, como em car[o]ço e car[ɔ]ços) também se incluem neste item A.

Especificamente em relação a harmonização vocálica, devido ao fato de elaborarmos regras de transcrição fonética segundo o dialeto da Rede Globo, consideraremos a inclusão do processo de harmonização vocálica em nossa regras de transcrição apenas futuramente. Assim, pretendemos como um de nossos trabalhos futuros, a elaboração de um *cópus* da variante linguística em estudo e posterior análise desse fenômeno, consoante as classes gramaticais *nome* e *verbo*. Tal escolha categorial foi inspirada nos estudos realizados sobre a variante culta praticada no noroeste paulista por Carmo (2009) – que investigou as vogais pretônicas de verbos oriundos de dezesseis entrevistas do Banco de Dados Iboruna, resultado do Projeto ALIP (Amostra Linguística do Interior Paulista), e por Silveira (2008) que investigou as vogais pretônicas de nomes, que também tiveram como fonte o mesmo banco de dados utilizado por Carmo.³⁹ Esses estudos indicam que os mesmos processos fonológicos ocorrem de maneira distintas dadas as configurações morfológicas diferentes das palavras contidas nas classes gramaticais dos verbos e dos nomes.

B) Inserem segmentos.

Epêntese: caracteriza-se pela inserção de um segmento, em geral um [i] átono e breve, em determinadas sílabas do PB (Cf. CAGLIARI, 1981; LEE, 1993; COLLISCHONN, 1996; 2002; MASSINI-CAGLIARI, 2000).⁴⁰ A vogal em questão é inserida, conforme Cagliari (1981, p. 107), “entre uma oclusiva, uma nasal bilabial ou uma fricativa alveolar surda por um

³⁹ O Banco de Dados Iboruna (Iboruna, sendo um topônimo de origem tupi-guarani que se pretendeu atribuir a cidade de São José do Rio Preto por ocasião da comemoração do seu cinquentenário) tem por objetivo constituir-se de base para a descrição do português falado no interior paulista, mais especificamente na região noroeste do Estado de São Paulo. Os informantes que cederam suas amostras de fala, que hoje compõem o banco de dados denominado Iboruna, são provenientes das seguintes cidades: Bady Bassit (BAD), Cedral (CED), Guapiaçu (GUA), Ipiguá (IPI), Mirassol (MIR), Onda Verde (OND) e São José do Rio Preto (SJP). Por meio de uma *homepage*, é possível acessar livremente uma subamostra composta de 10 gravações da Amostra Comunidade e de duas gravações da Amostra de Interação, acompanhadas das respectivas transcrições e dos diários de campo. Homepage do Iboruna: <http://www.iboruna.ibilce.unesp.br/interna.php?Link=corpo.php&corpo=8>.

⁴⁰ Para mais informações sobre epêntese consonantal em PB, ver Cagliari e Massini-Cagliari (2000).

lado, e uma outra consoante por outro lado”. A inserção dessa vogal transforma estruturas silábicas do tipo CC em estruturas do tipo CVC.

C) Apagam segmentos.

Degeminação: consiste na fusão de duas vogais idênticas e, conseqüentemente, no encurtamento da vogal longa resultante (ex.: *álcool* ['aʊkou]). Há a previsão de bloqueio desse processo caso a segunda vogal esteja em posição tônica na estrutura silábica: *microônibus* [mikro'onibos].

1.1.5 Prosódia

Mateus et al. (2003, p. 1037) definem a prosódia como um ramo da Linguística que investiga as propriedades fônicas da cadeia da fala, as quais “contribuem para a interpretação do significado e determinam o ritmo da frase”, a saber: o tom, o acento (tônico) e a duração.

Mateus e Villalva (2006, p. 59-60) afirmam que os sons das línguas não possuem apenas as propriedades articulatórias que diferenciam um /a/ de um /i/ ou de um /o/, mas também apresentam propriedades prosódicas, tais como: a intensidade (a vogal pronunciada com maior intensidade é a que contém o acento da palavra), a duração (em certas línguas as vogais podem contrastar pelo tempo de pronúnciação, sendo umas breves e outras longas, e a altura ou tom (a sequência de tons das vogais de uma palavra ou frase constitui a entoação).

Em relação ao tom e à duração, as mesmas autoras afirmam que eles não permitem distinguir significados no português, contrariamente ao que acontece em outras línguas como o latim, no qual a duração da vogal numa mesma sequência pode indicar a função sintática da palavra: *rosã*, com vogal final breve, é nominativo (tem função de sujeito) e com vogal final longa, *rosā*, é ablativo (tem uma função complementar).

Quanto à intensidade, Mateus e Villalva (2006, p. 60) declaram que tal propriedade prosódica está relacionada com o acento tônico da palavra e marca uma sílaba que é pronunciada com mais força, que se torna proeminente na sequência de sílabas que forma a palavra. Ainda de acordo com as autoras, podemos observar no português a presença do acento em todas as palavras, com exceção dos clíticos, sendo ele o responsável por distinguir palavras que apresentam mesmas sequências de vogais, mas com acento tônico em sílabas diferentes, tal como podemos observar nos pares de palavras “secretária”/“secretaria” e

“dúvida”/“duvida”. Portanto, podemos observar que os fonemas constitutivos dessas palavras são os mesmos; entretanto, o único traço que as distingue, a rigor, é a posição do acento de intensidade.

Em português, a menor unidade prosódica é a sílaba, que como veremos na próxima seção, apresenta características particulares (por exemplo, apenas certas sequências de duas consoantes podem pertencer à mesma sílaba: /br/ integra a segunda sílaba de *pobre* mas a sequência /st/ pertence a duas sílabas na palavra *pasta*).

1.1.5.1 Sílaba

A sílaba pode ser definida como sendo “a estrutura linguística formada por uma vogal ou fonema com valor de vogal, eventualmente acompanhada de não vogais, isto é, semivogais ou consoantes” (BIZZOCCHI, 2012). Uma língua pode ter diversas formas de silabação, ou seja, divisão de suas palavras de acordo com o número de sílabas nela contida (CRYSTAL 2008, p. 467). São elas:

- i) monossílabas: palavras com uma sílaba (ex: “pá”);
- ii) dissílabas: palavras com duas sílabas (ex.: “ca.sa”);
- iii) trissílabas: palavras com três sílabas (ex.: “ca.ba.na”) e
- iv) polissílabas: palavras com mais de três sílabas (ex.: “com.ple.xi.da.de”).

Do ponto de vista acústico e articulatório, a descrição da sílaba constitui um dos maiores desafios da fonética. A ausência de critérios fonéticos claros, que permitam delimitar de modo preciso as sílabas no *continuum* sonoro, é evidenciada por muitos autores.

Cagliari (2009, p. 109), com base em Stetson (1951), define a sílaba como:

(...) resultados de movimentos musculares, quando os músculos da respiração modificam o processo respiratório adaptando-o ao processo de fala. Como consequência, o ar dos pulmões não sai em fluxo contínuo e pressão constante, mas em pequenos jatos que formam o suporte sobre o qual se montam os outros parâmetros da fala. A sílaba seria, portanto, o primeiro parâmetro articulatório a ser ativado e nenhum enunciado poderia, em princípio, ser pronunciado sem que fosse no início montado sobre sílabas (CAGLIARI, 2009, p. 109).

Para Cagliari (2009, p. 110), o esforço muscular que gera as sílabas pode ser caracterizado como um movimento de força que se intensifica e se reduz em cada sílaba, podendo-se, conseqüentemente, reconhecer três partes nesse movimento: de intensificação da

força, de limite máximo de força atingido e de redução progressiva da força. Em decorrência desse tipo de movimento, o autor interpreta a sílaba como sendo uma estrutura constituída por três partes: duas partes periféricas (intensificação e redução da força) e uma parte central ou nuclear (limite máximo da força).

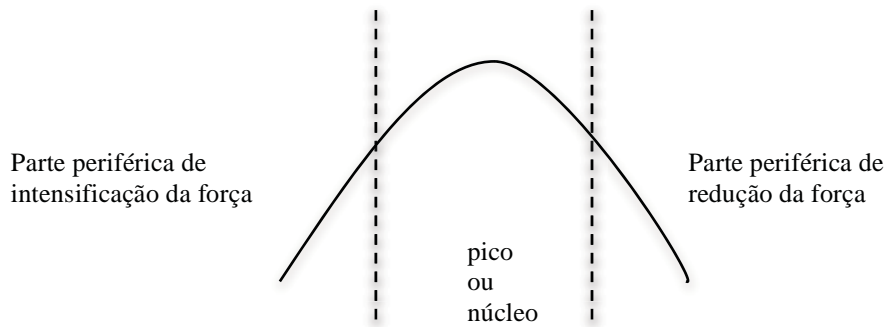


Figura 7. Esquema do esforço muscular e curva de força silábica.
Fonte: Cagliari (2009, p. 111)

A sílaba começa a ser considerada do ponto de vista fonológico pelos prosodicistas de Londres, a partir da Escola de Praga, passando pelo estruturalismo americano, teoria gerativa, fonologia segmental e pela métrica, perfazendo um trajeto de crescente importância no decorrer do tempo.

Na perspectiva da fonologia, há duas propostas para descrever a representação fonológica da sílaba: uma denominada por Câmara Jr. (1970) de *organização linear* e a *teoria hierarquizada* de Selkirk (1982).

Conforme o modelo linear de Câmara Jr. (1970), a estrutura silábica apresenta um ponto central do qual derivam três posições, as quais correspondem, respectivamente, aos momentos de intensificação, força máxima e redução da perspectiva fonética.

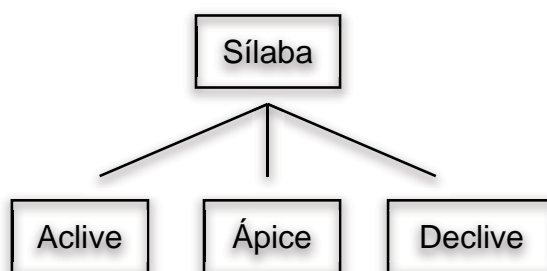


Figura 8. Modelo Linear de sílaba (Câmara Jr., 1970)

No PB, segundo Câmara Jr. (1970), o ápice é obrigatoriamente preenchido por qualquer vogal enquanto que o aclave pode ser preenchido por (quase) qualquer consoante.⁴¹ A posição de declive pode ser ocupada pelos arquifonemas /N/ e /S/, pelos fonemas /r/ e /l/, bem como pelas semivogais /y/ e /w/, que embora sejam vogais por natureza, funcionam como consoante, porque travam a sílaba, representando uma modificação final ou inicial do seu centro, com emissão reduzida, o que as caracteriza como glide, ou vogal pela metade.

O modelo hierárquico proposto por Selkirk (1982) orienta-se pela vertente gerativista, buscando explicar a organização da sílaba de um ponto de vista que supõe haver características universais (comuns) entre o modo como as diferentes línguas do mundo definem seus modelos silábicos. Segundo essa perspectiva, a sílaba seria um constituinte da estrutura fonológica da língua da qual partem dois constituintes imediatos: o ataque e a rima, que por sua vez, ramifica-se em núcleo e coda. Assim, a sílaba adquiriu status fonológico. Os segmentos passam, então, a não ser só um conjunto desorganizado de traços, mas um conjunto que possui uma estrutura interna organizada hierarquicamente.

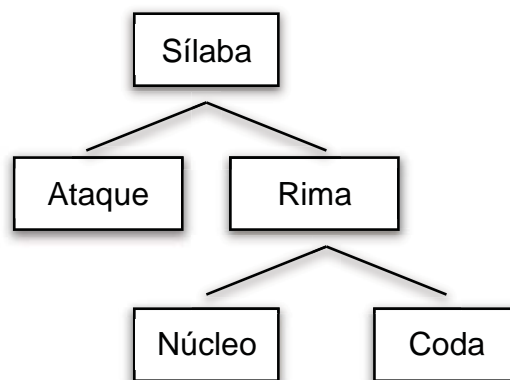


Figura 9. Modelo Hierárquico de sílaba (SELKIRK, 1982)

Na Figura 9, uma sílaba pode ser representada como uma unidade prosódica hierarquizada e estruturada, sendo os seus constituintes maiores o *Ataque* e a *Rima*. O

⁴¹ O encontro consonantal <dl> não existe em português. Para o encontro <vr>, não se observa a possibilidade de ocorrência em início de palavra. Para o encontro consonantal <vl>, a única possibilidade é a de nomes próprios de origem estrangeira. Com o encontro consonantal <tl>, temos a ocorrência de palavras que iniciam por ele apenas em uma onomatopeia tlim (imitação de sino, campainha, etc.), no nome de um indivíduo dos tlingites (povo indígena do Alasca, ou da língua falada por este povo indígena (FERREIRA, 2004).

primeiro, denominado Ataque (ou *Onset*), pode conter até duas consoantes. O segundo, chamado Rima, pode ser preenchido por vogais e consoantes, cujo número e sequência serão dados por especificidades da própria língua, como o molde silábico e filtros, e divide-se em *Núcleo* e *Coda*. O único constituinte de preenchimento obrigatório é o Núcleo, que pode ser simples, preenchido apenas por uma vogal (V); ou complexo, preenchido por uma vogal e um glide (G) (MATEUS; D'ANDRADE, 2003). Em outras línguas, entretanto, algumas consoantes podem preencher essa posição. Apesar de todos os segmentos [+consonânticos] no Português poderem ocupar a posição de Ataque, há, no entanto, restrições fonotáticas quanto à combinação desses segmentos no preenchimento da Coda. As sílabas que terminam em vogal são chamadas de sílabas *livres* ou *abertas* e as que terminam em consoante, de *travadas* ou *fechadas*.

As sílabas abertas predominam nas línguas do mundo e entre elas parece preponderar o tipo CV (consoante + vogal). Há línguas, como o japonês e muitas línguas africanas, que só aceitam sílabas desse tipo (SOUSA, 1998, p. 125).

Ao examinar o ataque silábico do PB, temos que considerar que ele pode ser preenchido por uma consoante (*ataque simples*, também chamada de sílaba aberta devido à ausência da coda, por exemplo, *pa.to*), podendo qualquer consoante ocupar essa posição. Todavia, as consoantes /t/, /k/ e /p/ não ocorrem em início de palavra, apenas em seu interior, mais precisamente entre vogais (i.e., *fa.ro.fa*, *mo.lho*, *so.nho*). Ademais, as sílabas em início de palavra como os fonemas /k/ e /p/ só ocorrem em vocábulos que são empréstimos no português, como <nhoque> (empréstimo italiano) e <lhama> (empréstimo hispânico). Além disso, na pronúncia de tais palavras acaba sendo inserida uma vogal inicial ([i'ɲokɪ] e [i'λẽma]), segundo afirma Cristófaros-Silva (2002). Denomina-se *ataque complexo* o ataque silábico preenchido por duas consoantes, devendo a segunda ser obrigatoriamente uma lateral líquida: /l/ ou uma não lateral /r/, independentemente de a sílaba ocorrer em posição inicial ou medial (i.e., *pra.to*, *pla.tô*).

A posição pós-vocálica de coda silábica pode ser ocupada por uma (*coda simples*) ou mais consoantes (*coda complexa*). Também nessa posição observam-se restrições para a ocorrência de fonemas do PB. Os segmentos que ocorrem em coda silábica são normalmente representados pelos arquifonemas /N/, /S/ e /R/ dada a variação decorrente nessa posição.

A consideração do arquifonema /N/ é dependente do ponto de vista teórico considerado. Como vimos anteriormente (seção 1.1.3.3), há uma corrente que argumenta que o sistema fonológico do PB apresenta sete vogais orais e cinco nasais. Assim, /ã/ e /a/ seriam fonemas

em nosso idioma, pois formam pares mínimos em palavras como <lá> e <lã>, uma vez que o que distingue as duas palavras são as vogais oral e nasal. Para essa corrente teórica não há necessidade de empregar o arquifonema /N/. Vimos também que uma outra corrente considera que o sistema fonológico do PB apresenta sete vogais orais e que a vogal nasal seria constituída por um segmento oral seguido de um segmento consonantal ([m, n, ŋ]), cujos traços fonológicos seriam neutralizados pelo arquifonema /N/. Isso porque tais segmentos consonantais são fonemas distintos no PB. Tal consoante nasal assimila o traço da consoante subsequente, sendo realizada como consoante bilabial diante de consoante bilabiais (bomba ['bõ^mba]); alveolar diante de consoantes alveolares (tonta ['tõⁿta]) e velar diante de consoantes velares (conga ['kõ^ŋga]).

O arquifonema /S/ é utilizado para indicar neutralização, ou seja, expressar falta de contraste fonêmico em ambientes específicos entre os fonemas /s/, /ʃ/ /z/ e /ʒ/, visto que tais fonemas, quando em determinados contextos, podem ser empregados alternativamente sem que haja modificação no significado de uma palavra. O arquifonema /S/ pode ser utilizado para indicar a presença de [s] ou [ʃ], dependendo do dialeto em que é pronunciado, em final de sílaba sucedido por consoante desvozeada ou em final de palavra ([ˈpas] – dialeto paulista, [ˈpaʃ] – dialeto carioca; ([ˈpasta] – dialeto paulista, [ˈpaʃta] – dialeto carioca). Pode ser ainda utilizado como [z] e [ʒ], dependendo do dialeto em que é proferido, em final de sílaba seguido por consoante vozeada, como por exemplo, [dʒisˈviʊ] – dialeto paulista; [dʒiʃˈviʊ] – dialeto carioca).

- O arquifonema /R/ é empregado também na posição de coda simples, pois em PB há contraste fonêmico entre o “r fraco” e o “r forte”. Assim, temos a título de exemplo, a transcrição fonológica /ˈmaR/ conforme suas possibilidades fonéticas decorrentes do dialeto falado: [ˈmah] para a pronúncia de Belo Horizonte, [ˈmax] para o Rio de Janeiro, [ˈmaɪ] para o falar caipira (falado no interior de São Paulo, sul de Minas Gerais, leste do Mato Grosso do Sul, sul de Goiás e norte do Paraná) e [ˈmar] para o paulistano.

Por fim, o arquifonema /L/ é empregado na posição pós-vocálica em decorrência da variante velar [ɫ] presente em algumas pronúncias do sul do Brasil e da variante vocalizada [ʊ] encontrada na maior parte do país.

A partir da caracterização do ataque e coda silábicos, podemos fazer o levantamento dos tipos de sílabas existentes no PB. Temos, portanto, as seguintes sílabas simples (constituídas apenas pelo núcleo silábico), complexas (cujo núcleo é seguido ou precedido por consoantes), abertas ou livres (quando não apresentam coda silábica) e fechadas ou travadas (quando possuem coda silábica): V (a), CV (pa), CCV (pla), VC (as), CCVC (plás), CVC (pas), CVCC

(pers), VG⁴² (au), VGC (aus), CVG (pau), CVGC (cais), CGV (qua), CGVC (ques), CGVGC (quais).

1.1.5.2 Acento

Ao analisar a Gramática Tradicional, podemos encontrar o acento tônico (primário) definido como “o modo de proferir um som ou grupo de sons com mais relevo do que outros” (BECHARA, 2009, p. 86). A mesma afirmação é feita por Rocha Lima (1972, p. 28), que define o acento como “a maior força expiratória com que uma sílaba se opõe às que lhe ficam contíguas no corpo dos vocábulos”. Assim, do ponto de vista fisiológico, acentuar uma sílaba é o mesmo que aplicar maior esforço muscular para sua enunciação, resultando em uma corrente de ar mais intensa.

Bechara (2009) diz que nossa marca acentual é de intensidade e não de tonicidade ou musicalidade, afirmando:

(...) que o acento é de *intensidade* (acento de força, acento dinâmico, acento expiratório ou icto), quando o relevo consiste no maior esforço expiratório. Diz-se que o acento é *musical* (acento de *altura* ou *tom*), quando o relevo consiste na elevação ou maior altura da voz.

O português e as demais línguas românicas, o inglês, o alemão, são línguas de *acento de intensidade*, o latim e o grego, por outro lado, possuem acento musical (Bechara, 2009, p. 86).

O português, o inglês e o alemão são exemplos de línguas de acento de intensidade, ou seja, nem todas as sílabas de uma palavra são pronunciadas da mesma forma, com a mesma intensidade. Como salienta Câmara Jr. (1964), apesar de o português ser uma língua cuja marca acentual é de intensidade, a gramática utiliza, no estudo do acento, a nomenclatura do grego, que é uma língua tonal.⁴³

Por essa razão, nossa gramática tradicional utiliza os termos *tônica* e *átona* (com tom e sem tom) para indicar a vogal com intensidade maior ou menor, respectivamente. Consequentemente, em português, quanto à posição do acento tônico, as palavras com duas ou mais sílabas podem ser classificadas em: oxítonas, quando o acento recai na última sílaba

⁴² A letra G maiúscula é utilizada aqui para simbolizar os glides.

⁴³ Diferentemente do português, os acentos em grego não significam uma maior intensidade da voz, e sim uma pequena elevação do tom da voz, conforme o tipo de acento (agudo, grave ou circunflexo). Fonte: SANTOS, Amador-Angel Garcia. *Gramática do grego do Novo Testamento*. Edições Loyola, 2002, p. 8.

(< tam.bém>, < a.mor>); paroxítonas, quando o acento recai na penúltima sílaba (ex.: < a.má.vel>, < de.mo.cra.ta>) e proparoxítonas, quando o acento recai na antepenúltima sílaba das palavras (< hi.pó.cri.ta>, < só.li.da>). Mais raramente, o acento pode incidir antes da antepenúltima sílaba do vocábulo, como acontece nos casos de epêntese vocálica (< té.c.ni.co>) e da clitização no sistema verbal (< fi.zé.sse.mos.-lhe>).⁴⁴

Ainda quanto ao acento, cabem algumas considerações sobre o que se denomina na literatura de *grupo de força*. Trata-se do fato de certas palavras, quando proferidas, perderem ou enfraquecerem seu acento de intensidade próprio, subordinando-se ao acento de outra palavra adjacente. As palavras compostas *girassol* e *aguardente*, por exemplo, resultaram de uma interpretação de um grupo de força e constituem, atualmente, palavras únicas. Assim:

O grande problema, no âmbito da língua oral, é que por vocábulo se entendem duas entidades diferentes. De um lado, há o vocábulo fonológico, que corresponde a uma divisão espontânea na cadeia da emissão vocal. De outro lado, há o vocábulo formal ou mórfico, quando um segmento fônico se individualiza em função de um significado específico que lhe é atribuído na língua. Há certa correspondência entre as duas entidades, mas elas não coincidem sempre e rigorosamente (CÂMARA JR, 1969, p. 34).

Portanto, para caracterizar a palavra fonológica no registro formal da pronúncia padrão do PB, Câmara Jr. (1970, p. 53) propõe uma pauta prosódica delineada em termos de algarismos. Sugere que marquemos a vogal tônica com um valor acentual 3. Esse valor é estabelecido em contraste com as demais vogais que tenham proeminência acentual, ou seja, as vogais pretônicas e postônicas. As vogais pretônicas são marcadas com o valor acentual 1, enquanto que as postônicas são marcadas com proeminência acentual 0. O valor acentual 2 é utilizado nos casos em que se observa uma sequência de vocábulos. Em uma sequência de segmentos ininterruptos, essa pauta permite identificar palavras fonológicas:

⁴⁴ No português, como todas as palavras proparoxítonas são acentuadas graficamente, não oferecem obstáculos ao processo de determinação automática da sílaba tônica. No entanto, as dificuldades surgem quando é necessário determinar se uma palavra não acentuada é oxítônica ou paroxítônica, principalmente quando ocorre encontro vocálico na última ou penúltima sílaba (CHBANE, 1994); nesses casos, faz-se necessária a existência de regras mais elaboradas.

(1) a. <i>habilidade</i>	<i>celebridade</i>	<i>arrocho</i>
a.bi.li.da.de	ce.le.bri.da.de	a.rro.cho
1 1 1 3 0	1 1 1 3 0	1 3 0
b. <i>hábil idade</i>	<i>célebre idade</i>	<i>ar roxo</i>
a.bi.li.da.de	ce.le.bri.da.de	a.rro.cho
2 0 1 3 0	2 0 0 3 0	2 3 0

Em (1a) podemos observar exemplos de vocábulos nos quais as vogais pretônicas apresentam valor 1, as tônicas valor 3 e as postônicas valor 0. Já em (1b) os vocábulos constituem um grupo de força no qual devemos marcar apenas uma proeminência acentual. Dessa maneira, a vogal que em (1a) possuía valor 3, quando em (1b) tem seu valor reduzido para 2.

Além de apresentar o valor demarcativo do acento, que estabelece a existência de vocábulos fonológicos, Câmara Jr. (1970) afirma que o acento tônico possui também valor distintivo, ou seja, é possível distinguir palavras conforme a posição do acento, como podemos observar nas palavras *sábia* x *sabia* x *sabiá*. O acento marca diferença no significado de palavras que podem pertencer a uma mesma classe gramatical (*sábia*, *sabiá*). Além dessa distinção semântica, a posição do acento contribui para a distinção entre formas nominais (*sábia*, *sabiá*) e verbais (*sabia*).

O modelo fonológico gerativo foi desenvolvido na década de 60, a partir da teoria gerativa clássica proposta por Chomsky, *Aspects of the theory of syntax* (1965). Em linhas gerais, Chomsky revolucionou os estudos linguísticos ao determinar que o componente sintático deveria ser o foco nas análises linguísticas, e não o sonoro. Além disso, introduziu a noção de *regras linguísticas* (regras que relacionam o som e seu significado), competência (o conhecimento da língua que o falante carrega), desempenho (o uso real que o falante faz da língua) e Gramática Universal (organização mental em relação à linguagem compartilhada por todos os humanos). Para Chomsky, as semelhanças presentes nas línguas naturais ocorreriam devido a uma essência linguística genética comum aos homens.

Segundo Massini-Cagliari (1992, p. 124), os estudiosos do gerativismo se dedicavam a formular regras a respeito da colocação do acento de palavra. E, segundo o que a gramática gerativa considerava “mais relevante para a avaliação das descrições, quanto mais simples,

quanto menor o número de regras e quanto menos símbolos utilizassem, melhor seriam as regras propostas”.

Ainda de acordo com a autora, o primeiro trabalho a respeito da acentuação do português, desenvolvido sob a perspectiva teórica da fonologia gerativa, foi o de Mateus (1975). Após críticas recebidas pelo fato de suas regras não abordarem a questão da colocação do acento em palavras proparoxítonas (esdrúxulas) e oxítonas (agudas), Mateus (1993 apud Massini-Cagliari, 1992) revê sua proposta sob uma perspectiva denominada pela própria autora de “fonologia gerativa revista/ampliada”.

No Brasil, conforme Massini-Cagliari (1992, p. 125), tem-se o trabalho de Costa (1978) que apresenta uma abordagem do acento do português dentro da perspectiva da fonologia gerativa padrão e da fonologia gerativa natural.⁴⁵ Ainda dentro da primeira perspectiva, a autora cita o trabalho de Maia (1981), “que resolve o problema da atribuição de acento às oxítonas elaborando uma regra que atribui acento à última vogal do radical” (MASSINI-CAGLIARI, 1992, p. 126). Com o advento da fonologia não-linear, em particular da fonologia métrica (Cf. BISOL, 1992; CAGLIARI, 1999) e da fonologia lexical (LEE, 1995), há uma mudança no enfoque da descrição do acento que antes era linear, segmental, e que passa a ser uma descrição não-linear, que considera os vários níveis hierárquicos (segmentos, sílabas, pés, palavras fonológicas, etc.).

Do ponto de vista fonológico, apesar de a grande maioria das unidades lexicais que compõem o léxico português, tal como em muitas outras línguas, se constituírem como palavras prosódicas (portadoras de acento de palavra), no entanto, há um grupo de vocábulos que ao serem pronunciados perdem seu acento próprio e se agregam aos vocábulos vizinhos. Por esse motivo são chamados de clíticos. Mais detalhes sobre esse assunto são apresentados a seguir.

1.1.5.3 Clítico

Sob o ponto de vista fonológico, os clíticos são considerados vocábulos desprovidos de acento prosódico, o que de certa forma permite explicar a tendência que possuem de “apoiar-se” no vocábulo adjacente acentuado (à direita ou à esquerda), ou seja, na palavra hospedeira ou hospedeiro do clítico. O processo que liga o clítico ao seu hospedeiro chama-se clitização.

⁴⁵ O modelo naturalista surgiu para propor novas interpretações para os processos fonológicos até então considerados complexos e, muitas vezes, abstratos. Para mais informações sobre o modelo, vide: (CRISTÓFARO-SILVA e YEHA, 2009).

Muitos pronomes pessoais de nosso idioma (nomeadamente, *o, a, os, as, lhe, lhes, se, me, te, nos, vos*) são clíticos, pois não constituem por si próprios uma unidade fonológica associada a um acento (isto é, uma *palavra fonológica*). Além destes pronomes, são também formas clíticas no português outros itens lexicais, tais como os artigos (definidos e indefinidos), os pronomes interrogativos (*que* e *por que*), o pronome relativo *que*, as conjunções *que, se, mas, e, ou*, as preposições *de, para, por, com, em* e o quantificador *cada* (Cf. MARTINS, 2013).

Bechara (2009, p. 89) afirma que em PB são, em geral, átonas e proclíticas as seguintes classes de palavras: artigos (definidos e indefinidos) numerais, pronomes, verbos auxiliares, (certos) advérbios, (certas) preposições e (certas) conjunções. Em relação aos enclíticos, ou seja, os vocábulos sujeitos à acentuação do vocábulo posterior, formando ambos um só vocábulo fonético, o autor declara que são enclíticas as formas pronominais *me, te, se, nos, vos, o, a, os, as, lhe, lhes* quando pospostas ao vocábulo tônico. No entanto, os pronomes oblíquos *comigo, contigo, conosco* e *convosco* resultam da combinação da preposição *com* e dos pronomes oblíquos correspondentes, não constituindo clíticos.

Todas essas formas sem acento próprio necessitam “se unir” (cliticizar) a uma palavra adjacente acentuada. O resultado gerado pela cliticização, ou seja, a unidade formada pelo clítico e o seu hospedeiro (por exemplo, numa frase como *o livro é meu*, a sequência *o + livro*) é uma nova palavra do ponto de vista da fonologia, na medida em que é uma unidade fonológica que contém um só acento de palavra (o mesmo da palavra hospedeira original). A cliticização é uma operação necessária, porque, do ponto de vista da fonologia, qualquer item lexical integrado numa frase tem de pertencer a uma palavra prosódica, ou seja, a uma unidade fonológica acentuada (MARTINS, 2013).

A maior parte dos clíticos em português cliticizam à esquerda da palavra acentuada, sendo portanto, denominados proclíticos (*o livro*). Quando localizados em posição posterior a da palavra acentuada são chamados de enclíticos (*vende-se*). Há ainda os casos em que a palavra acentuada é uma forma verbal do futuro ou do condicional, nos quais ocorrem em uma posição interna ao verbo (*dir-me-ás*). Nestes casos, tais pronomes são designados de mesoclíticos. Com exceção dos pronomes oblíquos, cuja posição pode ser proclítica, mesoclítica ou enclítica, a posição dos clíticos no PB é sempre proclítica.

Alguns clíticos, entretanto, são dissílabos: as contrações <pele> e <pela>, o artigo <uma>, as conjunções <como> e <porque> e as preposições <para> e <pera> (arcaico). Assim, a determinação dos casos em que se trata de um vocábulo átono (clítico) ou tônico

(palavra fonológica) torna-se um caso particular que só pode ser decidido quando proferidos em uma sentença.

Devido ao fato de o sistema PETRUS realizar a transcrição fonética de palavras isoladas, casos como esses são transcritos com a indicação de suas possíveis classes gramaticais, mas não apresentam a indicação da sílaba tônica, uma vez que constituem vocábulos átonos. Por exemplo, o vocábulo <para>, quando submetido para transcrição em nosso sistema, terá, além da forma transcrita, uma indicação (mas não desambiguação, uma vez que seria necessário acesso a sentença da qual faz parte) de suas possíveis classes gramaticais, quais sejam, prefixo, preposição e verbo. Nos casos (classes gramaticais) em que o dissílabo é tônico, há a marcação da sílaba tônica. Caso contrário, sendo o dissílabo um caso átono, essa informação não é mostrada ao usuário.

1.2 TRANSCRIÇÃO FONÉTICA

No princípio, antes do advento dos gravadores, as transcrições fonéticas eram utilizadas como forma de registro (gravação) da pronúncia de palavras. Logo, as transcrições fonéticas tornaram-se parte do ferramental padrão para a pesquisa em linguística descritiva (DESHPANDE, 1994), utilizada para investigar, por exemplo, o estilo de pronúncia de diferentes falantes de acordo com sua origem.

Além do campo da Fonética (LADEFOGED, 2003) e da Fonologia (LADEFOGED; MADDIESON, 1996), as transcrições fonéticas podem auxiliar grandemente em pesquisas das áreas da sociolinguística (CARDOSO, 2009), desordens da fala (WERTZNER, 2007), terapia da fala (HOWARD & HESELWOOD, 2002) e do ensino-aprendizagem de línguas ao indicar, nos dicionários, a pronúncia correta das palavras, ou ainda, em livros, ao mostrar os sons de uma língua que utiliza, por exemplo, um alfabeto não latino, contribuindo para que o aprendiz realize a pronúncia correta dos vocábulos no idioma que está aprendendo.

O meio de comunicação mais importante entre pessoas de diferentes culturas é a língua. Assim, aprender novas línguas possibilita realizar uma comunicação mais efetiva no mundo globalizado de hoje. Para compreender de forma efetiva uma língua estrangeira, os aprendizes precisam aprender seu vocabulário, gramática e pronúncia. Em relação ao último tópico, compreender o alfabeto fonético da língua que se deseja dominar habilita o aprendiz a extrair informações importantes sobre pronúncia a partir de um dicionário, por exemplo.

A esse respeito, Wells (1996, p. 239) diz que:

Hoje em dia os alunos de línguas estrangeiras deveriam ter muitas oportunidades de ouvir a língua falada, e não apenas a de seu professor e companheiros alunos. Televisão, fitas de vídeo, cassetes e CDs dão aos alunos de hoje uma vantagem, que as gerações anteriores não tiveram. No entanto, a simples exposição a um material de língua autêntico, embora certamente melhore a capacidade de compreensão do aluno, ele não é suficiente para garantir um bom domínio produtivo da língua ou uma boa pronúncia. Praticamente todos podem se beneficiar de ensino explícito de pronúncia, no qual o uso de transcrição fonética tem um papel importante. (WELLS, 1996, p. 239, tradução minha).⁴⁶

Para descrever de forma fidedigna a realização da fala, independentemente da língua em questão, o instrumento mais utilizado é a transcrição fonética. Realizar transcrição fonética é utilizar símbolos fonéticos para representar os sons da fala de forma sistemática. Assim, o objetivo principal de uma transcrição fonética é permitir que se possa ler, por meio dos símbolos fonéticos, qualquer língua, da forma exata como foi pronunciada, com a condição de que o leitor conheça os símbolos utilizados e suas convenções.

Uma palavra (UL) dificilmente é pronunciada da mesma maneira em todo o nosso país, devido às muitas variações linguísticas decorrentes da variedade de idades, escolaridade, regiões geográficas, entre outros fatores. Entretanto, o alfabeto que utilizamos na escrita tradicional não está apto para representar todos os tipos de sons existentes, porque – ainda que seja baseado nos sons das palavras, e não na ideia que estas representam (MASSINI-CAGLIARI, 1997, p. 26) – não há equivalência perfeita entre os sons e as letras. Para que fosse possível simbolizar a pronúncia real de um som via escrita, foram criados sistemas de notação gráfica dos sons, com símbolos e convenções que os representam, conhecidos como alfabetos fonéticos.

O *Alfabeto Fonético Internacional* (referenciado pela sigla AFI e pela sigla em inglês IPA, de *International Phonetic Alphabet*), vem sendo promovido desde 1886 pela Associação Fonética Internacional, como uma forma de representação padronizada dos sons constitutivos de todas as línguas naturais. Quando ocorreu a criação desse alfabeto fonético, a associação não havia sido criada, constituindo-se apenas por um grupo de professores franceses e ingleses, liderados pelo linguista Paul Passy.

⁴⁶ Trecho Original: “Nowadays learners of foreign languages ought to have ample opportunities of hearing the language spoken, and not just by their teacher and their fellow-pupils. Television, videotapes, cassettes and CDs give today’s learners an advantage, which earlier generations did not have. However, mere exposure to authentic language material, while it will certainly improve a learner’s comprehension ability, is not sufficient to ensure a good productive command of the language or a good pronunciation. Almost everyone can benefit from explicit pronunciation teaching, in which the use of phonetic transcription has an important role.” (Wells, 1996). Assim como esta, todas as traduções contidas neste trabalho são de minha responsabilidade.

O que motivou a criação do IPA foi o fato de os membros criadores dessa associação estarem focados em criar símbolos fonéticos para os quais diferentes articulações seriam associadas. Assim, poderiam promover a utilização da notação fonética em escolas com o objetivo de que as crianças pudessem adquirir uma pronúncia mais realista das línguas estrangeiras que aprendiam, além de ser uma maneira de ensinar as crianças mais novas a ler.

Muitos dos símbolos do IPA são derivados do alfabeto romano, e alguns do grego. O IPA não é apenas um alfabeto, mas também um conjunto de princípios de transcrição, os quais se diferenciam conforme as necessidades da transcrição a ser realizada. Assim, o mesmo enunciado poderá ser transcrito de diferentes maneiras, todas conforme os princípios do IPA. A partir de sua alteração/atualização mais recente ocorrida em 2005, passou a contar com 107 letras, 52 diacríticos, e quatro marcas de prosódia.⁴⁷ De tempos em tempos, símbolos são adicionados, removidos ou modificados pela Associação Fonética Internacional.

Apesar de o sistema IPA ser o mais utilizado, o *Speech Assessment Methods Phonetic Alphabet* (SAMPA) também é bem aceito e amplamente utilizado em sistemas computacionais que envolvem (lidam com) representação fonética. O SAMPA começou a ser desenvolvido no final da década de 1980, por um grupo de foneticistas do Projeto ESPRIT 1541,⁴⁸ com o objetivo de facilitar o intercâmbio de dados das transcrições utilizadas nas tecnologias que envolviam fonética, ou seja, criar um sistema de notação fonética computável (legível e processável por computador). Trata-se, portanto, de um alfabeto fonético baseado no IPA que utiliza caracteres ASCII de 7 bits ou ASCII estendido de 8 bits.⁴⁹

Quando o sistema Unicode não é apropriado, o SAMPA e o X-SAMPA⁵⁰ (Extended SAMPA) constituem a base mais robusta para codificação de notação fonética. O SAMPA consiste basicamente em um mapeamento dos símbolos do IPA em caracteres ASCII. Atualmente, o SAMPA contém sons de 26 idiomas, aproximadamente. Existem outros sistemas com o mesmo objetivo do SAMPA, como o Worldbet⁵¹ e o IPA/ASCII,⁵² também conhecido como formato Kirshenbaum.

⁴⁷ Alfabeto IPA: <http://www.langsci.ucl.ac.uk/ipa/ipachart.html>.

⁴⁸ Projeto ESPRIT: <https://aclweb.org/anthology/H/H91/H91-1007.pdf>.

⁴⁹ A memória do computador conserva todos os dados sob a forma numérica. Não existe um método para armazenar diretamente os caracteres. Assim, cada caractere possui o seu respectivo código numérico: o código ASCII (*American Standard Code for Information Interchange*). A codificação ASCII é usada para representar textos em computadores, equipamentos de comunicação, entre outros dispositivos que trabalham com texto. Desenvolvida a partir de 1960, grande parte das codificações de caracteres modernas herdaram como base. O código ASCII básico representava os caracteres em 7 bits (ou seja, 128 caracteres possíveis, de 0 a 127). Como tal código foi concebido para o inglês, portanto, não possui caracteres acentuados, por exemplo. Desse modo, para codificar caracteres específicos recorre-se ao ASCII estendido de 8 bits.

⁵⁰ Sobre X-SAMPA, vide <http://www.phon.ucl.ac.uk/home/sampa/x-sampa.htm>.

⁵¹ Sobre Worldbet, vide: <http://www.ling.ohio-state.edu/~edwards/WorldBet/worldbet.pdf>.

Ainda em relação aos códigos existentes para representação fonética em sistemas computacionais, resta citar o código Unicode, um sistema de codificação dos caracteres em 16 bits, inventado em 1991. O sistema Unicode permite representar qualquer caractere por um código em 16 bits, independentemente do sistema ou linguagem de programação. Por conta dessa vantagem, adotamo-lo para o sistema PETRUS.⁵³

Criada originalmente como uma ferramenta para preservação da herança cultural e religiosa, a transcrição fonética gradualmente passou a ser utilizada também com outros propósitos, tais como, ensino-aprendizagem de línguas, projetos de investigação e descrição de língua e, quando automatizada, contribui grandemente para a confecção de sistemas de acesso remoto ao correio eletrônico por telefone, serviços bancários e de compra de produtos também via telefone, sistemas de ajuda a pessoas com problemas motores, visuais ou de fala, sistemas de localização (GPS), interação com celulares, televisores e *videogames* via voz, entre outros. Como resultado dos diferentes propósitos a que se destina, pode ser realizada com diferentes graus de precisão, havendo, portanto, diferentes formas de transcrição para um mesmo enunciado. Assim, o termo genérico “transcrição fonética”, de acordo com Ladefoged (2006), pode abranger transcrições do tipo:

- *Fonêmica* ou *Alofônica*, segundo o tipo de símbolos fonéticos utilizados;
- *Estreita* ou *Larga*, segundo o nível de detalhe do conjunto de símbolos;
- *Sistemática* ou *Impressionista*, segundo o propósito para o qual são criadas.

A transcrição do tipo *fonêmica* representa a fala em termos de unidades básicas contrastantes, ou seja, *fonemas* do sistema sonoro de uma língua. Como já vimos, os fonemas podem ser definidos como símbolos abstratos que representam a menor unidade de fala capaz de distinguir significado. Em português, por exemplo, as palavras <mato> e <pato> possuem significado diferente apenas por conta do uso alternado dos sons [m] e [p]. Por isso, /m/ e /p/ são considerados dois fonemas distintos nesse idioma. Embora ocorram sobreposições entre

⁵² Sobre IPA/ASCII, vide: <http://www.kirshenbaum.net/IPA/index.html>.

⁵³ O Unicode foi projetado com base no código ASCII, sem contudo, ficar limitado ao alfabeto latino: O Unicode é capaz de codificar todos os caracteres presentes em qualquer uma das linguagens escritas existentes no mundo atualmente. Para tanto ele trabalha com um código de 16-bit ao invés dos 7 bits do código ASCII. Esta pequena diferença no número de bits faz com que o Unicode consiga reconhecer mais de 65.000 caracteres, enquanto que o ASCII só é capaz de reconhecer 128. A ausência de códigos diferenciados para caracteres especiais faz do Unicode um código mais simples e ao mesmo tempo mais eficiente. Os sistemas operacionais estão gradualmente adotando o Unicode.

os inventários fonêmicos das línguas naturais, cada língua é caracterizada por um inventário específico.⁵⁴

Transcrições que representam a pronúncia padrão dos fonemas em palavras sem refletir qualquer pronúncia alternativa devido à aplicação de regras fonológicas são denominadas de *transcrições fonêmicas*. Transcrições que representam a realização efetiva de fonemas, na medida em que são produto da aplicação de regras fonológicas são denominadas *transcrições alofônicas*.

A quantidade de símbolos diferentes com os quais se criam transcrições determina o grau de detalhamento que a transcrição pode representar. As transcrições que se limitam à identificação geral dos sons produzidos e utiliza um conjunto pequeno de símbolos fonéticos por não incluir, por exemplo, variações fonéticas decorrentes do contexto, são chamadas de *transcrições fonéticas largas*. A *transcrição fonética larga* apresenta apenas o conjunto de dados fonéticos que são relevantes para caracterizar o contraste fonológico, ou seja, a distinção entre palavras (LADEFOGED, 2006).

Transcrições fonéticas geradas com um conjunto de símbolos mais elaborado, composto por um maior número de símbolos e diacríticos, para marcar variações dos sons decorrentes do contexto (por exemplo, junto de vogais ou consoantes, no final de sílaba) no qual os sons foram produzidos são denominadas *transcrições fonéticas estreitas*. Esse tipo de transcrição é mais fiel ao uso da língua e permite apresentar produções particulares de certos fonemas, que podem variar consoante o falante ou o contexto fonético (International Phonetic Association, 2007). Uma *transcrição fonética estreita* pode representar, por exemplo, a realização específica das vogais ou consoantes finais e iniciais de palavras que encontram-se em sequência, devido à ocorrência de um processo fonético denominado ‘sândi externo’ (<casa amarela> [kazama'rela]) ou não representar certas vogais reduzidas que são suprimidas na oralidade (< pode crer > [pɔd'kre]).⁵⁵

Transcrições fonéticas são nomeadas como *transcrições sistemáticas* se elas representarem o sistema sonoro ou fonológico de uma língua. Transcrições *fonêmicas* e

⁵⁴ Assim, para uma transcrição fonética e uma fonológica da palavra *salto*, obteríamos os seguintes resultados, respectivamente: [ˈsawtu] e /ˈsaLto/. A respeito desses exemplos, é importante destacar que no primeiro caso, entre colchetes retos, foram marcados os sons produzidos em uma fala específica. Essa transcrição poderia ser diferente se o falante tivesse utilizado outros sons para pronunciar a mesma palavra, como o uma consoante fricativa palatal no final da sílaba (ex.: [ˈsaɫu]). Já a transcrição fonológica, entre barras oblíquas, não visa abarcar esse tipo de variação. Portanto, há apenas uma transcrição fonológica para cada palavra, em que são explicitados os sons que têm função distintiva, que são padronizados pela maior recorrência de uso. Assim, o som da consoante <l> em final de sílaba, que pode ser pronunciado [h] ou [w], é representado pelo arquifonema /L/.

⁵⁵ Embora os termos “larga” e “estreita” possam nos dar uma ideia de extremos, na verdade há um *continuum* de transcrições mais ou menos precisas.

alofônicas são *transcrições sistemáticas* porque elas representam apenas o contraste sonoro de uma língua ou as variantes fonológicas desses sons. Em contrapartida, as *transcrições impressionísticas* são foneticamente motivadas, pois seu propósito consiste em descrever a pronúncia dos fones com a maior quantidade possível de símbolos fonéticos, quanto o conjunto simbólico utilizado permitir; diferentemente da transcrição sistemática, cujo objetivo é descrever os sons no âmbito fonológico de uma língua específica.

Para os propósitos desta tese, é importante notar que ela aborda a investigação sobre e para a realização de uma *transcrição fonética larga* de modo automático.

1.3 TRANSCRIÇÃO MANUAL, AUTOMÁTICA E SEMIAUTOMÁTICA

As transcrições fonéticas podem ser obtidas de três formas: manual, semiautomática ou automaticamente. Enquanto as transcrições fonéticas manuais dependem exclusivamente do esforço e empenho de transcritores humanos, os procedimentos semiautomáticos e automáticos requerem o uso de um sistema computacional de transcrição, cujo *output* pode ser tanto tomado como correto e ser utilizado exatamente como foi produzido (transcrição automática) ou ser manualmente verificado e corrigido por transcritores humanos (transcrição semiautomática).

Ainda que a qualidade da transcrição fonética manual seja considerada em geral uma boa referência para avaliar a qualidade das transcrições produzidas automaticamente, a transcrição fonética de *cópus*, tradicionalmente realizada de maneira manual, é demorada e trabalhosa e, portanto, pode ser considerado um empreendimento caro tanto do ponto de vista da execução da tarefa quanto da contratação de especialistas para sua realização.

A esse respeito, Binnenpoorte (2006 apud Van Bael, 2007) diz que para a transcrição de um minuto de conversa coloquial gasta-se em torno de sessenta minutos para transcrevê-la. Ademais, a transcrição fonética manual está diretamente associada a importantes inconvenientes, tais como o fato de ser um processo moroso que exige grande atenção durante sua execução, promove ocorrência de erros humanos, quer por distração quer por cansaço do anotador (CUCCHIARINI, 1993).

Além disso, a opção pela realização de transcrição fonética manual pode dar margem à ocorrência de discordância(s) intrapessoal (uma vez que o mesmo trecho de enunciado pode ser transcrito de forma diferente pela mesma pessoa em dois momentos diferentes, quando se

tem um *córpus* de grandes dimensões, por exemplo) e interpessoal (o mesmo trecho transcrito de forma distinta por diferentes anotadores) (SHRIBERG; LOF, 1991). A falta de coerência entre anotadores, que podem deixar sua subjetividade marcada durante a anotação, pode ser decorrente, muitas vezes, do uso de um protocolo de anotação não tão claro e objetivo ou, ainda, ser resultado do tipo de treinamento que receberam para realizar a transcrição fonética em questão.

Somando-se a esses problemas as dimensões dos *córpus* pretendidos serem cada vez maiores, tanto linguistas quanto engenheiros especialistas no desenvolvimento de tecnologias de voz sentiram a necessidade de recorrer a sistemas computacionais que executassem o processo de transcrição fonética de maneira automática (Cf. BINNENPOORTE, 2006). Conforme veremos a seguir, diferentes técnicas podem ser utilizadas para a obtenção de transcrições fonéticas de modo automático.

1.4 ESTADO DA ARTE EM TRANSCRIÇÃO FONÉTICA AUTOMÁTICA

O processo de transformar um texto de entrada, que se apresenta sob o formato ortográfico convencional, numa sequência de fones é chamado de conversão grafema-fone ou transcrição letra-som.

No entanto, a conversão grafema-fone comporta algumas dificuldades, uma vez que não existe uma correspondência biunívoca entre os grafemas e a sua realização oral, sendo particularmente difícil em línguas cujo mapeamento de grafemas em fones é altamente irregular, tais como o inglês e o francês. Mesmo em línguas com ortografias de base fonológica, como é o caso do português (MATEUS; VILLALVA, 2006), a correspondência entre seus sistemas ortográfico e fonético não é perfeita.

Contudo, em uma atualidade motivada para o uso de várias aplicações tecnológicas ativadas pela voz, como a que vivemos, tanto a indústria quanto pesquisadores da área de tecnologia de fala têm se dedicado à questão da conversão grafema-fone de forma automática.

Nesse contexto, diferentes abordagens têm sido propostas para a construção de sistemas de conversão grafema-fone nas mais diferentes línguas. A escolha de um determinado método depende, sobremaneira, da relação grafema-fone existente na língua-alvo na qual o sistema conversor será construído; dentre outros fatores, tais como os recursos disponíveis, a memória e a capacidade de armazenamento do sistema computacional que irá realizar tal processamento.

Diante disso, nas seções seguintes serão apresentadas, de forma concisa, as técnicas mais utilizadas na conversão grafema-fone. Considerando a dificuldade inerente em esboçar o estado da arte em transcrição fonética automática atual, dado o rápido avanço da tecnologia e do conhecimento na área de Síntese de Fala – a partir da qual a maioria dos sistemas de transcrição automática é derivada – procuramos destacar, da melhor forma que nos foi possível, as principais tendências e trabalhos.

1.4.1 Sistemas baseados em dicionários

A conversão de grafemas em fones pode parecer, em um primeiro momento, uma tarefa simples, que poderia ser realizada por meio da busca de palavras em um dicionário anotado foneticamente.⁵⁶ Assim, sempre que uma palavra estivesse contida no referido dicionário, então seria transcrita. No caso de línguas em que o mapeamento de letras em fone(ma)s é altamente irregular, como no inglês e no francês, uma das estratégias utilizadas para determinar a pronúncia de uma palavra consiste em armazenar a maior quantidade possível de palavras transcritas foneticamente em um dicionário.

O *CMU Pronouncing Dictionary* (também conhecido como *cmudict*) é um exemplo desse tipo de dicionário fonético. Trata-se de um repertório de palavras com pronúncias norte-americanas, legível por computador e de domínio público, criado pela Carnegie Mellon University (CMU). É comumente utilizado em aplicações que envolvem processamento de fala em inglês americano, tais como, o Sistema Festival⁵⁷ e o sistema de reconhecimento de fala CMUSphinx.⁵⁸ Sua última versão disponibilizada contém 133.746 entradas, organizadas da seguinte maneira: palavra <dois espaços> pronúncia. Possui um conjunto de 39 fones representados (transcritos) com símbolos do conjunto ARPAbet.⁵⁹ As vogais recebem também uma marca de acento: 0 para não tônica, 1 para acento primário e 2 para acento secundário.

Para o PB, temos dois dicionários disponibilizados também para uso gratuito. Encontram-se disponíveis para *download* no “FalaBrasil”, grupo pertencente ao “Laboratório de Processamento de Sinais” da Universidade Federal do Pará (UFPA), cujo objetivo principal é desenvolver e disponibilizar livremente recursos úteis para o desenvolvimento de

⁵⁶ No contexto da conversão grafema-fone, o termo ‘dicionário fonético’ é utilizado para caracterizar uma lista de palavras cuja transcrição fonética é apresentada à frente de cada palavra.

⁵⁷ The Festival Speech Synthesis System: <http://www.cstr.ed.ac.uk/projects/festival/>.

⁵⁸ CMUSphinx Open Source Toolkit for Speech Recognition: <http://cmusphinx.sourceforge.net/wiki/>.

⁵⁹ Arpabet: <http://en.wikipedia.org/wiki/Arpabet>.

Tecnologias de Fala para o PB.⁶⁰ Um dos dicionários disponibilizados é denominado UFPAdic.2.0 e possui 32 fones baseados no alfabeto SAMPA. O outro, UFPAdic.3.0, possui 38 fones também representados com símbolos do alfabeto SAMPA. Tais dicionários apresentam a configuração palavra <dois espaços> pronúncia, mas diferentemente do *CMU Pronouncing Dictionary* (cmudict), não há qualquer marcação quanto à tonicidade das palavras.

Como alternativa, para reduzir a quantidade de entradas em um dicionário fonético, é comum a utilização de um dicionário de morfemas, em vez de um dicionário de pronúncia com palavras completas. O morfema pode ser definido como a unidade mínima dotada de significado que constitui uma palavra. Considere, a título de exemplo, a palavra ‘independentes’. Decompondo-a em unidades básicas (morfemas), observamos que ela é constituída pelo radical ‘dependente’, que é precedido pelo prefixo ‘in-’ e sucedido pela desinência de plural ‘-s’.

Do ponto de vista prático, um dicionário de morfemas apresenta a vantagem de ser muito mais compacto do que um dicionário de pronúncias composto por palavras completas. Nesse caso, a transcrição fonética é obtida a partir de um conjunto de regras flexionais, derivacionais e de composição morfofonêmicas que especificam as modificações sofridas pelos morfemas quando combinados para formarem palavras (DUTOIT, 1994). Essa abordagem tem sido adotada pelo sistema MITalk (ALLEN et al., 1987), um sistema desenvolvido para realizar a conversão automática de um texto em inglês em uma fala sintética. Para tanto, um dicionário com 12.000 morfemas do inglês foi suficiente para transcrever por volta de 95% das palavras utilizadas como entrada no sistema MITalk.

Apesar de simples, no caso de a palavra a ser transcrita estar contida no dicionário, essa estratégia apresenta algumas limitações. De fato, se construído manualmente, torna-se um processo longo e custoso; além disso, o tamanho desses dicionários impede a sua utilização em sistemas embarcados.⁶¹ Ademais, um sistema de transcrição baseado em dicionário tem desempenho limitado, pois não está apto a transcrever palavras que nele não estejam contidas. Por fim, desenvolver um dicionário que contenha todas as palavras de uma língua com suas respectivas flexões seria impossível, dado que todos os dias novas palavras são criadas e

⁶⁰ FalaBrasil: <http://www.laps.ufpa.br/falabrasil/downloads.php>.

⁶¹ Um sistema embarcado (ou sistema embutido) é um sistema microprocessado no qual o computador é completamente encapsulado ou dedicado ao dispositivo ou sistema que ele controla. Diferente de computadores de propósito geral, como o computador pessoal, um sistema embarcado realiza um conjunto de tarefas predefinidas, geralmente com requisitos específicos. Já que o sistema é dedicado a tarefas específicas, através de engenharia pode-se otimizar o projeto reduzindo tamanho, recursos computacionais e custo do produto. Fisicamente, os sistemas embarcados passam desde MP3 *players* a semáforos.

incorporadas às línguas.

1.4.2 Sistemas baseados em regras

Há estratégias que podem ser utilizadas para evitar a utilização de dicionários de pronúncia extensos. A aplicação de regras de transcrição, por exemplo, tem por objetivo tratar das correspondências entre letras e sons.

No caso do português, em que há uma grande regularidade entre a representação ortográfica e a transcrição fonética, pode-se determinar a pronúncia correta da maioria das palavras através da aplicação pura e simples de regras de transcrição (OLIVEIRA, 1996; CASEIRO et al., 2002; BARBOSA et al., 2003; TEIXEIRA, 2004; SILVA et al., 2006; BRAGA, 2008; SAMPAIO NETO, 2011).

Braga (2008) afirma que a utilização de regras linguísticas é, sem dúvida, a melhor estratégia para resolver o problema da conversão grafema-fone em PE, pois é mais econômica em termos de memória computacional e permite desempenhos maiores que os métodos automáticos. De acordo com Taylor (2005), essa técnica é também bastante eficaz para lidar com as palavras novas, que vão sendo permanentemente adicionadas ao léxico.

Oliveira (2009, p. 113) advoga em defesa da utilização dessa mesma técnica ao dizer que o desenvolvimento de algoritmos de conversão grafema-fone, com base em regras, pode ser ainda um tópico de investigação de grande interesse para a Linguística. Isso porque um sistema desse gênero poderia, e deveria ser encarado, como uma ferramenta para ensaios e avaliações de diferentes teorias linguísticas, pois o pesquisador teria a oportunidade de implementar computacionalmente uma gramática ou um modelo fonológico e acessá-lo sempre que necessário para monitorar e modificar as regras em tempo real, analisando e tentando corrigir os erros produzidos.

A aplicação de regras linguísticas na conversão grafema-fone parte do princípio de que uma sequência de grafemas pertencente a uma dada palavra pode ser convertida em uma sequência fonética a partir da análise do contexto dos grafemas que lhe são adjacentes. Assim, a transcrição de um texto ou palavra é feita a partir da aplicação de regras de produção do tipo: $A(B)C \rightarrow D$, dado que a realização de um fone está condicionada pelas propriedades dos fones a ele adjacentes. Portanto, o fone B será mapeado para o fone D quando seu contexto anterior for A e o posterior um C. Segundo Cagliari (2002, p. 29), as regras de contextualização de elementos podem também vir representadas como no seguinte esquema

formal: A → B/ CD. Significa que o elemento que precede a flecha ocorre no contexto descrito após a barra inclinada, isto é, entre os elementos D e C. Existem seqüências de grafemas no português que podem ser facilmente transcritas a partir da aplicação de regras desse tipo, como por exemplo, /t/ → tʃ / /i/.

Especificamente em relação ao português, o tratamento das consoantes costuma ser bem mais simples do que o das vogais, exceção feita à consoante ‘x’, cujo tratamento demanda um número bastante grande de regras, apesar de, ainda assim, não serem suficientes para transcrevê-la corretamente em todos os casos. Já as vogais <e> e <o> não acentuadas graficamente e em posição tônica de palavras podem alternar entre dois tipos de pronúncia possíveis, ora podendo ser pronunciadas de forma semiaberta [ɛ/ɔ], ora semifechada [e/o]. Em posição átona, <e> e <o> podem alternar ainda entre [e, ɛ, ɪ] e [o, ɔ, ʊ], respectivamente. As regras necessárias para dar conta da transcrição correta nesses casos são complexas e em número elevado, e nem sempre satisfatórias para todas as possibilidades (leia-se: contextos fonéticos) de ocorrência dessas vogais.

Vale ressaltar que a precisão dessa abordagem depende fortemente da qualidade com que as regras são geradas. A elaboração de regras significativas, que capturem a relação intrincada entre a forma gráfica e sua pronúncia, é uma tarefa desafiadora por si só, dependente de um conhecimento profundo a respeito da relação grafema-fone da língua em questão. Por isso, em geral, há a presença de linguista(s) tanto no desenvolvimento quanto na supervisão dos resultados gerados por essa abordagem baseada em regras.

Especificamente para o PB, podemos destacar o *Ortofon*, primeiro sistema proposto para realizar a conversão grafema-fone em PB, desenvolvido pelos investigadores do Laboratório de Fonética e Psicolinguística (LAFAPE) da Universidade Estadual de Campinas (UNICAMP) (ALBANO e MOREIRA, 1996; BARBOSA et al., 2003).⁶² O Ortofon faz parte de um sistema de síntese de fala construído para o PB denominado *Aiuruetê*, *software* de voz digital capaz de ler textos escritos em português, que foi desenvolvido em parceria entre as áreas de linguística e engenharia elétrica da UNICAMP (para mais informações sobre, vide BARBOSA et al., 1999). O Ortofon é composto por um conjunto de regras linguísticas e um dicionário de exceções. De acordo com os seus desenvolvedores (VIOLARO et al., 1996; BARBOSA et al., 1999), tal conjunto de regras está apto a reescrever fonicamente os caracteres ortográficos, marcar a acentuação lexical e as fronteiras silábicas. Quanto ao seu dicionário de exceções, afirmam ser constituído por 1.383 verbetes do *Minidicionário*

⁶² LAFAPE - Laboratório de Fonética e Psicolinguística da UNICAMP: <http://lafape.iel.unicamp.br>.

Aurélio, a partir do qual busca e transcreve palavras que são exceções às regras de transcrição. Infelizmente, o Ortofon não se encontra disponível gratuitamente, apenas para uso interno dos membros do grupo que o desenvolveu.

Mais recentemente, podemos citar o conversor grafema-fone baseado em regras, desenvolvido por Silva et al. (2006), que realiza também a marcação da tonicidade. Conforme afirmam os autores, o sistema por eles desenvolvido foi testado com um extrato do cópuz CETENFolha⁶³ e 97,44% dos fones gerados pelo sistema estavam corretos.

As regras propostas para marcação da tonicidade também foram testadas com um conjunto de 4.825 palavras desse mesmo cópuz, tendo sido obtida uma taxa de acerto de 98,58%.

Em relação ao PE, podemos citar os conversores grafema-fone desenvolvidos por Oliveira et al. (1991), Oliveira (1996), Teixeira et al. (1998), Teixeira (2004) e Braga (2008), que apresentam em comum o uso de um conjunto de regras para a conversão grafema-fone, outro para a separação silábica e um terceiro para atribuição da tonicidade.

Embora mais econômica em termos de memória, e de ser muito eficiente para a tarefa de conversão grafema-fone em português, no entanto, essa técnica apresenta algumas desvantagens. O número de regras é relativamente alto: por volta de 1500 para o inglês (BISANI; NEY, 2008), mais de 600 para o francês (DIVAY; VITALE, 1997), 112 para o romeno (STEFAN-ADRRAIN; DORU-PETRU, 2009) e, no nosso caso, para a construção do sistema PETRUS, foram necessárias 153 regras (sem contabilizar suas subdivisões) apenas para a transcrição fonética, além das regras para atribuição de tonicidade e de silabificação, aproveitadas de Silva et al. (2006).

Apesar de ser uma metodologia que contribui para a construção de sistemas de conversão mais robustos, a necessidade de conhecimento linguístico muito especializado para seu desenvolvimento converge para dois inconvenientes. O primeiro deles está relacionado com o fato de a concepção das regras ser uma tarefa morosa, difícil e cansativa, que requer conhecimento linguístico profundo sobre a língua-alvo. O segundo, diz respeito às irregularidades frequentes das línguas naturais, que precisam ser tratadas, por meio do uso de um dicionário de exceções, composto por palavras de pronúncia irregular (não cobertas pelas regras criadas), bem como por siglas ou estrangeirismos. Tal dicionário precisa ser

⁶³ CETENFolha (Corpus de Extractos de Textos Electrónicos NILC/Folha de S. Paulo) é um cópuz de cerca de 24 milhões de palavras em português brasileiro, criado com base nos textos do jornal Folha de S. Paulo que fazem parte do cópuz NILC/São Carlos, compilado pelo Núcleo Interinstitucional de Linguística Computacional (NILC).

constantemente incrementado, haja vista o fato de novas palavras serem criadas/incorporadas ao nosso idioma com certa frequência.

1.4.3 Sistemas baseados em dados

Com o advento da inteligência artificial, surgiram novas soluções que se propuseram também a resolver a questão da conversão grafema-fone automática. As mesmas foram criadas com base no conceito de que a partir de um grupo de palavras transcritas, é possível construir um sistema computacional que incorporará o conhecimento obtido por meio de um conjunto de exemplos (chamado de conjunto de treino).

Na prática, esses sistemas são treinados com pequenos dicionários fonéticos, gerados manualmente, contendo as palavras mais frequentes da língua-alvo. Quanto maior for o tamanho do dicionário utilizado, mais eficiente tornar-se-á o método. Assim, os maiores desafios nessa abordagem são o de se produzir dicionários para treino cada vez maiores, bem como o de produzir resultados que sejam melhores do que os apontados para a tarefa de conversão grafema-fone com a utilização da abordagem baseada em regras.

Entre os métodos baseados em dados para geração de conversores grafema-fone(ma), temos as seguintes abordagens utilizadas para o PE:

Redes neurais. A utilização de Redes Neurais em aplicações destinadas à conversão grafema-fone(ma) teve início na década de 80. Siravenha (2009) afirma que o sistema NETTALK que visava ao aprendizado de pronúncia da língua inglesa, publicado em 1987 por Sejnowski e Rosenberg, é o trabalho pioneiro que adota a técnica de redes neurais na realização do processo de conversão grafema-fonema. Segundo Siravenha (2009), uma rede multicamada treinada por um algoritmo *backpropagation* cumpria as tarefas de alinhamento e aprendizagem do referido *software*, e cada entrada da rede era baseada em um grafema cujo contexto era determinado pelas entradas (grafemas) vizinhas, esperando-se, dessa maneira, que os grafemas estivessem alinhados corretamente. Um exemplo de conversor gerado por essa técnica é o produzido por Trancoso et al. (1994).

Árvores de decisão. São modelos estatísticos que utilizam um treinamento supervisionado para a classificação e previsão de dados. Em outras palavras, em sua construção é utilizado um conjunto de treinamento formado por entradas e saídas. Árvores de decisão são geralmente aplicadas junto a grandes bases de dados. Para tanto, regularidades

implícitas presentes na base de dados devem ser descobertas automaticamente e expressas, predominantemente, na forma de regras. Conhecimentos de Inteligência Artificial e Estatística são comumente empregados para a geração de árvores de decisão. As árvores de decisão estão entre os mais populares algoritmos de inferência e têm sido aplicadas em áreas como diagnóstico médico e risco de crédito (Cf. Mitchell, 1997), e delas pode-se extrair regras do tipo “se-então” que podem ser implementadas computacionalmente. Oliveira et al. (2001) utilizaram essa técnica para a criação de um sistema transcritor.

Máquinas de estados finitos. Uma *máquina de estados finitos* ou *Autômato Finito* é um tipo de modelagem de um comportamento composto por estados, transições e ações. Um estado armazena informações sobre o passado, isto é, ele reflete as mudanças desde a entrada num estado, no início do sistema, até o momento presente. Uma transição indica uma mudança de estado e é descrita por uma condição que precisa ser realizada para que a transição ocorra. Uma ação é a descrição de uma atividade que deve ser realizada em um determinado momento. As máquinas de estados finitos foram originalmente definidas na Teoria de Autômatos e depois retomadas na Teoria da Computação. Em ciência da computação, máquinas de estados finitos são amplamente utilizadas na modelagem de comportamento de um dado aplicativo, no *design* de sistemas digitais de *hardware*, na engenharia de *software*, em compiladores, em protocolos de rede, em sistemas neurológicos e, em linguística, para descrever as gramáticas das línguas naturais. Caseiro (2003) e Oliveira et al. (2004) utilizam essa técnica para desenvolver um conversor para o PE.

Transformation-Based Learning (TBL). O *Aprendizado Baseado em Transformação* é uma abordagem que reúne tanto características referentes ao uso de regras fonológicas (uma vez que o TBL é baseado em regras que especificam quais rótulos devem ser associados a quais palavras), quanto ao uso de probabilidades, já que se trata também de uma técnica de aprendizado automático em que as regras são induzidas automaticamente a partir dos dados. O TBL é uma técnica supervisionada de aprendizado de máquina, isto é, pressupõe um *cópus* de treinamento já anotado, no nosso caso, foneticamente, para que as regras sejam aprendidas e aprimoradas automaticamente. A lógica de seu funcionamento é bastante simples, pois, inicialmente, anota-se um *cópus* automaticamente, assumindo-se que a *tag* (etiqueta) de uma palavra é a sua *tag* mais frequente (segundo um *cópus*/léxico). Em seguida, verificam-se os erros cometidos (comparando-se com a anotação humana correspondente) e, dentre todas as

possibilidades de correção, monta-se uma regra de correção com maior precisão. A seguir, aplica-se essa regra nova em todo o *cópus*. No próximo passo, verificam-se novamente os erros cometidos e monta-se uma segunda regra de correção com maior precisão, e assim por diante, até não se obter mais melhora de *performance*. Oliveira (2009) utiliza essa técnica para desenvolver seu sistema conversor.

Memory-Based Learning (MBL). O Aprendizado Baseado em Memória é outro método de autoaprendizagem, que pode ser considerado individualmente ou combinado com o TBL citado anteriormente. Trata-se de um método, que, de acordo com Siravenha (2009, p. 16), surgiu a partir da ideia de que exemplos podem servir como base para problemas de PLN. Assim, esses exemplos são treinados e posteriormente guardados sem que seja feito qualquer tipo de alteração. No momento do processo de classificação, os exemplos mais similares aos dados de treino são arquivados e sua classificação é usada para classificar novos exemplos. Teixeira et al. (2006a; 2006b) utilizam tal técnica para desenvolver um conversor grafema-fone dedicado ao PE.

Modelos de Máxima entropia. É uma abordagem muito utilizada para a resolução de ambiguidades na área de PLN. Com a adoção dessa abordagem, muitos problemas podem ser reformulados como uma tarefa de classificação, que consiste em, a partir de um contexto, prever uma classe correta. Essa abordagem foi adotada por Barros e Weiss (2006) para a construção de um sistema de conversão para o PE, de um sistema silabificador, e também de um sistema que marca a sílaba tônica em palavras.

Em relação ao PB, temos conhecimento sobre o uso das seguintes técnicas baseadas em dados para a confecção de sistemas conversores grafema-fone:

CARTs – Classification and Regression Trees. É uma técnica de aprendizado de máquina que já provou ser útil para prever a prosódia em sistemas que realizam síntese de voz (MANA; QUAZZA, 1995). O conhecimento extraído por essa técnica é representado em uma árvore binária, sendo esse tipo de representação interessante quando é necessário integrar um sistema conversor grafema-fone em aplicações de tempo real, já que as regras são armazenadas em um formato compacto, o que poupa memória, tornando o tempo necessário de execução pequeno. O sistema produzido por Mana et al. (2001) utiliza essa técnica.

Programação Genética. A descoberta automática de regras para a conversão grafema-fone, a partir de exemplos, por meio da Programação Genética é uma atividade que envolve a combinação de instruções, tipicamente condicionais (IF), lógicas (AND, OR) e variáveis de entrada como letras e fonemas, buscando obter programas que produzam os fonemas corretos, em relação a um conjunto de letras. O sistema desenvolvido por Frazen e Barone (2002) utiliza programação genética.

Aprendizado de Máquina. Hosn et al. (2006) adotam uma abordagem de aprendizado de máquina constituída por dois passos, responsáveis por derivar automaticamente, a partir de um cópula de treino, um classificador, que pode ser utilizado para realizar a tarefa de conversão grafema-fone. De acordo com os autores, no primeiro passo dessa abordagem, há a necessidade de se realizar um alinhamento entre grafemas e seus respectivos fonemas, o qual foi realizado segundo o método descrito em Damper et al. (2004). Hosn et al. (2006) afirmam que esse alinhamento é crucial para todo o esquema de treinamento de sistemas conversores que adotam esse tipo de abordagem, uma vez que é a partir dele que são construídos os dados dos quais os métodos de aprendizagem extraem automaticamente as regras de conversão. No segundo passo, os autores utilizam o *software* para mineração de dados denominado WEKA.⁶⁴ A mineração de dados pode ser definida, em poucas palavras, como sendo a transformação de grandes quantidades de dados em padrões e regras significativos. Ela (mineração de dados) não é usada estritamente para produzir um número absoluto, mas sim para criar um modelo que permite detectar padrões, prever a saída, e tirar conclusões com base em dados. Hosn et al. (2006) utilizaram os algoritmos J48 e Naive Bayes da suíte WEKA, para construir dois classificadores, nos quais cada fone representa uma classe. Ao classificar cada fone, os autores obtiveram quatro tipos de resultados com relação ao desempenho dos classificadores desenvolvidos:

1. Verdadeiro positivo: quantidade de fones que estão no conjunto de teste e que o sistema classificador conseguiu identificar;
2. Falso negativo: quantidade de fones que estão no conjunto de teste mas que o sistema não conseguiu identificar;

⁶⁴ Weka, sigla utilizada para designar “Waikato Environment for Knowledge Analysis”, é uma suíte formada por um conjunto de implementações de algoritmos de diversas técnicas de mineração de dados. Algoritmos como Id3, C45, J48, ADTree, UserClassifier, PredictionNode, Splitter, ClassifierTree, M5Primer, Prism, Part, OneR são classificadores presentes no Weka que produzem regras ou árvores de decisão. O WEKA é totalmente gratuito e produto da Universidade de Waikato (Nova Zelândia), que foi implementado pela primeira vez em 1997. Encontra-se disponível em: <http://www.cs.waikato.ac.nz/ml/weka>. Acessado em 28 de Agosto de 2014.

3. Falso positivo: quantidade de fones que não estão no conjunto de teste mas que o sistema identificou;
4. Verdadeiro negativo: quantidade de fones que não estão no conjunto de teste e que o sistema não identificou.

Para medir a efetividade das abordagens utilizadas para resolver problemas de aprendizado de máquina, são utilizadas as seguintes métricas de avaliação: acurácia (*accuracy*), precisão (*precision*), revocação (*recall*) e *F-measure* (F1). Acurácia mede a soma total de acertos em todas as categorias. Precisão é a razão dos elementos classificados corretamente na categoria dividido pelo número de elementos classificados nessa categoria, enquanto a revocação é o número de acertos corretamente classificadas na categoria dividido pelo número de elementos da categoria. Já F1 sumariza as métricas precisão e revocação.

A diferença principal entre a abordagem baseada em regras geradas manualmente e a grande maioria desses métodos para a geração automática de regras de transcrição com base em dados reside, logicamente, no modo como as regras são desenvolvidas: no primeiro caso, elas são explicitamente formuladas à mão por especialistas, enquanto que métodos automáticos implicam em uma aprendizagem a partir de dados (dicionários fonéticos). A opção por adotar uma abordagem com regras geradas manual ou automaticamente depende sobremaneira dos recursos (dados, tempo, pessoas envolvidas) disponíveis para sua execução.

Em síntese, podemos dizer que a transcrição mediante regras parece ser uma estratégia bastante eficaz em línguas que se caracterizam por uma elevada regularidade entre a representação ortográfica e a transcrição fonética, como é o caso do português. Assim, podemos justificar a utilização preferencial de regras (geradas manualmente ou via automática) em sistemas de conversão grafema-fone(ma) dedicados ao PE (Oliveira et al. (1991; 1992; 1996); Teixeira et al. (1998; 2006a, 2006b); Barbosa et al. (2003); Braga (2008); Veiga et al. (2011, 2013); Simões, 1999; Albano & Moreira, 1996; Albano & Aquino, 1997)) e ao PB (Simões (1999); Mana et al. (2001); Barbosa et al. (2003); Vasilévski (2008), Almeida (2005); Silva et al. (2006); Latsch (2011), Sampaio Neto (2011)).

Justificamos a adoção dessa abordagem para a construção do sistema PETRUS, tendo como base as seguintes quatro premissas: a primeira delas diz respeito ao fato de o português ser uma língua que apresenta regularidades na correspondência entre letras e sons; a segunda, tem como base a economia, em termos de memória computacional, que a abordagem por regras permite em detrimento do uso da abordagem baseada em dicionário. Já a terceira diz respeito ao fato de que os conversores grafema-fone(ma) produzidos com o uso de métodos estatísticos (aprendizagem automática das regras de transcrição) para o PB não mostraram

uma taxa de acerto maior do que as reportadas por sistemas construídos com base em regras. E por último, o fato de a abordagem baseada em regras ser sempre capaz de ler uma palavra nova.

1.4.4 Sistemas híbridos

Como o próprio nome sugere, consiste no uso combinado de duas ou mais abordagens para gerar um sistema conversor, de modo que se consiga diminuir os problemas decorrentes do uso de abordagens em separado. Veiga (2013), por exemplo, propõe o uso de uma abordagem híbrida para a solução de problemas da conversão grafema-fone em PE, utilizando um modelo de base estatística para a conversão grafema-fonema ao qual são incorporadas regras fonológicas para marcação da vogal tônica. Segundo o autor, a inclusão de informação sobre a tonicidade da vogal foi decisiva para o aumento do desempenho de seu conversor. O sistema desenvolvido por Veiga encontra-se disponível *online* para uso gratuito.⁶⁵

1.4.5 Por que um novo sistema de transcrição?

Em geral, os sistemas de conversão grafema-fone são desenvolvidos como um dos módulos necessários ao bom funcionamento de um sistema de síntese de fala.⁶⁶ O que se apresenta neste trabalho é uma proposta de ampliação da possibilidade de sua aplicação: inserção de transcrição fonética automática em dicionários do PB. Para tanto, poderíamos partir do pressuposto de que nosso objetivo poderia ter sido atendido facilmente com a incorporação prática e simples de um sistema já existente, como, por exemplo, um dos sistemas já citados nesse capítulo. Entretanto, não é tão simples quanto possa parecer.

Uma revisão bibliográfica realizada acerca de sistemas conversores dedicados à conversão grafema-fone para português (europeu e brasileiro) nos indicou um cenário desfavorável para o livre acesso a tais sistemas. Em geral, esses sistemas foram produzidos para serem utilizados por um grupo de pesquisa específico, com acesso restrito apenas aos seus membros, como é o caso, por exemplo, do sistema Ortofon, já mencionado na seção 1.4.2. A mesma revisão possibilitou verificar uma escassa partilha dos algoritmos

⁶⁵ Sistema conversor desenvolvido por Veiga (2013): <http://lsi.co.it.pt/spl/g2p/grafone.html>.

⁶⁶ Para mais informações sobre módulos que constituem um sistema de síntese de fala geral, vide Braga (2008).

implementados pelos referidos sistemas. Exceção à regra, no entanto, é o sistema de conversão produzido pelo grupo “FalaBrasil”. Trata-se de um sistema de conversão com determinação de vogal tônica. Segundo Sampaio Neto (2011, p. 25), foi gerado a partir de uma codificação do conjunto de regras descritas por Silva et al. (2006), com algumas adaptações (regras relativas a transcrição de <a, i, u, x>). O conversor proposto por Sampaio Neto lida apenas com palavras isoladas e não realiza análise coarticulatória entre elas. Entretanto, ressalva o autor, que as regras de seu conversor não refletem nenhum dialeto específico do PB. Assim, como um de seus trabalhos futuros, aponta a intenção de implementar uma nova versão de seu conversor, adaptando as regras desenvolvidas ao dialeto falado na cidade de Belém do Pará. O conversor (um arquivo executável) encontra-se publicamente disponível para uso.⁶⁷

Poderíamos, portanto, partir do sistema desenvolvido por Sampaio Neto (2011) para produzir o sistema transcritor aqui proposto, sendo necessário apenas a construção de uma interface gráfica para interação por usuários não familiarizados com linguagens de programação, já que sua utilização atual se dá por meio de um arquivo executável. Contudo, tal opção se tornou inviável devido aos seguintes inconvenientes:

- O alfabeto fonético utilizado pelo sistema de Sampaio Neto (2011) é derivado do sistema de notação SAMPA e, em obras lexicográficas, há preferência pelo uso do IPA. O alfabeto SAMPA foi opção do autor, pois trata-se de uma alfabeto fonético computável, ou seja, legível por computador, portanto, uma opção importante de ser utilizada na confecção de um sistema computacional que visa o processamento de fala. Como dito anteriormente, o objetivo do grupo “FalaBrasil” é o de produzir recursos para o desenvolvimento de tecnologias de fala para o PB, diferente do nosso, que visa a atender a demanda de lexicógrafos e demais interessados em obter transcrição fonética automática em IPA.
- Não há uma variedade linguística considerada na confecção do sistema, como relatado por Sampaio Neto (2011). Esse fato é comum em síntese de fala, por isso ocorrem certos estranhamentos quando ouvimos um sistema de síntese de voz (como o GPS) falar com ‘s’ carioca e ‘r’ paulista, por exemplo, ou nos dizer para contornar a “rotunda” (em PE) e não a “rotatória” (em PB). Entretanto, quando se deseja inserir transcrição fonética em uma obra lexicográfica, é importante fazer a opção por uma variedade linguística. Assim, construir um sistema dedicado à inserção de transcrição fonética em dicionários implica

⁶⁷ Conversor publicamente disponível em: <http://www.laps.ufpa.br/falabrasil/downloads.php>.

uma série de escolhas que não são simples, já que envolvem aspectos linguísticos importantes (por conta disso, muitos materiais se abstêm de oferecer as informações sobre a pronúncia das entradas ao consulente, por exemplo). Tais aspectos compreendem desde o sistema sonoro da língua-alvo, tanto em nível segmental (fonema) quanto suprasegmental (acento lexical primário e entoação), até as questões relacionadas com a variedade linguística que será adotada como parâmetro.

- Além disso, há o fato de o sistema aqui pretendido ser fruto de um projeto de doutorado em Estudos Linguísticos. Assim, nada mais natural que aproveitar o conhecimento linguístico adquirido sobre a língua, para partir de uma variedade eleita e produzir manualmente as regras que permitem converter palavras grafadas nos respectivos correlatos sonoros dessa variedade, que podem facilmente ser implementadas no formato de algoritmos, dado seu formato lógico (se X, então Y, caso contrário, Z), por um especialista em linguagem de programação, como foi o caso deste trabalho.
- Por fim, podemos citar o fato de, apesar de ser sólida a investigação na área de conversão grafema-fone em português, notam-se, ainda, erros recorrentes nos sistemas desenvolvidos. No que diz respeito aos tipos de erros produzidos, constatou-se que a maior parte deles tem relação com a decisão do timbre⁶⁸ vocálico dos grafemas <e> e <o> decorrente do contexto em que podem ocorrer. Seguem-se a estes, erros de transcrição dos sons possíveis de <x>, siglas e estrangeirismos.

Nesse contexto, decidimos construir o sistema PETRUS, que difere dos sistemas existentes para o PB, pois:

- Possui acesso livre e gratuito via *Web*, cuja interface permite que mesmo um usuário não familiarizado com linhas de comando consiga interagir com o sistema PETRUS;
- Dedicar-se a auxiliar lexicógrafos, principalmente, na inclusão de transcrição fonética nas obras que são produzidas para o PB. Portanto, considerando os objetivos do público-alvo que queremos atender, foi adotado para o PETRUS o sistema de notação fonética IPA, amplamente utilizado por linguistas, fonoaudiólogos, lexicógrafos, entre outros profissionais, e não um alfabeto computável;

⁶⁸ Timbre é o efeito acústico resultante da distância entre o dorso da língua e o véu palatino, fazendo com que a cavidade bucal adquira função de caixa de ressonância. O timbre é o traço distintivo das vogais. Na vogal de timbre aberto, a língua se acha baixa, como é o caso das vogais /a/, /ɛ/, /ɔ/. Na vogal de timbre fechado, a língua se eleva, como em /e/, /o/, /i/, /u/. A distinção entre abertas e fechadas só se dá nas vogais tônicas e subtônicas; nas átonas desaparece a diferença entre /ɔ/ e /o/, /ɛ/ e /e/, e o /a/ reduzido é proferido com menos nitidez, como se pode depreender comparando-se os dois tipos em <casa>, em que o primeiro é aberto e o segundo é reduzido (BECHARA, 2009, p. 63).

- Independentemente da discussão que há em relação ao que se considera a variedade padrão do PB, adotamos para este trabalho o falar da Rede Globo. Trata-se de uma variedade inventada pela mídia, que se caracteriza pela pronúncia das pessoas cultas (com alto nível de escolaridade) expostas na mídia, em telejornais e em programas culturais, portanto, o mais comum a ser ouvido. De fato, os dicionários monolíngues do PB não apresentam, via de regra, transcrição fonética, realizam somente indicações acerca da abertura ou fechamento de vogais tônicas, da consoante <x> e de estrangeirismos. Frente a isso, a construção de um sistema para transcrição fonética automática da pronúncia brasileira em dicionários do PB, como o aqui proposto, revela-se uma importante iniciativa para a promoção e difusão dessa variedade do português no cenário internacional, via obra lexicográfica. Dada essa proposta, a busca de uma norma de pronúncia padrão do PB afigurou-se como uma decisão premente neste trabalho.

Apoiando-nos em reflexões feitas sobre a investigação dos conversores grafema-fone existentes para o PB e o PE, e nas teorias linguísticas apresentadas ao longo deste trabalho, cujos conteúdos se relacionam às características da variante linguística adotada, elaboramos como proposta de trabalho desta pesquisa um ambiente *Web* gratuito de suporte à transcrição fonética automática das unidades lexicais que fazem parte da cabeça do verbete de dicionários mono ou plurilíngues do PB, o PETRUS, que encontra-se disponível atualmente em: <http://www.nilc.icmc.usp.br/petrus/>.

O próximo capítulo é dedicado à apresentação detalhada do processo de conversão realizado pelo sistema proposto.

CAPÍTULO 2

METODOLOGIA

A concretização dos objetivos indicados na Introdução implicou na realização de um conjunto de tarefas acerca da construção de recursos linguísticos necessários ao funcionamento pleno da ferramenta de conversão grafema-fone aqui proposta que passamos a apresentar.

A realização da conversão automática de grafemas em fones é um tema que está longe de estar resolvido. Como vimos, não se trata de um simples mapeamento entre uma unidade gráfica e outra sonora, dada a não existência de uma correspondência biunívoca entre os grafemas e a sua realização oral, mas sim, de uma classe de problemas que a conversão grafema-fone inclui além da transcrição automática, a identificação da sílaba tônica, o alinhamento grafema-fone, a silabação, a análise morfológica. Desse modo, o conhecimento linguístico necessário ao desenvolvimento de cada um deles será apresentado a seguir, bem como a sequência de etapas contidas em seu funcionamento para que se obtenha, ao final, uma unidade léxica transcrita foneticamente, conforme a pronúncia escolhida para este trabalho.

A linguagem de programação utilizada para sua implementação foi Python⁶⁹ e Django,⁷⁰ o framework utilizado, o qual possibilita acesso multiplataforma, ou seja, o ambiente PETRUS pode ser acessado via computador, *tablet* ou *smartphone*.

⁶⁹ Python é uma linguagem de programação orientada a objetos robusta e dinâmica, desenhada para ser rápida tanto para aprender/entender quanto para programar. Python permite ter uma codificação limpa e uma sintaxe uniforme. Além disso, é uma linguagem muito usada para aplicações na Web e para o processamento da língua natural (PLN). Por exemplo, Python fornece a biblioteca NLTK (*Natural Language ToolKit*) que contém vários algoritmos de PLN usados na literatura. Para mais informações, vide: <http://www.python.org/>.

⁷⁰ Um *framework* é um conjunto de classes e funcionalidades implementadas numa linguagem de programação específica usadas para auxiliar o desenvolvimento de software. Django é um *framework* implementado em Python para o desenvolvimento de aplicações Web. A opção pelo Django se deu porque é um *framework* completo e bem documentado, em outras palavras, Django possui uma grande quantidade de funcionalidades especificadas de maneira completa e inteligível. Para mais informações sobre o Django, vide: <https://www.djangoproject.com/>.

O conteúdo das seções e subseções seguintes se subdividem em: a) apresentação de um diagrama para ilustrar a etapa que será descrita, b) descrição de como a etapa é realizada de modo automático pelo sistema proposto, e quando possível, d) um estado da arte sobre a etapa-alvo em outros sistemas de conversão grafema-fone do PB e do PE. Assim, tem-se uma ideia geral de todo o processo realizado de forma automática pelo PETRUS.

2.1 DIAGRAMA DO PROCESSO DE TRANSCRIÇÃO DO PETRUS

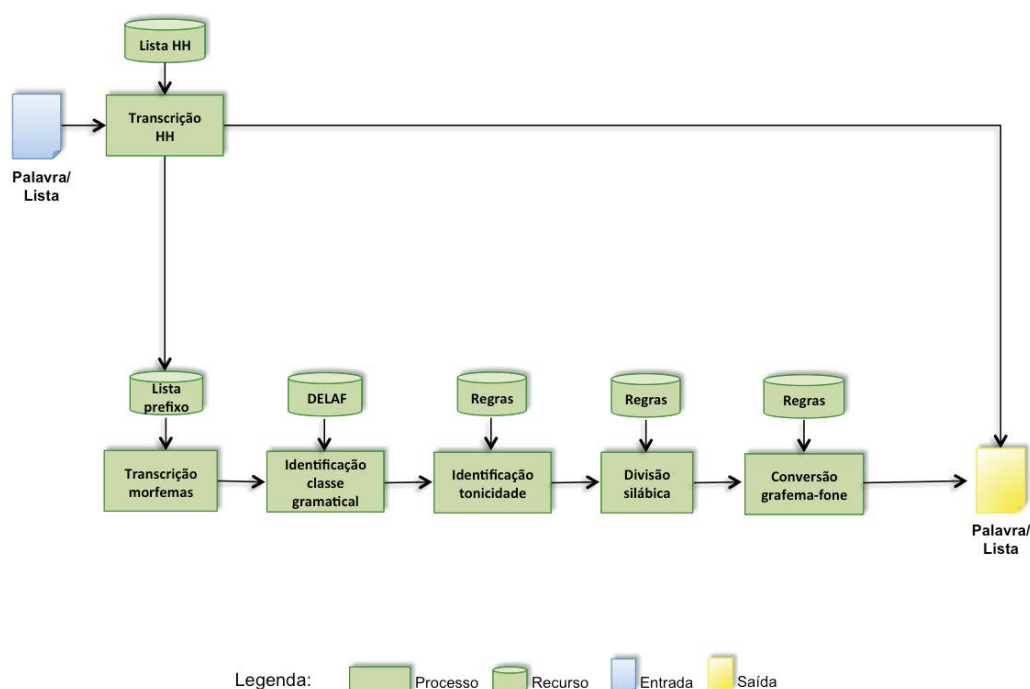


Figura 10. Diagrama do processo de transcrição fonética automática do sistema PETRUS (*Phonetic Transcriber for User Support*), composto por seis módulos.

Conforme pode ser observado a partir da arquitetura ilustrada na Figura 10, há seis etapas desde a submissão de uma palavra como entrada do sistema PETRUS até a obtenção de uma saída transcrita foneticamente de maneira automática. A depender da dificuldade de transcrição das palavras submetidas, poderão, ou não, passar por todas as etapas do sistema até que sua transcrição fonética esteja terminada. Cada uma dessas etapas compreendem as seguintes atividades específicas:

Etapa 1: Transcrição de homógrafos heterófonos (HH) – Primeira etapa do processo de transcrição fonética automática do sistema PETRUS. Consiste na identificação e posterior transcrição fonética de palavras homógrafas heterófonas, ou seja, palavras que possuem mesma grafia, mas pronúncia diferente. Se a palavra de entrada constar da lista de nosso sistema, será diretamente disponibilizada ao usuário, já com sua transcrição fonética, com referência explícita às fronteiras silábicas e sílaba tônica, sem que precise passar pelos outros módulos do sistema. Caso contrário, a palavra de entrada passará pelas cinco etapas seguintes, de modo a se obter, ao final, sua transcrição fonética, divisão silábica e classificação gramatical.

Etapa 2: Transcrição de morfemas localizados à esquerda da base – Etapa responsável pela transcrição fonética apenas de morfemas localizados à esquerda da base, uma vez que tais elementos mórficos não seguem as mesmas regras gerais utilizadas para a transcrição fonética de palavras do PB, quanto pelo comportamento das vogais médias [e] e [o], que podem possuir pronúncia aberta ou fechada. Como tal, precisam ser tratados separadamente.

Etapa 3: Identificação da classe gramatical – Refere-se à identificação automática da classe gramatical da palavra de entrada do sistema. Trata-se de uma informação importante para ser adicionada em uma obra lexicográfica, uma vez que indica a função que a palavra pode exercer na língua-alvo. Essa etapa pode ser importante também para uso interno pelo próprio sistema quando da atribuição da sonoridade às vogais médias tônicas que são problemáticas para a conversão automática em PB. Partindo do fato de que as vogais médio-altas [e, o] e médio-baixas [ɛ, ɔ] de nomes e verbos sofrem alternâncias regidas por regras, um estudo de caso, a ser desenvolvido a partir deste trabalho, pretenderá, futuramente, depreender os fatores fonéticos e morfológicos condicionadores de tais mudanças e formalizá-los como regras para transcrição fonética automática a serem implementadas no módulo 6 do ambiente PETRUS. Assim, o sistema poderá apresentar conjuntos de regras para transcrição fonética, de vogais tônicas, conforme a classe gramatical que a palavra em análise possui.

Etapa 4: Identificação automática da sílaba tônica – Etapa que realiza a identificação e marcação da vogal tônica da palavra de entrada do sistema PETRUS. Tal informação é disponibilizada para o usuário do sistema, mas que também serve à implementação de algumas regras de transcrição fonética.

Etapa 5: Divisão silábica automática – Este módulo se dedica a realizar a divisão silábica automática da palavra de entrada, através da aplicação de um algoritmo para silabificação em PE proposto por Mateus; D’Andrade (2003) e adaptado para o PB por Silva (2011).

Etapa 6: conversão grafema-fone – No tocante à conversão grafema-fone, adotamos uma metodologia baseada em regras – aproveitando a regularidade ortográfica brasileira – combinada com a técnica de dicionário de exceções (lista de HHs), com palavras cujas transcrições fogem às regras elaboradas manualmente. Com efeito, essa solução viabilizou o desenvolvimento de uma primeira versão do nosso sistema de transcrição automática, com desempenho aceitável.

Na sequência, todas essas etapas são descritas com maiores detalhes.

2.1.1 Etapa 1 - Transcrição fonética de homógrafos heterófonos

Por meio da Figura 11, ilustramos a primeira etapa de transcrição fonética no sistema PETRUS. Podemos observar que o primeiro passo consiste em verificar se se trata de um caso de HH. Para isso, o sistema recorre a uma lista composta por pares de HH transcritos foneticamente de acordo com a classe gramatical que possuem. Se confirmada essa hipótese, o sistema transcreve foneticamente tanto a palavra como seu par com base nessa lista. Caso contrário, tal palavra passará por todas as outras etapas (módulos) necessárias para que sua transcrição fonética seja realizada.

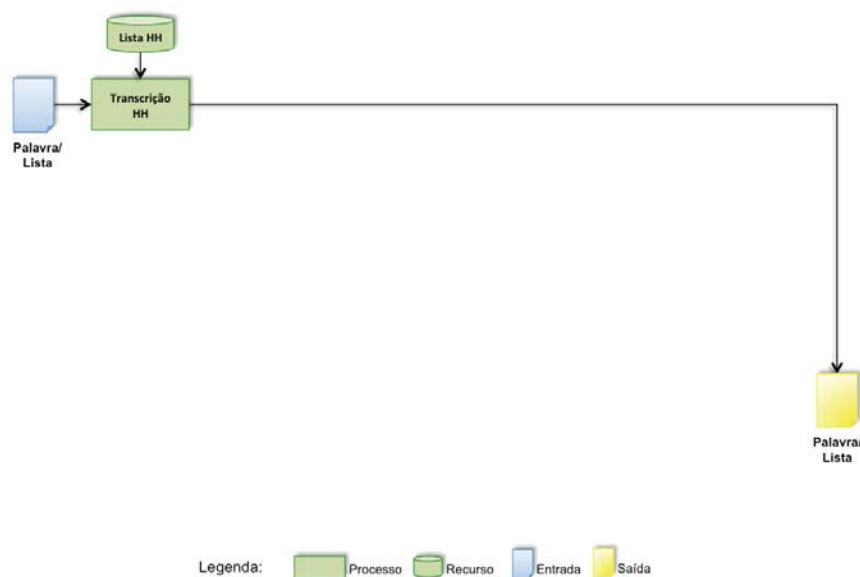


Figura 11. Primeira etapa de transcrição fonética do sistema PETRUS.

A decisão que os sistemas de conversão grafema-fone têm que tomar sobre a forma correta de transcrição de palavras HH não é um problema de solução trivial.

Em PB, sempre que um sistema como esse possui como *input* um caso de HH, os algoritmos responsáveis por realizar a conversão da sua forma ortográfica para seu equivalente fonético precisam optar entre duas situações possíveis: transcrever foneticamente a vogal tônica de tais palavras como [e]/[o] ou como [ɛ]/[ɔ]. Isso porque palavras HHs, tais como <corde> e <apelo> possuem mesma grafia, mas pronúncia e significado diferentes, dependendo de sua classe gramatical, tal como em “c[o]rte” – nome/“c[ɔ]rte” – verbo e “ap[e]lo” – nome/“ap[ɛ]lo” – verbo. Se não forem corretamente analisadas por um sistema que converte grafemas em fones, ou mesmo por um sintetizador de voz, poderão ser transcritas ou pronunciadas de forma inadequada.

Ainda que a quantidade de HHs existentes represente uma porcentagem muito pequena, quando comparados ao léxico geral de uma língua no contexto da conversão grafema-fone, a recorrência de erros quanto à transcrição fonética automática de HHs pode dar margem a uma avaliação inadequada da eficiência do sistema pelo usuário. Portanto, minimizar a ocorrência desse tipo de erro é fundamental para se obter uma melhor eficiência e, conseqüentemente, uma maior credibilidade do sistema pelo usuário.

A questão da resolução dos casos de ambigüidade que envolvem HHs é um assunto explorado pela comunidade científica de diversas línguas, dada a sua importância em síntese de voz para diversos idiomas. Yarowsky (1996) apresenta uma tipologia para a classificação

de pares de HHs em inglês, bem como as técnicas computacionais utilizadas tradicionalmente na desambiguação. Em Tesprasit et al. (2003), o tema da desambiguação de HHs é abordado considerando os idiomas japonês, tailandês e chinês, nos quais as palavras não apresentam delimitação de fronteira entre elas.

Em relação ao português, para a vertente europeia, temos os trabalhos de Ribeiro et al. (2002), que utilizam informação morfossintática, enquanto que Braga et al. (2007) fazem uso tanto de informação morfossintática quanto semântica para resolver os casos de ambiguidades. Já no PB, Seara et al. (2001, 2002) utilizam um analisador morfossintático que visa auxiliar na solução de problemas de alternâncias vocálicas entre verbos e substantivos, sem abordar a desambiguação semântica de homógrafos. Em Ferrari et al. (2003) é sugerida a aplicação de uma metodologia linguística com base na análise de córpus e na Gramática Cognitiva, para solucionar a questão da variação fonética de HHs. Apesar de ser um método que permite realizar a desambiguação morfossintática e semântica, o método proposto foi testado apenas com um único exemplo (a palavra <sede>), além de se revelar uma metodologia muito cara do ponto de vista computacional de implementação dos algoritmos que farão a desambiguação automática, pois para cada par de homógrafo existente é necessário realizar um estudo exaustivo das ocorrências contextuais. Silva et al. (2012) apresentam um conjunto de algoritmos baseado em regras para desambiguar casos de HHs em sistemas de síntese de voz produzidos para o PB. Os algoritmos foram implementados para resolver 111 pares de HHs organizados em 23 tipos de algoritmos de desambiguação.

Inspirados no trabalho de Silva et al. (2012), propomos realizar uma tipologia de HHs em PB, conforme a particularidade de sua oposição gramatical e alternância vocálica tônica com relação às vogais médias presentes nos 1.181 pares coletados do Dicionário de Palavras Homógrafas, de Walmírio de Macedo (1961), e do Novo Dicionário de Acentuação das Palavras Homógrafas Heterófonas, de Pandiá Pându (1972). Vale dizer que pretendemos, como trabalho futuro, enriquecer essa lista com HHs provenientes de obras lexicográficas mais recentes que as utilizadas até o presente momento, a saber: *Dicionário de Homônimos e Parônimos* (1987), de autoria de Osmar Barbosa e a obra *Um significado só é pouco: dicionário de formas homônimas do português contemporâneo do Brasil* (2010), publicado por Claudia Zavaglia.

Diante dos objetivos almejados neste projeto, diferentemente do proposto por Silva et al. (2012), não pretendemos desambiguá-los, mas sim, transcrever foneticamente o caso de palavra HH submetido como input ao PETRUS, bem como de seu par. Assim, partimos do pressuposto de que seja mais pertinente indicar aos lexicógrafos, ou até mesmo a um aprendiz

de língua, as possíveis pronúncias de pares de HHs do que realizar sua desambiguação.

Como o propósito deste projeto não envolve desambiguação, toda vez que for submetida ao sistema PETRUS uma palavra que esteja contida em sua base de dados com os HHs coletados, será apresentado como saída as duas pronúncias possíveis para cada palavra do par, associadas às suas respectivas classes gramaticais. Assim, por exemplo, se a palavra <gosto> for submetida como entrada, o usuário receberá como saída duas transcrições possíveis ([ˈgostu] e [ˈgɔstu]), associadas as suas classes gramaticais, dado ser um caso de HH.

Para a construção desse módulo, o primeiro passo consistiu na elaboração de uma lista com pares de HHs do PB, a partir de dois dicionários publicados sobre o referido tema: Macedo (1961) e Pându (1972), perfazendo um total de 1812 pares de HHs coletados, que depois de extraídos os pares repetidos nas duas obras, chegamos a um total de 1.181, sendo 1.040 pares com classe gramatical distinta e 141 pares com mesma classe gramatical.

Na Tabela 9 podemos observar o número de pares de HHs contidos em cada um dos dicionários citados e a quantidade final de pares de HHs utilizada neste trabalho, após a exclusão dos pares repetidos.

Tabela 9. Pares de homógrafos heterófonos coletados

	<i>Dicionário (1961)</i>	<i>Dicionário (1972)</i>
Pares de HHs	681	1.131
Pares de HHs com mesma categoria gramatical	49	92
Subtração de pares repetidos		772
Total		1181

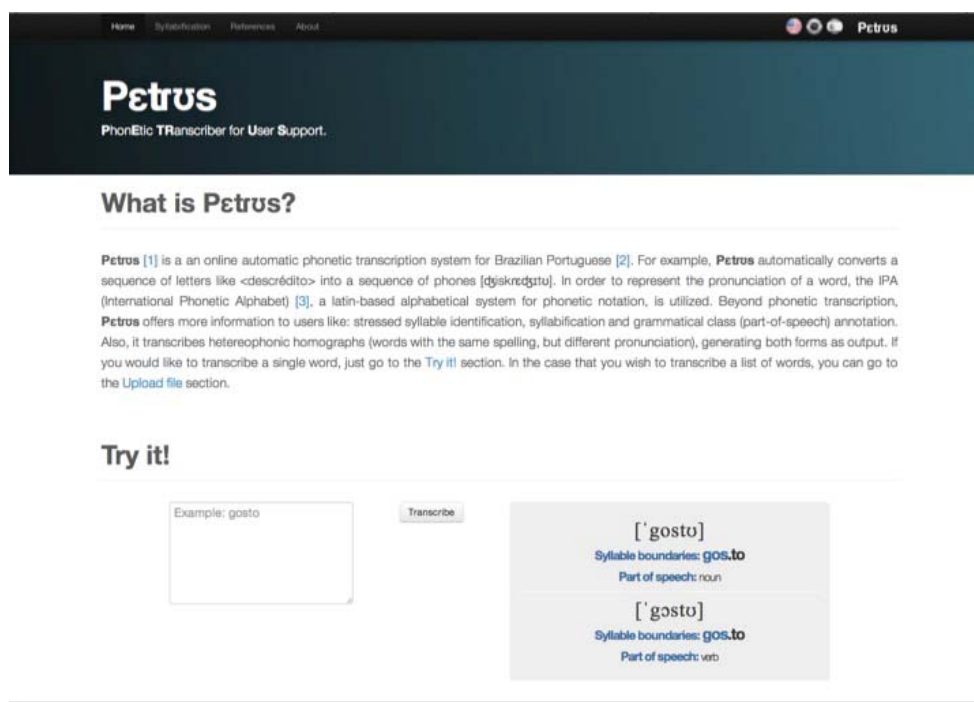
Após a elaboração da lista de HHs, o passo seguinte consistiu em realizar manualmente a transcrição fonética dos pares nela contidos, considerando o dialeto adotado. Foi também atribuída uma classificação morfossintática, conforme a pronúncia e significado dos HHs transcritos. Um exemplo da transcrição e atribuição morfossintática realizada pode ser observada nos pares de HHs contidos na Tabela 10 que segue:

Tabela 10. Pares de homógrafos heterófonos transcritos foneticamente.

<i>Palavra</i>	<i>Classe Gramatical</i>	<i>Transcrição Fonética (IPA)</i>
<abadessa>	Nome	[aba'desa]
<abadessa>	Verbo	[aba'desa]
<acarreto>	Nome	aka'xetɔ]
<acarreto>	Verbo	[aka'xetɔ]
<acerca>	Nome	[a'sexka]
<acerca>	Verbo	[a'sexka]
<acerto>	Nome	[a'sextɔ]
<acerto>	Verbo	[a'sextɔ]
<adereço>	Nome	[ade'resɔ]
<adereço>	Verbo	[ade'resɔ]

Por fim, com base no conhecimento linguístico produzido nesse primeiro módulo, foi criado um algoritmo responsável por verificar se a palavra submetida para transcrição no sistema PETRUS é, ou não, HH.

Essa verificação é feita pela busca da palavra na lista de HH compilada e transcrita para o sistema. Se a palavra estiver contida na referida lista, será fornecido ao usuário como *output* as duas pronúncias possíveis, associadas às suas respectivas classes gramaticais, sejam elas idênticas ou distintas. Assim, se o usuário do PETRUS digitar a palavra <gosto> e acionar o botão “transcrever” desse sistema, receberá como *output*:

**Figura 12.** Tela do PETRUS com a transcrição fonética da palavra <gosto>.

Posteriormente, nos inspiramos no trabalho de Silva et al. (2012) e organizamos tipologicamente o conjunto de HHs coletados nos dicionários supracitados, segundo a natureza de sua oposição gramatical e alternância vocálica das vogais médias presentes em sílaba tônica não marcada por acentuação gráfica.

Ao final, obtivemos um total de 27 tipos de HHs, dentre eles, os dois grupos que apresentaram maior quantidade de HHs foram escolhidos e implementados em forma de algoritmo em um classificador piloto para suporte à desambiguação de HHs em sistemas de síntese e reconhecimento de voz do PB, apresentado no STIL 2013.⁷¹ Para mais detalhes sobre o experimento realizado, vide Shulby et al. (2013).

2.1.2 Etapa 2 - Transcrição fonética de morfemas à esquerda da base

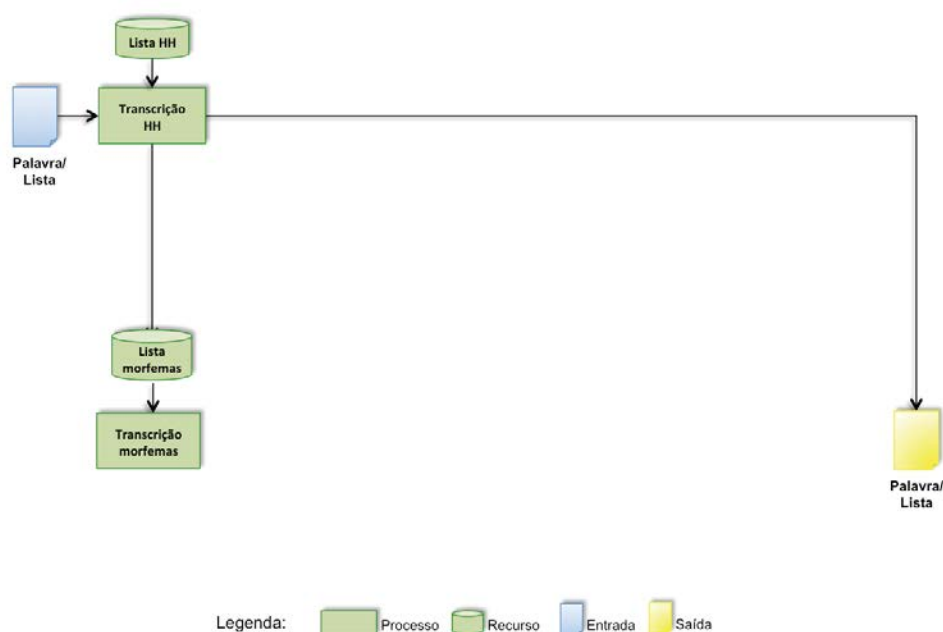


Figura 13. Segunda etapa de transcrição fonética no sistema PETRUS.

Por meio da Figura 13, ilustramos a segunda etapa no sistema PETRUS. Podemos observar que o segundo passo consiste em verificar se há a presença de um morfema à esquerda da base da palavra em análise.

⁷¹ STIL – IX Brazilian Symposium in Information and Human Language Technology. Anais disponíveis em: <http://www.nilc.icmc.usp.br/til/stil2013English/ProceedingsSTIL2013.pdf>.

A conversão das vogais médias em seus correlatos fonéticos está entre os principais problemas que envolvem transcrição fonética automática da língua portuguesa falada no Brasil. A dificuldade está justamente no fato de que, embora as vogais médias sejam pronunciadas de maneira distinta, dado seu contexto de ocorrência, são grafadas exatamente da mesma maneira, ou seja, nem sempre há um sinal gráfico (agudo ou circunflexo) que indique como devem ser pronunciadas. O mesmo se observa em relação à pronúncia dessas vogais em posição pretônica, dependendo da sua tonicidade e localização em relação à palavra-base.⁷² Em posição pretônica, as vogais médias podem ser constituintes de palavras prefixadas (ex.: *prefixo*) ou estarem presentes em elementos de composição gramatical (ex.: *etnosociologia*).

Frente a isso, decidimos criar uma etapa dentro do sistema PETRUS dedicada especificamente à transcrição fonética automática de morfemas localizados à esquerda da base e que contenham vogais médias em sua constituição. Isso porque, conforme afirma Oliveira (1996), a realização fonética de prefixos, principalmente os de origem grega, não segue as regras gerais utilizadas para a transcrição de palavras. Um exemplo que o autor cita é o caso do prefixo tele-. Em palavras como *telecomunicações*, *teleguiado*, etc., a primeira vogal média <e> do prefixo é pronunciada com som aberto [ɛ]; em contrapartida, em palavras frequentes como *telefone*, *televisão*, *telepatia*, etc., a inserção do prefixo acarretou o fechamento dessa primeira vogal, sendo pronunciada como [e]. Outro exemplo a ser citado é a não ocorrência do processo fonológico de neutralização da vogal pretônica. Uma das regras características do PB é a ocorrência de neutralização da pretônica (CÂMARA JR, 1970; WETZELS, 1991; 1992), que determina que o contraste entre as vogais médio-altas e médio-baixas não aconteça em posição não-acentuada. Essa regra, contudo, não se aplica às vogais médias dos seguintes prefixos:

- (a) n[ɛ]o + liberal → n[ɛ]o liberal *n[e]oliberal ;
- (b) pr[ɛ] + aviso → pr[ɛ] aviso *pr[e]aviso;
- (c) p[ɔ]s + graduado → p[ɔ]s graduado *p[o]sgraduado;
- (d) pr[ɔ]to + mártir → pr[ɔ]to mártir *pr[o]tomártir.

⁷² Como é o caso, por exemplo, do prefixo pré- que pode se realizar como vogal média aberta quando se localiza a esquerda da palavra e sem integrá-la (ex.: 'pr[ɛ]-natal') ou com vogal média fechada quando não recebe acento, integrando-se a palavra de base (ex.: pr[e]destinado).

Dessa forma, na presente seção, apresentamos um delineamento sobre o processo de obtenção da transcrição fonética de morfemas localizados à esquerda da base que apresentam vogais médias. O objetivo dessa etapa é, portanto, realizar a transcrição fonética de morfemas em início de palavras, constituídos por vogais médias, como proposta para minimizar os erros causados na transcrição das referidas vogais em posição pretônica.

Para executar tal tarefa, um algoritmo identificador compara a subcadeia inicial de cada palavra com uma lista de morfemas contida no PETRUS, transcrevendo-a, em caso positivo.

A Tabela 11 apresenta os morfemas localizados à esquerda da base tratados pelo sistema PETRUS, provenientes de uma recolha manual em gramáticas (Bechara (2009); Cunha e Cintra (2002); Rocha Lima (2011) e manuais *online* de ensino do PB, como o “Só Português” e o “Gramática Online”.⁷³ Conforme podemos observar na mesma tabela, apenas aqueles que apresentam vogais médias foram considerados.

Tabela 11. Morfemas localizados à esquerda da base constituídos pelas vogais médias <e> ou <o>, transcritos pelo PETRUS.

	<i>Morfemas à esquerda da base</i>	<i>Transcrição fonética</i>	<i>Exemplo</i>
1.	<i>acro-</i>	[akro]	acropatia
2.	<i>aero-</i>	[aero]	aerodinâmica
3.	<i>agro-</i>	[agro]	agronomia
4.	<i>andro-</i>	[ẽdro]	androgenia
5.	<i>ante-</i>	[ẽte]	antebraço
6.	<i>antropo-</i>	[ẽtropo]	antropologia
7.	<i>auto-</i>	[aũto]	autoadesivo
8.	<i>bem-</i>	[bẽĩ]	bem-aventurado
9.	<i>biblio-</i>	[biblio]	biblioterapia
10.	<i>bio-</i>	[bio]	biodiversidade
11.	<i>bronco-</i>	[brõõko]	broncodilatação
12.	<i>cardio-</i>	[kardʒio]	cardiorrespiratório
13.	<i>cloro-</i>	[klõro]	clorocarbonato
14.	<i>cripto-</i>	[kripito]	criptografia
15.	<i>des-</i>	[dʒis] ou [dʒiz]	despedaçar, desumano
16.	<i>eco-</i>	[eko]	ecologia ⁷⁴
17.	<i>etno-</i>	[etfino]	etnolinguista
18.	<i>eletro-</i>	[elẽtro]	eletrotécnico ⁷⁵

⁷³ “Só Português”: <<http://www.soportugues.com.br/secoes/morf/morf7.php>>; Gramática Online: <<http://www.gramaticaonline.com.br/Page.aspx?id=9&iddetalhe=172&idsubcat=44&idcateg=3>>.

⁷⁴ Exceção a essas regras são as palavras <ecochato> e <ecorrenewação>, cujos prefixos devem ter a vogal média transcrita como [ɛ].

⁷⁵ Exceto nas palavras <eletrodo>, <eletrônico> e seus derivados, nas quais o prefixo é transcrito como [eletro], com vogal média fechada.

19.	<i>endo-</i>	[ẽido]	endovenoso
20.	<i>entre-</i>	[ẽĩtre]	entrelinhas
21.	<i>epi-</i>	[epi]	epiderme
22.	<i>etno-</i>	[eʃino]	etnorracial
23.	<i>eu-</i>	[eʊ]	eufonia
24.	<i>es-</i>	[is] ou [ez]	esquentar, esôfago ⁷⁶
25.	<i>ex-</i>	[es]	ex-presidente
26.	<i>extra-</i>	[estra]	extraterrestre
27.	<i>ferro</i>	[fexʊ]	ferroelétrico
28.	<i>físio-</i>	[fizio]	fisioterapia
29.	<i>fito-</i>	[fito]	fitoterapia
30.	<i>foto-</i>	[fʊto]	fotomagnético ⁷⁷
31.	<i>gastro-</i>	[gastro]	gastroenterite
32.	<i>helio-</i>	[elio]	heliocêntrico
33.	<i>hemi-</i>	[emi]	hemisfério
34.	<i>hemo-</i>	[emo]	hemodiálise
35.	<i>hepta-</i>	[ɛpita]	heptacampeão
36.	<i>hetero-</i>	[ɛtero]	heterossexual
37.	<i>hidro-</i>	[idro]	hidrodinâmica
38.	<i>hiper-</i>	[ipex]	hipertensão
39.	<i>hipno-</i>	[ipino]	hipnoterapia
40.	<i>hipo-</i>	[ipo]	hipodérmico
41.	<i>homo-</i>	[omo]	homoerotismo
42.	<i>idio-</i>	[idʒio]	idiossincrático
43.	<i>iso-</i>	[izo]	isonomia
44.	<i>logo</i>	[logo]	logografia
45.	<i>macro-</i>	[makro]	macrobiótico
46.	<i>mega-</i>	[mɛga]	megalomania
47.	<i>meso-</i>	[mezo]	mesoderma
48.	<i>meta-</i>	[meta]	metalinguagem
49.	<i>metro-</i>	[metro]	metrossexual
50.	<i>micro-</i>	[mikro]	microempresa
51.	<i>mio-</i>	[mio]	miopatia
52.	<i>mono-</i>	[mono]	monossílabo
53.	<i>morfo-</i>	[morfo]	morfossintaxe
54.	<i>moto-</i>	[mʊto]	motoniveladora
55.	<i>necro-</i>	[nekro]	necrofobia
56.	<i>neo-</i>	[neo]	neoclássico
57.	<i>odonto-</i>	[odõõto]	odontogenia
58.	<i>oftalmo-</i>	[ofitaʊmo]	oftalmografia
59.	<i>oni-</i>	[oni]	onipresente
60.	<i>onto-</i>	[õõto]	ontogênese
61.	<i>orto-</i>	[orto]	ortopedia
62.	<i>oto-</i>	[oto]	ototóxico
63.	<i>pato-</i>	[pato]	patologia
64.	<i>per-</i>	[pex]	percorrer

⁷⁶ Se o prefixo *es-* é sucedido por consoante, então é transcrito como [is], caso contrário, transcrito como [ez].

⁷⁷ Exceto na palavra <fotografar> e suas variações.

65.	<i>peri-</i>	[peri]	pericárdio
66.	<i>piro-</i>	[piro]	piromania
67.	<i>pneumo-</i>	[pneomo]	pneumoterapia
68.	<i>poli-</i>	[poli]	politécnico
69.	<i>pos-</i>	[pos]	postônico
70.	<i>pseudo-</i>	[piseodo]	pseudocirrose
71.	<i>psico-</i>	[psiko]	psicopatologia
72.	<i>re-</i>	[xe]	recomeçar
73.	<i>recém-</i>	[xesẽĩ]	recém-casado
74.	<i>retro-</i>	[xetro]	retroalimentação
75.	<i>rino-</i>	[xino]	rinoceronte
76.	<i>semi-</i>	[semi]	semiconsoante
77.	<i>sob-</i>	[sobi]	sobdominante
78.	<i>sobre-</i>	[sobre]	sobrecarga
79.	<i>socio-</i>	[sosiu]	sociopolítico
80.	<i>soto-</i>	[soto]	sotoposto
81.	<i>sub-</i>	[subi]	subproduto
82.	<i>super-</i>	[supex]	supermercado
83.	<i>tecno-</i>	[tekno]	tecnoestrutura
84.	<i>tele-</i>	[tele]	teleconferência ⁷⁸
85.	<i>teo-</i>	[teo]	teologia
86.	<i>termo-</i>	[teymo]	termomagnético
87.	<i>topo-</i>	[topo]	topofobia
88.	<i>trans-</i>	[trẽs] ou [trẽz]	transamazonense, transcodificar
89.	<i>tres-</i>	[trez] ou [tres]	tresentender, trespassar
90.	<i>vice-</i>	[vise]	vice-presidente
91.	<i>video-</i>	[video]	videoconferência
92.	<i>xeno-</i>	[feno]	xenofobia
93.	<i>zoo-</i>	[zoo]	zoologia

Posteriormente, formalizamos, por meio de regras de transcrição, a realização fonética de cada um dos morfemas contidos na Tabela 11, considerando a variedade linguística adotada neste estudo. A seguir, será apresentado o módulo do PETRUS responsável por atribuir a classe gramatical às palavras de entrada.

2.1.3 Etapa 3 - Identificação automática de categorias gramaticais

Este módulo tem como objetivo identificar as categorias gramaticais das palavras submetidas ao sistema PETRUS para análise, possibilitando ao usuário de nosso sistema

⁷⁸ Exceto a palavra <telefone> e seus derivados.

incorporar de forma automática informação gramatical ao verbete de sua obra lexicográfica. Futuramente, essa informação será interessante para o próprio sistema de transcrição, uma vez que será a partir da classificação gramatical que o PETRUS poderá atribuir a pronúncia das vogais médias tônicas em nomes e verbos que não caracterizam casos de HH.

Vejam a ilustração da etapa de atribuição automática de classe gramatical na figura que segue:

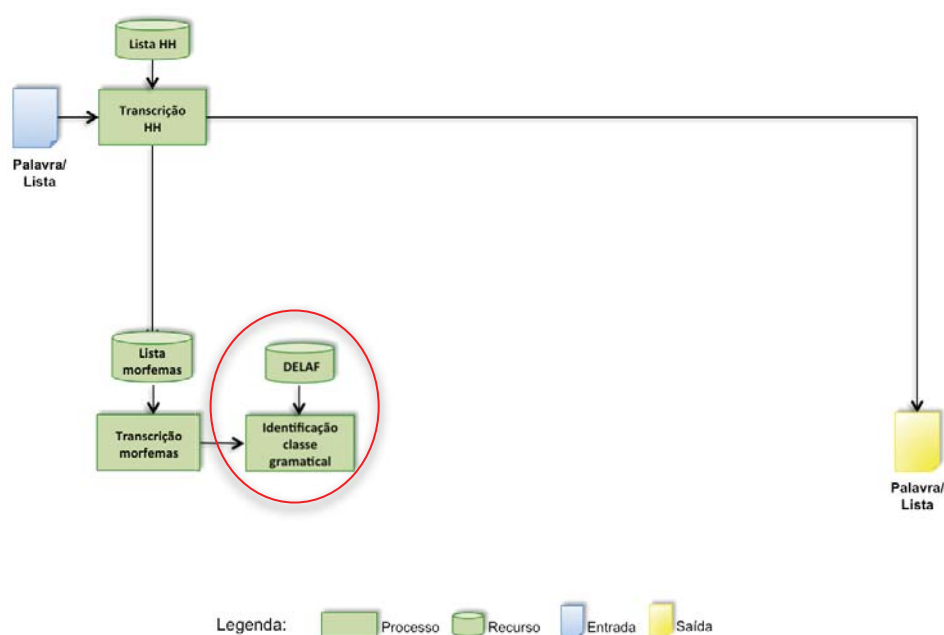


Figura 14. Atribuição da classe gramatical às palavras de entrada do sistema PETRUS.

Conforme se pode observar na Figura 14, o léxico computacional DELAF composto por 880.000 palavras simples flexionadas, produzido por Muniz (2004), é o único recurso linguístico necessário ao desenvolvimento deste módulo atualmente.

O DELAF é uma versão em português do Brasil do DELA (Dictionnaire Electronique du LADL⁷⁹), um padrão de dicionários eletrônicos (também conhecido como léxicos computacionais) que é utilizado pela ferramenta de análise de córpus, o INTEX, e mais atualmente pela ferramenta de código fonte aberto, o Unitex.⁸⁰

O Unitex é um ambiente de desenvolvimento linguístico que inclui dicionários e gramáticas de grande cobertura para várias línguas, tais como o Espanhol, o Inglês, o Francês,

⁷⁹ LADL – Laboratoire d’Automatique Documentaire et Linguistique, localizado na França.

⁸⁰ Unitex. Disponível em: <http://www-igm.univ-mlv.fr/~unitex/>.

o Grego, o Russo, o Português Europeu e o Tailandês e processa textos com milhões de palavras em tempo real. Ele inclui ferramentas para criar e manter recursos lexicais, criar e manter gramáticas de remoção de ambigüidades, e esses dicionários e gramáticas podem ser aplicados aos textos para localizar padrões morfológicos, lexicais e sintáticos, remover ambigüidades, e etiquetar palavras simples e compostas. Portanto, o objetivo do trabalho de Muniz (2004) foi justamente construir léxicos computacionais para o PB baseado no formalismo DELA.

De acordo com Muniz (2004, p. 5), o léxico computacional, ou dicionário, é uma estrutura de dados composta por unidades lexicais de uma língua bem como por informações correspondentes a esses itens. Em relação às unidades lexicais, essas podem ser tanto palavras isoladas (ex.: casa, mesa) quanto agrupamento de palavras com significado específico (ex.: vez ou outra, mão de vaca). Quanto ao tipo de informações às quais podem ser associadas, o autor destaca as referentes a categoria gramatical (*part-of-speech*) do item, além de valores com variáveis morfossintáticas, tais como, gênero, número, grau, pessoa, tempo, modo, regência verbal ou nominal, entre outras. O mesmo autor acrescenta ainda que descrições ou representações semânticas também podem ser associadas à UL, bem como representações contextuais, apesar de serem mais raras.

Ainda de acordo com Muniz (2004, p. 17), as formas flexionadas de todas as entradas simples constituem o dicionário DELAF, gerado automaticamente a partir do dicionário DELAS.⁸¹ Dessa forma, em relação às entradas do DELAF, estas são constituídas da UL, seu lema, a sua categoria gramatical e a flexão que corresponde a essa forma (por exemplo: as flexões das entradas mato, N001D026A01 e beijar,V005 no DELAF são apresentadas como beijar,beijar.V:W1s:W3s:U1s:U3s e mato,mato.N:ms, respectivamente.

Como vimos, o dicionário DELAF é formado essencialmente por palavras às quais estão associadas informações gramaticais e é utilizado na atribuição da classe gramatical de palavras submetidas ao PETRUS para transcrição. Assim, implementou-se um algoritmo que

⁸¹ As entradas do dicionário DELAS possuem a estrutura <palavra>, <descrição formal>, onde *palavra* representa a forma canônica (o lema) de uma UL simples. A vírgula separa o lema do código de flexão, como podemos observar nos exemplos retirados do dicionários DELAS produzido para o PB: mato, N001D026A01 - beijar, V005. A primeira informação sobre o item lexical é a indicação da categoria gramatical a que pertence. O código numérico indica o modelo de flexão da categoria gramatical em questão. Os códigos D e A, seguidos de um código numérico adicional, permitem gerar, respectivamente, o(s) diminutivo(s) e o(s) aumentativo(s) adequado(s). Vale lembrar que, no caso de uma palavra se enquadrar em mais de uma categoria gramatical, para cada categoria, deverá haver uma entrada no DELAS, por exemplo, a entrada flexionada mato pode ser gerada pela seguintes entradas no DELAS: matar,V030 e mato,N001D026A01. (MUNIZ, 2004, p. 17-18).

possibilita que, dada uma palavra de entrada, o sistema PETRUS pode identificar rapidamente sua classe gramatical, fazendo uma busca rápida no dicionário citado.

A Figura 15 mostra as classes gramaticais que a palavra <casa> pode assumir em nosso idioma.

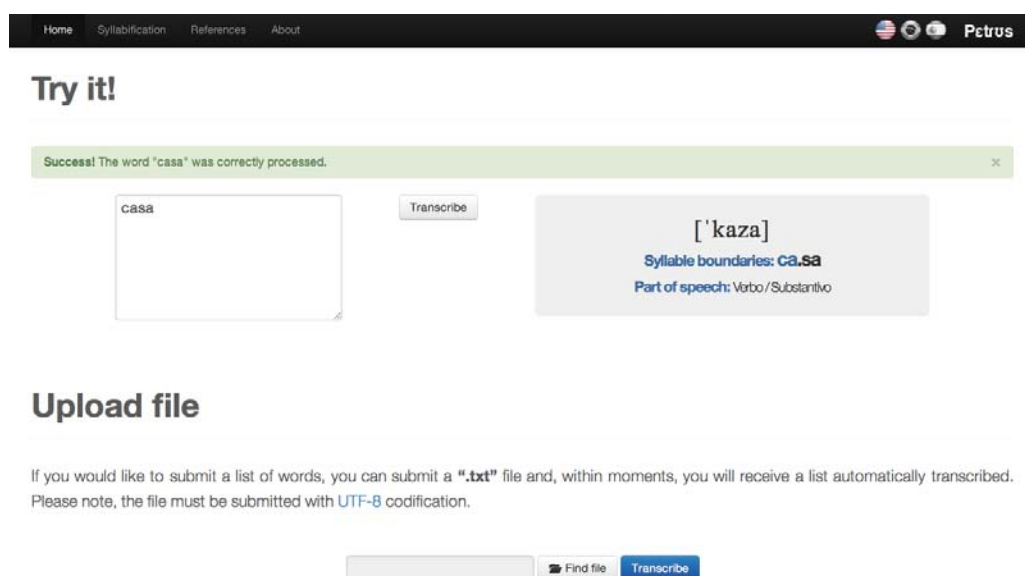


Figura 15. Tela do PETRUS com a análise da classe gramatical da palavra <casa>.

Podemos observar que <casa> pode desempenhar tanto o papel de nome, como em “A *casa* está bela”, quanto de verbo, como por exemplo em “Ele *casa* neste mês”. Por se tratar de UL simples, fora de contexto, as classes gramaticais possíveis foram extraídas a partir do dicionário implementado, o DELAF.

Entretanto, para o caso de uma sentença ser apresentada como entrada no sistema PETRUS, será utilizado o módulo *Part-of-Speech (POS) tagger* do sistema “Natural Language Processing with neural networks” (*nlpnet*), elaborado por Fonseca e Rosa (2013), para atribuir as classes gramaticais, já que o contexto de ocorrência das palavras pode ser considerado para a realização dessa tarefa.⁸² Essa função será habilitada quando, futuramente, unidades maiores que uma UL poderão ser tratadas pelo referido sistema.

⁸² O *nlpnet* é uma biblioteca Python para tarefas que envolvem o processamento de linguagem natural baseada em redes neurais. Atualmente, ela realiza classificação morfossintática (*part-of-speech tagging*) e anotação semântica. Para mais informações sobre o *nlpnet* vide: <http://www.nilc.icmc.usp.br/nlpnet/#module-nlpnet>.

2.1.4 Etapa 4 - Identificação automática da sílaba tônica

Na transcrição fonética promovida pelo sistema PETRUS, são marcadas apenas as vogais tônicas, ou seja, apenas o acento primário é identificado. Como os acentos secundários não são marcados, nos casos de ocorrência de palavras compostas, só a última receberá o acento, como acontece por exemplo em <vale-refeição> [valixeferi'sẽõ] e <guarda-chuva> [goaxda'juva]. Da mesma maneira, os diminutivos sufixados <menininho, cafezinho>, advérbios de modo <fortemente, alegremente>, compostos <afro-brasileiro, surdo-mudo> recebem apenas a marcação de uma sílaba tônica. Em caso de ocorrência de mais de uma palavra, o usuário é advertido para que seja submetida apenas uma por análise. Neste caso, uma mensagem de erro aparecerá advertindo que se deve inserir apenas uma UL simples para análise. No que diz respeito aos clíticos (palavras átonas do PB, como as preposições, os artigos, os pronomes oblíquos, etc.), não recebem a marca de acento.

Conforme ilustrado na Figura 16, temos a realização da marcação automática da vogal tônica em palavras submetidas para análise no PETRUS, a partir da aplicação de um conjunto de regras propostas por Silva (2011).

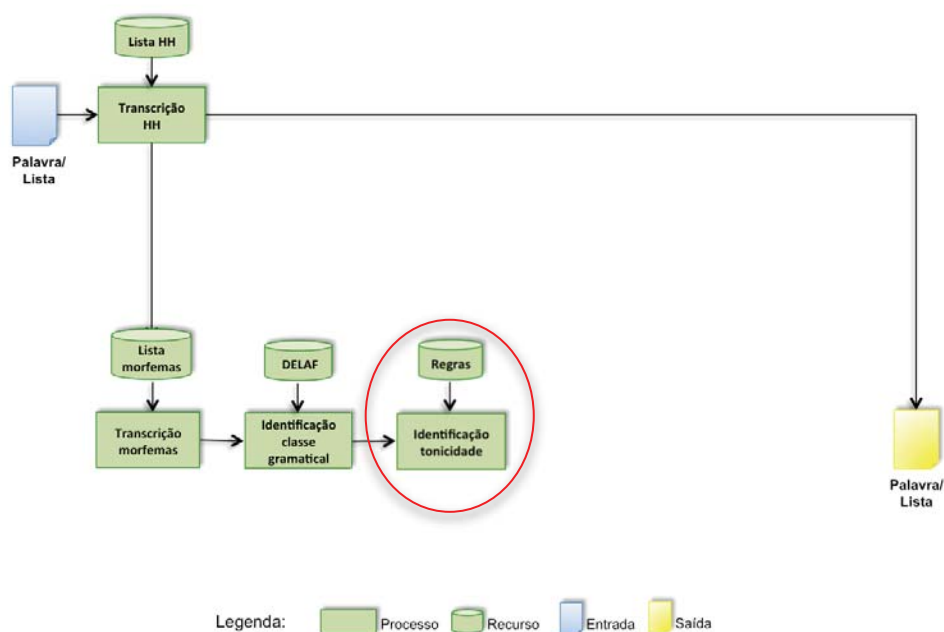


Figura 16. Realização da marcação automática da vogal tônica em palavras submetidas ao PETRUS.

A indicação da sílaba tônica é uma informação interessante para ser incluída em obras lexicográficas do PB, para o aprendizado de pronúncia. No *Dicionário online Caldas Aulete*,⁸³ por exemplo, apesar de não apresentar transcrição fonética de suas entradas, a partir da qual haveria a indicação da sílaba tônica, há, no entanto, o recurso ao uso de uma grafia itálica da sílaba tônica, conforme podemos observar na tela abaixo, reproduzida do referido dicionário:

The image shows a screenshot of the Caldas Aulete online dictionary. At the top, there are navigation tabs: 'Sua língua na Internet', 'Dicionário Aulete', 'Gramática básica', and 'Dicionário analógico'. Below these are social media icons for Facebook, Twitter, and Google+, along with a 'Compartilhar' button. A secondary navigation bar contains links for 'Página principal', 'O que é', 'Palavra do dia', 'Downloads', and 'Convide um amigo'. The main content area features the 'Aulete DIGITAL' logo and a search input field. Below the search bar, there are buttons for 'Verbetes Atualizado' and 'Verbetes Original'. The word 'caricatura' is displayed in a large font, with its syllable '(ca.ri.ca.tu.ra)' circled in red. To the right of the word are three 'A' icons. Below the word, there are three numbered definitions: 1. 'Desenho que representa uma pessoa ou um acontecimento com traços deformados, muitas vezes exagerados, revelando com isso aspectos característicos, grotescos ou ridículos do que é retratado.' 2. 'Fig. Pessoa de aparência e/ou modos grotescos ou ridículos.' 3. 'Exemplo malsucedido do que (algo ou alguém) pretende ou deveria ser: *Ele é uma*'. To the right of the main content, there is a 'Lexikon' section titled 'obras de referência' with the sub-heading 'Palavras análogas'. This section contains a mind map diagram with 'caricatura' at the center, connected to five related terms: 'Ridicularia', 'Ridiculização', 'Cópia', 'Exagero', and 'Arremedo'.

Figura 17. Tela do dicionário Caldas Aulete com indicação da sílaba tônica em itálico.

A indicação da sílaba tônica tem impacto direto também no módulo que realiza a conversão grafema-fone do sistema PETRUS, na medida em que algumas regras para conversão utilizam a informação da tonicidade. Exemplos a esse respeito podem ser encontrados na Tabela 13.

Dentre os sistemas de marcação de sílabas tônicas desenvolvidos para o português, temos para o PE, Oliveira et al. (1991) que utilizam 18 regras previamente descritas em Viana e Andrade (1985), referindo-se à importância desse módulo de marcação da tônica para a então versão do sistema DIXI.⁸⁴ Ainda para o PE, Teixeira et al. (1998) descrevem um algoritmo de marcação de sílaba tônica composto por apenas três regras, mais uma tabela de exceções; contudo não é mencionado o desempenho do sistema nem as tabelas de exceções foram publicadas. Mais recentemente, Barros e Weiss (2006) utilizaram uma abordagem estatística para derivar as regras de marcação de tônicas, adotando para isso um modelo por

⁸³ *Dicionário online Caldas Aulete*, disponível em: <http://www.aulete.com.br/>.

⁸⁴ DIXI: O primeiro sistema de síntese de fala a partir de texto desenvolvido para a língua portuguesa. Para mais informações, consultar Oliveira (1996).

máxima entropia e um córpus composto por 4.219 palavras com sílaba tônica identificadas. Segundo publicado pelos autores, esse método estatístico apresentou uma taxa de acerto de 85,57%. Posteriormente, Veiga et al. (2011; 2013), seguindo os pressupostos teóricos apresentados por Mateus e D'Andrade (2000), propõem um conjunto de 6 regras para marcação da vogal tônica em palavras. Entretanto, a taxa de acerto não é mencionada.

Em relação ao PB, Chbane (1994) destaca a importância da determinação da sílaba tônica para o estudo da prosódia e para a formulação de modelos que controlem a variação dos parâmetros prosódicos a fim de imprimir à fala sintetizada uma maior naturalidade. Para tanto, o autor utilizou em seu trabalho 6 regras para atribuição da sílaba tônica e como teste aplicou-as em 200 palavras previamente selecionadas. O seu desenvolvimento sofreu refinamentos sucessivos até que produzisse 100% de acerto na determinação da sílaba tônica das palavras que não apresentam hiato. Mais recentemente, Silva et al. (2006) apresentaram um algoritmo composto por 20 regras para determinação da tonicidade, com uma taxa de acerto de 98.53%, em 4.825 palavras recolhidas aleatoriamente do córpus CETEN-Folha. Posteriormente, com base no trabalho de Silva et al. (2006), Siravenha (2009) apresenta um conjunto de 29 expressões regulares para determinar a vogal tônica. A autora afirma que os erros mais observados referem-se às palavras estrangeiras e nomes próprios, mas não cita valores numéricos para as taxas de acerto e erro obtidas com as expressões. Mais recentemente, Vasilévski e Araújo (2013) apresentam uma amostra do uso de um algoritmo que identifica a posição tônica de palavras, considerando critérios fonológicos, ou seja, partindo da transcrição fonológica de palavras do PB. Entretanto, apenas um pequeno trecho do algoritmo é publicado. Em relação ao seu desempenho, não é feita nenhuma consideração, apresenta-se, como teste, a aplicação do algoritmo para estudo prosódico da música “Construção”, de Chico Buarque.

Para a primeira versão do sistema PETRUS aqui descrita, decidimos implementar as regras publicadas por Silva et al. (2006) devido à completa documentação e disponibilização dos algoritmos desenvolvidos, que promoveram uma rápida implementação ao nosso sistema. Após serem implementadas e testadas em um córpus de estudo com 52.525 palavras com vogal tônica marcada,⁸⁵ constatamos uma taxa de acerto de 93%.

Vale mencionar também, que as regras implementadas são específicas para atribuição de tonicidade em palavras simples do PB.

⁸⁵ Córpus de teste construído com base no Dicionário online Caldas Aulete: <http://www.aulete.com.br>.

2.1.5 Etapa 5 – Divisão silábica automática

A Figura 18 ilustra o módulo responsável por realizar divisão silábica de forma automática, que, como podemos observar, foi habilitado no sistema PETRUS através dos algoritmos para silabificação de palavras do PB, também desenvolvidos e disponibilizados por Silva (2011).

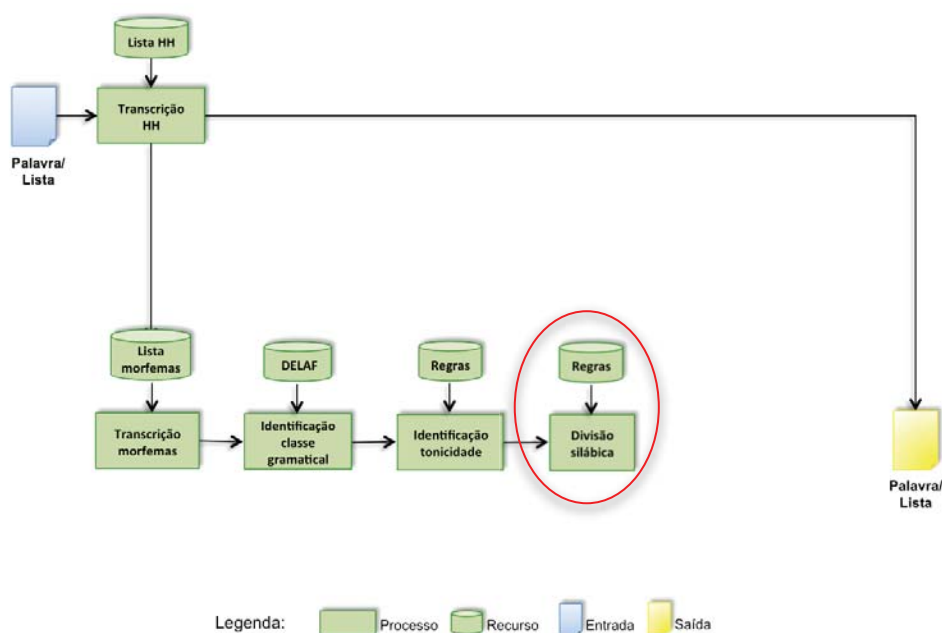


Figura 18. Divisão silábica automática das palavras de entrada do PETRUS.

A divisão silábica consiste na identificação e delimitação das sílabas que constituem cada UL. Como vimos na seção 1.1.4.1, o constituinte obrigatório da sílaba é uma vogal que designa seu núcleo, a qual pode vir, ou não, acompanhada por outra vogal, foneticamente denominada de glide, formando um ditongo, como ocorre na palavra <lei.te>. Esse mesmo núcleo pode ser precedido por uma ou mais consoantes, como em <ca.sa> e <pra.to>, bem como ser sucedido por uma ou mais consoantes, como <ár.vo.re> e <pers.pec.ti.va>. No entanto, a divisão silábica de palavras com determinadas sequências de vogais e consoantes nem sempre é tão intuitiva: o correto seria <saudar> (sa.u.dar) ou (sau.dar)? Segundo o dicionário Aulete digital,⁸⁶ o segundo exemplo (sau.dar) é apontado como sendo correto. Já o

⁸⁶ Dicionário Aulete Digital. Disponível em: <<http://www.aulete.com.br>>. Acesso em: 24 nov 2014.

Portal da Língua Portuguesa⁸⁷ adverte que *saudar* utilizado como verbo deve ser silabificado como <sa.u.dar> e, quando utilizado como nome masculino, deve ser silabificado como <sau.dar>. Um outro exemplo: a divisão silábica está correta em <obstruir> (obs.tru.ir) ou (o.bs.tru.ir)? Tanto no *Dicionário online Caldas Aulete*⁸⁸ como no *Portal da Língua Portuguesa*,⁸⁹ o primeiro exemplo (obs.tru.ir) é apresentado ao usuário.

De qualquer maneira, independentemente de que critérios considerar para a sua definição, dos pontos de vista teóricos sobre seu *status* linguístico, e da dificuldade de se realizar automaticamente a divisão silábica em palavras, a sílaba tem tido seu papel no desenvolvimento de diferentes práticas que envolvem PLN e tecnologias de fala (PEARSON et al., 2000).

O desenvolvimento de um reconhecedor de voz frequentemente envolve o trabalho com unidades menores do que uma palavra, tais como as sílabas (GREENBERG, 1999); mais ainda, a silabificação é uma parte essencial de sistemas de síntese de fala. De acordo com Sampaio Neto (2011, p. 26), é essencial, em qualquer língua para a qual se queira construir um sistema de síntese de voz, realizar mapeamento das sílabas, uma vez que essa informação é de grande monta tanto para a transcrição fonética, quanto para a geração de prosódia, já que em uma sílaba tônica há aumento de duração e intensidade. A respeito do papel da divisão silábica em sistemas que realizam a conversão grafema-fone de forma automática, Oliveira (2009) afirma:

À semelhança do que acontece para outras línguas - e.g. inglês (MARCHAND; DAMPER, 2007; BARTLETT et alii, 2008), francês (BERINGER, 2004) e alemão (LIBOSSEK; SCHIEL, 2000) - também no caso do PE, o desempenho dos sistemas de conversão grafema-fone(ma) aumenta significativamente com a integração de informação silábica (OLIVEIRA, 2009, p. 105).

Portanto, do ponto de vista da engenharia, há boas razões práticas que motivam a busca por algoritmos cada vez mais poderosos para marcar automaticamente as fronteiras silábicas.

No âmbito da aplicação ao sistema PETRUS, a informação silábica tem auxiliado na implementação de determinadas regras de conversão grafema-fone, que foram construídas considerando esse tipo de informação, tais como as relacionadas com a geração de ditongos.

⁸⁷ Portal da Língua Portuguesa. Disponível em: <<http://www.portaldalinguaportuguesa.org/about.html?action=syllables&act=list&search=saudar>>. Acesso em: 24 nov. 2014.

⁸⁸ Dicionário online Caldas Aulete: <http://www.aulete.com.br>.

⁸⁹ Portal da Língua Portuguesa: <http://www.portaldalinguaportuguesa.org>.

Em relação ao usuário do sistema PETRUS, acreditamos que disponibilizar uma funcionalidade como essa seja de grande valia, sendo esse usuário lexicógrafo ou não. Assim, para os lexicógrafos, esse tipo de informação poderá ser disponibilizada nas obras lexicográficas que produzem, sem que para isso tenham que realizar essa tarefa de forma manual, em todos os verbetes. Para linguistas, esse recurso pode ser aproveitado, por exemplo, em estudos sobre prosódia. Para aprendizes de português, um silabificador poderá auxiliá-los quanto à identificação e pronúncia de ditongos e hiatos no PB, uma questão problemática de nossa língua para estrangeiros. Até mesmo falantes nativos do português apresentam dificuldades de reconhecer se há hiatos ou ditongos em palavras não conhecidas pelo falante, bem como em palavras já conhecidas, como “ruim”, “gratuito”, “circuito”, por exemplo.

Entre os principais silabificadores existentes para o PE, podemos citar os de Gouveia et al. (2000), Oliveira et al. (2007) e Braga e Resende Jr (2007), enquanto que para o PB, temos os de Madureira et al. (1999), Silva et al. (2008), Silva e Oliveira (2012), Silva (2011) e Rocha e Neto (2013), que é uma versão contendo as 20 regras de silabificação propostas por Silva et al. (2008) mais duas regras para tratamento de encontros vocálicos, já descritas em Monte et al. (2011). Todos foram desenvolvidos com vistas à aplicação em sistemas de síntese de voz para o português.

Partindo da revisão bibliográfica sobre silabificadores já produzidos para o PB, cujas regras e algoritmos estejam disponíveis livremente para uso, realizamos testes para verificar a precisão de dois deles: Silva (2011) e Rocha; Neto (2013). Para isso, foi construído um corpus de teste, composto por 19.056 palavras, cuja silabificação realizamos manualmente. Os resultados obtidos na comparação são apresentados na tabela seguinte.

Tabela 12. Comparação do desempenho de sistemas de divisão silábica automática para o PB

	<i>Silva (2011)</i>	<i>Rocha; Neto (2013)</i>
<i>Palavras corretamente silabificadas</i>	18.041	16.781
<i>Palavras não silabificadas corretamente</i>	1.015	2.264
<i>Taxa de acerto</i>	94.67%	88.11%

Diante dos resultados obtidos nos testes realizados, adotamos para o sistema PETRUS o conjunto composto por 25 regras de Silva (2011), devido à completa documentação e disponibilização dos algoritmos, que promoveram uma rápida

implementação ao nosso sistema. Tais regras foram construídas com base, primeiramente, na busca de vogais na palavra em análise, seguida pela análise dos caracteres adjacentes a vogal para só então o algoritmo decidir qual a ação que deve ser realizada. A decisão a ser tomada envolve a união, ou não, da vogal em análise aos grafemas a ela adjacentes. Assim, uma nova sílaba é formada, caso os grafemas possuam uma vogal anexada. Das 25 regras, as seis primeiras tratam da questão das vogais em início de sílaba, enquanto que as dezenove restantes são dedicadas ao tratamento das vogais localizadas em posição interna ou final.

Existem diferenças entre o conjunto de regras adotados para uma divisão silábica para efeitos de translineação⁹⁰ e uma divisão silábica feita com base fonológica. Os algoritmos de silabificação propostos por Silva (2011, p. 39) foram desenvolvidos “com base na ortografia do PB, mas prevalecendo a fonologia das palavras, na tentativa de conciliar as teorias fonológicas da língua com as necessidades dos sistemas TTS” (de síntese de fala). Com isso, as palavras <obstrução>, <arredondar> e <assado> são silabificadas como (o.bs.tru.ção), (a.rre.don.dar) (a.ssa.do). Todavia, se considerássemos as normas ortográficas, os critérios para a divisão também seriam modificados; portanto, pelos critérios ortográficos, só poderíamos separar as sílabas dessas palavras da seguinte forma: (obs.tru.ção), (ar.re.don.dar) (as.sa.do). Como desejamos uma análise fonético-fonológica das palavras para realizar a divisão silábica das palavras de entrada do sistema PETRUS, informação essa que é importante para o próximo módulo deste sistema, o de conversão grafia-fone, o silabificador adotado (desenvolvido por Silva (2011)) atende aos nossos requisitos.

Uma questão que merece nota são os encontros vocálicos e a sua caracterização na divisão silábica enquanto ditongos (vogais pronunciadas em uma mesma sílaba). Ditongos crescentes, como <náusea>, <ciência> e <etéreo> são problemáticos, uma vez que a pronúncia que se imprime à glide é indecisa e variável. Trata-se de um fenômeno resultante da dupla possibilidade de se articular a sequência vocálica tanto como ditongo crescente (náu.sea, ci.ên.cia, e.té.reo) quanto como hiato (náu.se.a, ci.ên.ci.a, e.té.re.o), conforme a rapidez (caso dos ditongos) ou lentidão (caso dos hiatos) com que são pronunciadas.

Segundo Bechara (2011), as pronúncias ditongadas como (náu.sea, ci.ên.cia, e.té.reo) são mais comuns no falar rápido dos portugueses atualmente; entre brasileiros, o ditongo é menos usado, em favor do hiato. Justifica essa preferência mediante o fato de falarmos mais claro e pausadamente, recordando a definição de Eça de Queirós sobre o PB ser “o português

⁹⁰ Translineação: mudança, na escrita, de uma linha para outra, ficando parte da palavra no final da linha superior e parte no início da linha inferior.

com açúcar”.

De acordo com Bisol (1989), os ditongos verdadeiros são os decrescentes (vogal + glide), pois os crescentes (glide + vogal) variam com o hiato. A autora complementa sua afirmação ao dizer que os ditongos decrescentes são ditongos fonológicos, ou seja, aparecem tanto na pronúncia quanto na escrita. Em geral, possuem pares mínimos com as vogais simples e atraem o acento lexical das formas morfológicamente simples. Assim, são ditongos as sequências vocálicas que aparecem, por exemplo, em <lei> e <laudo>, que se opõem às vogais simples das palavras <lê> e <lado>. Já em relação aos falsos ditongos (crescentes), a autora afirma se tratar de encontros vocálicos nos quais a glide pode ser apagada, como em <peixe> [ˈpeʃi] ~ [ˈpeʃɪ]⁹¹.

Por essas razões, foi adotado o procedimento de determinar apenas os ditongos decrescentes, enquanto que os ditongos crescentes são todos interpretados como hiatos. Dessa maneira, os encontros vocálicos terminados em <i> ou <u>, como em (bai.xo), (lei.te), (pa.péis), (foi.ce), (in.troi.to), (cir.cui.to), (cau.sa), (deu.sa), (mel.ro), (fil.tro), (sol.to), (vol.to), (cul.to), são sempre convertidos como ditongos decrescentes. Neste caso, as vogais <i> e <u> são convertidas para as semivogais [ɪ] e [ʊ], respectivamente. Há, no entanto, apenas dois casos em que esses encontros vocálicos são tratados como hiatos: quando as vogais <i> e <u> são sucedidas pela pelo dígrafo <nh> ou pela sequência <ir>. Por exemplo, (ba.i.nha), (gra.u.nha) e (cons.tru.ir), (sa.ir).⁹²

Neste estágio do sistema PETRUS, tem-se acesso a diferentes níveis de informação sobre a palavra em análise, desde a presença ou não de um prefixo até a categorização gramatical, a identificação da vogal tônica e da divisão silábica. Assim sendo, foi possível implementar regras para transcrição fonética automática com maior complexidade, incorporando tais tipos de informação no desenvolvimento dos algoritmos implementados.

Na Tabela 13 são apresentadas algumas regras de transcrição do PETRUS que consideram, para sua execução, as informações produzidas pelos módulos de identificação de vogal tônica e de divisão silábica.

⁹¹ O ditongo <ei> também é realizado como <e> diante de [r] como nas palavras isqu[e]ro, ch[e]ro e de [ʒ], como nas palavras qu[e]jo e b[e]jo.

⁹² A questão da interpretação da sequência glide + vogal como ditongo crescente ou hiato é uma questão ainda controversa na literatura. Por isso, apoiamos-nos em Bisol (1989) e decidimos adotar para este trabalho a postura de considerar apenas a identificação dos ditongos decrescentes para a realização da divisão silábica, ou seja, considerar a sequência glide + vogal como sendo a realização de um hiato. Contudo, não consideramos inadequada a interpretação dessa sequência como um ditongo crescente. Dependendo do critério utilizado, uma ou outra interpretação pode ser mais adequada. Nesse contexto, pensando no possível perfil do usuário do sistema PETRUS, temos como trabalho futuro, a opção de produzir, neste caso, um ou outro resultado: divisão silábica automática com ditongos crescentes ou hiatos. Assim, o sistema ganharia em flexibilidade e o usuário poderia optar pela forma que fosse mais adequada aos seus interesses e o sistema ganha flexibilidade.

Tabela 13. Regras de transcrição fonética que consideram informação de tonicidade e silábica

<i>Regras</i>	<i>Palavras (grafia e transcrição fonética)</i>
Redução das vogais átonas finais <e>, <o> para [ɨ], [ʊ].	<pa.ta> ['pata] <go.le> ['gɔli] <sa.po> ['sapʊ]
Redução de [e] pretônico para [ɨ] na primeira sílaba da palavra, com [e] sucedido por [s] ou [z] na sílaba seguinte.	<es.co.la> [ɨs'kɔla] <e.xa.me> [ɨ'zɛmi]
Redução de [e] para [ɨ] em sílaba pretônica inicial, quando [e] é precedido por [d] e seguido por [s] ou [z].	<des.cre.ve> [dʃɨs'krɛvi]
Nasalização da vogal tónica <a> quando a consoante da sílaba seguinte é nasal.	<ta.ma.nho> [ta'mɛ̃ɲʊ] <ca.na> ['kɛ̃na]

A seguir, apresentamos o último módulo que compõe o sistema PETRUS. Como poderá ser observado, a partir das informações geradas pelos cinco módulos até então descritos, associadas a um conjunto de regras linguísticas, o sistema ora proposto promove a conversão automática de letras em fones.

2.1.6 Etapa 6 - Conversão grafema-fone

Apesar de o português se situar entre as línguas cuja ortografia se mostra razoavelmente fonêmica, realizar transcrição fonética de forma automática não é tão simples quanto possa parecer. Isso porque, conforme afirmam Braga et al. (2003), apesar de a escrita da língua portuguesa ter natureza alfabética, isto é, baseada numa relação unívoca entre um som e um grafema, na verdade, essa relação não se verifica atualmente, devido à tensão existente entre o dinamismo da evolução fonética da língua e o carácter conservador próprio de todos os sistemas de escrita. Surgindo, conseqüentemente, dificuldades de conversão de alguns grafemas, e em especial dentro do vocalismo do português, visto que, para a mesma vogal do alfabeto ortográfico podem corresponder diferentes vogais do alfabeto fonético, segundo a sua posição na palavra, tonicidade e fones adjacentes.

Veja, por exemplo, o caso dos caracteres <e> e <o>, que podem apresentar pronúncia aberta ('f[ɛ]sta', 'c[ɔ]sta'), fechada ('ch[e]go', 'f[o]go') ou ainda serem pronunciados como <u> e <i>, respectivamente, como em 'tant[ʊ]' e 'tent[ɨ]'. Notório também é o caso do caractere <x>, que possui diferentes sons conforme seu contexto de ocorrência (<deixar> ~

[dei{fax}], <taxi> ~ [taksi], <exame> ~ [izẽmi], <excluir> ~ [iskluix], e entre vogais ou seguido de consoante tem pronúncia arbitrariamente variável, como em <baixo> ~ [bai{fo}], <axila> ~ [aksila], <máximo> ~ [masimo] e <êxito> ~ [ezito]). Há certas palavras do PB para as quais é ainda mais difícil deduzir a pronúncia adequada da vogal média tônica. Considere, por exemplo, as sentenças: “Como era a vida na *corte* francesa?” e “Quer que eu *corte* o seu cabelo?”. Para descobrir que o substantivo <cor-te> da primeira frase deve ser pronunciado como c[o]rte e que na segunda sentença o verbo deve ser pronunciado como c[ɔ]rte, é necessário realizar uma análise morfossintática, pois só então é possível identificar a classe gramatical e, posteriormente, a pronúncia adequada. Tais palavras constituem casos de homógrafos heterófonos, palavras que possuem mesma grafia, mas são pronunciadas de forma diferente. O módulo do sistema PETRUS responsável pela transcrição de homógrafos heterófonos descrito anteriormente pretende abordar esse tipo de problema. Vale lembrar que o nosso sistema não se dispõe a desambiguar os casos de homógrafos heterófonos, mas apenas indicá-los quando presentes, uma vez que para desambiguá-los seria necessário analisar o contexto do qual foram retirados, e o contexto abordado pelo PETRUS é o de analisar apenas palavras simples, fora de contexto. Como trabalho futuro, pretendemos aumentar o contexto de sua atuação.

Por fim, certas palavras são ainda mais complicadas de serem transcritas, pois a identificação de sua classe gramatical não é suficiente para determinar sua correta transcrição fonética. Nesses casos, se faz necessário o acesso à informação semântica da palavra e o contexto no qual se encontra inserida. Na sentença “Que sede tem essa torcida!”, dependendo do significado, a pronúncia de <sede> pode ser s[e]de ou s[ɛ]de. Apenas analisando o contexto da sentença podemos inferir a pronúncia mais adequada, pois em ambas as possibilidades, temos um substantivo. Trata-se de homógrafos heterófonos com mesma classe gramatical. O módulo do sistema PETRUS responsável pela transcrição de homógrafos heterófonos descrito anteriormente pretendeu abordar também esse perfil de problema, com uma lista, ainda que pequena, de pares de palavras com esse tipo de característica.

Palavras constituídas por morfemas à esquerda de sua base também fogem às regras de transcrição elaboradas. Considere, como exemplo, as palavras <teleconferência>, <fotocopiadora>, <socioeconômico> e <ecocardiograma>, nas quais a primeira vogal <e> ou <o> possuem pronúncia [ɛ] ou [ɔ], respectivamente, diferente do que acontece com palavras <televisão>, <fotografia>, <sociologia> e <ecologia>, que tiveram sua pronúncia modificada pela frequência de uso, por exemplo. Um algoritmo para identificação de morfemas que precedem sua base evitaria que as mesmas tivessem que fazer parte de um dicionário de

exceções. Frente a isso, decidimos criar o módulo específico para a transcrição de tais casos com o objetivo de minimizar a ocorrência desse tipo de problema. Assim, criamos um algoritmo que transcreve apenas os morfemas à esquerda de sua base, cujas vogais são pronunciadas como [ɔ] ou [ɛ].

Como já vimos, outros problemas relacionados às idiosincrasias da língua (DUTOIT, 2001) incluem: o fato de (i) uma sequência de caracteres poder corresponder a um único fone, como por exemplo com os dígrafos em <banha>, onde o <nh> é transcrito como apenas um fone; (ii) um caractere não corresponder a nenhum fone, como o <h> de <haja>; e (iii) o mesmo caractere poder ser transcrito de duas maneiras diferentes, de acordo com seus caracteres precedentes e subsequentes. Nas palavras <asma> e <casca>, por exemplo, o <s> deve ser transcrito com som de [z] e [s], respectivamente. O inverso também ocorre: um mesmo fone pode ser representado por grafemas distintos, como o grafema <g> em <gente> e o grafema <j> em <jarra> representam o fone [ʒ]. Há ainda o caso de um fonema ser representado por uma sequência de grafemas (considere como exemplo o [x] de <barro>, o [ʃ] de <chato>, o [ʎ] de <lhama> e o [ɲ] de <banha>), entre outros casos.

Podemos depreender do contexto acima descrito, que a tarefa de conversão de grafemas em fones no PB não se trata de um simples mapeamento entre uma unidade gráfica e seu correspondente sonoro (fonético). Ao contrário, trata-se de um cenário para o qual tanto a indústria quanto pesquisadores da área de tecnologia de fala têm se dedicado à minimização dos erros causados por módulos de conversão grafia-fone quer em Portugal (Trancoso et al. (1994); Teixeira et al. (1998); Caseiro (2003); Barros and Weiss (2006); Oliveira (2007); Braga (2008); Veiga et al. (2011, 2013)), quer no Brasil (Chbane (1994); Albano e Moreira (1996); Gomes (1998); Simões (1999); Ostermann Filho (2002); Seara et al. (2002); Barbosa et al. (2003); Almeida (2005); Hosn (2006); Silva et al. (2006); Vasilévski (2008); Siravenha (2009); Couto (2010); Latsch (2011)).

A Figura 19 ilustra o módulo de transcrição fonética automática desenvolvido para o sistema PETRUS. Para habilitar essa funcionalidade em nosso sistema, tivemos como aporte para sua execução uma bibliografia dedicada à pronúncia da Rede Globo, como os estudos de Silveira (2008) e de Cagliari (2011; 2012), suportado com um corpus de estudo, derivado a partir do UFPAdic 1.0,⁹³ um dicionário fonético composto pelas 65.532 palavras mais

⁹³ O UFPAdic1.0. Trata-se de uma lista de palavras transcritas foneticamente com o alfabeto SAMPA, sem considerar uma variedade linguística específica do PB, cujo formato em .txt, possui palavras de diferentes tipos: substantivos, adjetivos, verbos (conjugados ou não), preposições, nomes próprios, interjeições, estrangeirismos e

frequentes do *cópus* CETENFolha. Tal dicionário foi transcrito manualmente, neste trabalho, segundo a variedade linguística adotada e constitui nosso *cópus* de estudo, a partir do qual extraímos as regras implementadas no último módulo.

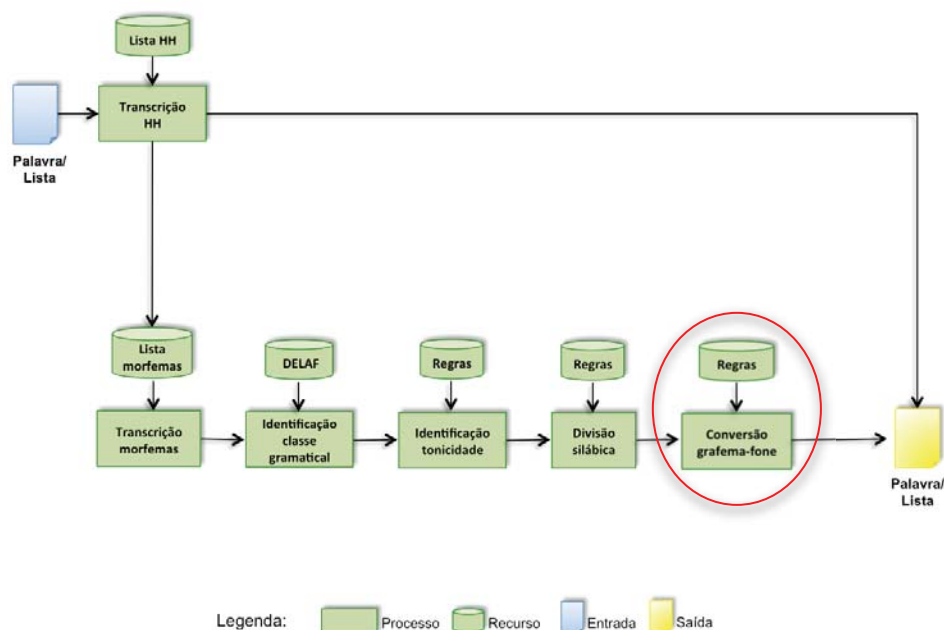


Figura 19. Etapa de transcrição fonética automática do PETRUS.

As informações adicionadas à palavra em análise até o Módulo 5 servem como *input* para a etapa de transcrição grafema-fone que ocorre no Módulo 6. O objetivo desse módulo consiste em determinar a sequência correta de símbolos fonéticos para representar adequadamente os caracteres gráficos contidos na palavra a ser analisada. A transcrição realizada pelo PETRUS é fonética, e não fonológica, de forma a representar os hábitos articulatórios da variedade linguística aqui considerada.

O conhecimento linguístico inerente ao nosso conversor grafema-fone é baseado em um conjunto de regras linguísticas dependentes de contexto que indicam como transcrever grafemas em suas unidades fonéticas correlatas, elaborado de forma manual, são executadas após ocorrer os processos de silabação, marcação de tonicidade e atribuição da classe gramatical à palavra-alvo.

siglas. Essas palavras foram escolhidas conforme sua frequência no CETENFolha, mas estão organizadas por ordem alfabética no UFPAdic1.0.

Foram criadas apenas regras que se aplicam nos limites de uma palavra, pois o processo de transcrição fonética automática realizado pelo sistema PETRUS atualmente é dedicado apenas à análise de unidades lexicais simples do PB.

A aplicação de tais regras parte do princípio de que é possível converter uma sequência de grafemas em uma sequência de símbolos fonéticos a partir da análise do contexto dos grafemas que lhe são adjacentes. Considere por exemplo:

Se <d> e <s>, então <e> = [ɪ]

Essa regra afirma, simplesmente, que se a vogal <e> aparece precedida por <d> e sucedida por <s>, então deve ser transcrita como [ɪ].

Muitas sequências de grafemas do PB são transcritas simplesmente a partir da aplicação de regras como essa, como é o caso das consoantes. Exceção se faz necessária em relação à transcrição automática da consoante <x>, que apesar de demandar um grande número de regras para sua transcrição, ainda assim, notam-se incorreções em alguns casos. Por exemplo, neste primeiro protótipo, o sistema PETRUS transcreve o <x> de <axila> como [x] e não como [ks], um erro comum em outros sistemas de conversão do PB, uma vez que nesse contexto intervocálico o som de <x> varia arbitrariamente. Como solução, poderemos utilizar um dicionário de exceções, com palavras frequentes do PB, para minimizar a ocorrência de erros no processo automático de transcrição fonética dessa consoante pelo PETRUS.

2.1.6.1 Conjunto de regras linguísticas para transcrição fonética

Nesta seção, apresentamos os símbolos fonéticos utilizados no conjunto de regras dedicados à transcrição fonética automática, além das regras propriamente ditas, que podem ser vistas na Tabela 14. Cada tabela corresponde a um grupo específico de fones e se apresentam organizadas da seguinte maneira: a coluna da esquerda contém os grafemas do PB, a seguir temos a coluna dos fones, sucedida pela colunas das regras de transcrição responsáveis pela conversão de grafemas do PB em seus respectivos fones, conforme a pronúncia da Rede Globo aqui adotada e, por fim, temos a última coluna, na qual constam exemplos gráficos com o grafema-alvo da regra em destaque (sublinhado).

Tabela 14. Regras de transcrição para os grafemas do PB

	Grafia	Fone	Regras	Exemplos gráficos
1.	a	[a]	-	ave
2.	aa	[a]	os casos de <aa> são transcritos como [a].	caatinga
3.	à	[a]	-	àquela
4.	á	[a]	-	árvore
5.	ã	[ẽ]	-	ímã
6.	a	[ẽ]	Quando ‘a’ é tônico e seguido por ‘m’, ‘n’ ou ‘nh’.	ban <u>ana</u> , <u>ca</u> ma, ba <u>na</u> ha
7.	â	[ẽ]		â <u>ma</u> go
8.	b	[b]	exceto quando sucedido por c, d, j, m, n, p, t, v.	<u>bo</u> lo, <u>B</u> rasil
9.	b	[bɪ]	quando sucedido por c, d, j, m, n, p, t, v.	o <u>bj</u> etiva, su <u>bm</u> arino
10.	bs	[bɪz]	quando precede <s> que está em sílaba tônica.	o <u>bs</u> équio, su <u>bs</u> ídio
11.	c	[s]	Quando preceder e, é, ê, i, í.	<u>ce</u> do, <u>ci</u> dade
12.	c	[kɪ]	Quando a sílaba seguinte é iniciada por consoante, exceto pelas consoantes “r” e “l”. Ou quando é a última letra da palavra.	pa <u>ct</u> o, tic-ta <u>c</u>
13.	c	[k]	Quando não preceder e, é, ê, i, í. Ou quando preceder “r”, “l” ou ç.	<u>ca</u> sa, <u>cr</u> avo, di <u>ç</u> ção
14.	ç	[s]	-	laço
15.	ch	[ʃ]	-	<u>ch</u> ave
16.	d	[d]	a. Antes das vogais a, â, ã, à, á, é, ê, ô, ó, o, u, ú; b. Ou se for seguida de consoante na mesma sílaba; c. Ou antes de <e> ou <o> tônico	<u>d</u> ata, <u>d</u> ente, ca <u>d</u> eia, <u>d</u> evo
17.	d	[dʒ]	a. Diante de <i>. b. Ou quando <e> é átono.	<u>d</u> ia, bo <u>d</u> e
18.	d	[dʒɪ]	Se não for sucedido por vogal na mesma sílaba	a <u>d</u> mirar
19.				
20.	ds	[dʒɪs]	-	a <u>d</u> stringente
21.	e	[e]	a. Se em posição pretônica não inicial e postônica não final; b. Ou se “e” + “m” ou “n” + vogal; c. Ou em posição postônica, mas não na última sílaba; d. Ou quando “e” inicia um pronome masculino, “e” é a vogal tônica; e. Ou quando “e” inicia a palavra + “x” + vogal; f. Ou quando “e” ou “ê” inicia a palavra + “x” + consoante; g. Ou a letra “e” + sequência “ne” ou “me”; h. Ou quando a letra “e” é seguida por “sa”, “ssa”, “za” em final de palavra; i. Ou quando “e” + letra qualquer + vogal tônica;	que <u>br</u> ante, et <u>er</u> ea em <u>er</u> ge, c <u>or</u> rego, e <u>l</u> e, ex <u>er</u> cício, err <u>ad</u> o, cre <u>m</u> e, rique <u>z</u> a, ex <u>ót</u> ico, engen <u>h</u> o, per <u>er</u> eca, por <u>qu</u> e, jap <u>on</u> esa bel <u>d</u> ade

			<p>j. Ou se “e” é seguido por “nh”;</p> <p>k. Ou se <e> ocorre em posição pretônica, mesmo se a vogal da sílaba tônica é média-baixa;</p> <p>l. Ou se (em posição inicial de palavra) e (a sílaba seguinte é iniciada por consoante, exceto “m”, “n”, “x”);</p> <p>m. Ou em final de palavra, quando é precedido da vogal ‘u’;</p> <p>n. Ou se “e” é sucedido por ‘sa, se, ssa, za, sas, ses, ssas, zas’.</p> <p>o. Quando “e” é átono e a sílaba é formada por “e” + “l”.</p>	
22.	e	[ɛ]	<p>a. Quando o “e” é tônico e a sílaba é formada por “e” + “l”.</p> <p>b. Quando é vogal tônica e é seguido por “la” ou “lo”, então e=e (ex: vela), exceto nas palavras “pelo” e “pela”, em que o “e” tem som de [e].</p> <p>c. As vogais [ɔ, ɛ] ocorrem em posição pretônica em formas derivadas com os sufixos -mente, -zinh, -inh ou -íssim quando o radical do substantivo/adjetivo apresenta [ɔ, ɛ]: <i>cipozinho</i> [sipɔ'zipɔ], <i>pezinho</i> [pe'zipɔ]</p>	me <u>l</u> , e <u>l</u> a, be <u>l</u> a,
23.	ee	[e]	Quando “ee” for átono.	re <u>e</u> scriver
24.	e	[ẽ]	Nos casos de nasalização, em que o <m>, <n> ou o <nh> aparecem na sílaba seguinte ao <e> e esse <e> é tônico	te <u>m</u> a, ve <u>n</u> ha, pe <u>n</u> a
25.	e	[ɪ]	<p>a. Em final de palavra;</p> <p>b. Em final de palavra + s;</p>	le <u>i</u> te, le <u>q</u> ues
26.	e	[i]	<p>a. Se a vogal [e] está em posição inicial de palavra e ocorre diante das fricativas alveolares [s, z].</p> <p>b. Ou se a vogal [e] está em posição inicial de palavra e ocorre diante de ‘xa’</p>	e <u>s</u> cola, e <u>x</u> ame
27.	é	[ɛ]	Exceto quando ocorre antes de <m> e <n>.	e <u>g</u> ua
28.	ê	[e]	-	vo <u>ç</u> ê
29.	f	[f]	Exceto se “f” é final de sílaba e a sílaba seguinte começa com consoante.	fa <u>ç</u> a
30.	f	[fi]	Quando “f” é final de sílaba e a sílaba seguinte começa com consoante.	a <u>ft</u> a
31.	g	[g]	<p>a. Antes de a, â, ã, à, á, ô, ó, o, u, ú;</p> <p>b. Ou antes das consoantes l,r;</p> <p>c. Ou antes dos encontros vocálicos <ui> e <ue></p>	ga <u>t</u> o, gl <u>ó</u> ria, gu <u>er</u> ra
32.	g	[gi]	Se “g” é a última letra da sílaba e é seguido por consoante, exceto “l”, “r”.	gn <u>o</u> mo
33.	g	[ʒ]	Se g for seguida por e, é, ê, i, í.	ge <u>l</u> o, gir <u>a</u> fa
34.	gu	[gu]	Seguido por <a>, <o>.	ling <u>u</u> agem, ex <u>í</u> guo
35.	gu	[g]	Antes de “e” e “i”.	gu <u>e</u> ixa, guit <u>a</u> rra
36.	gu (com som de gü)	[gu]	<p>a. Se a sequência <gue> é sucedida pelas consoantes <m, n, t, r, l>, então <u> é transcrito como [u].</p> <p>b. Ou se a sequência <gui> é sucedida pelas consoantes <d, ç, r, s, m, n>, então <u> é</p>	agu <u>e</u> nta, enx <u>a</u> gue, alc <u>a</u> guete ambigu <u>i</u> dade, arg <u>u</u> ir

			transcrito como [ʊ].	
37.	h		Em início de palavra não tem som.	h <u>ora</u>
38.	i	[i]	a. Se é final de sílaba e não precedido de vogal; b. Ou se é final de palavra seguido de r,l; c. Ou se é precedido de ku, gu, kü, gü	esquis <u>i</u> to, part <u>i</u> r, quit <u>a</u> do
39.	i	[ɪ]	a. Em sílaba átonas, principalmente, finais de palavras.	min <u>i</u>
40.	i	[ĩ]	Nos casos de nasalização, ou seja, em que o <m>, <n> ou o <nh> aparecem na sílaba seguinte ao <i> e esse <i> é <u>tônico</u>	mimo, nino, vinho
41.	j	[ʒ]		jac <u>a</u>
42.	k	[k]	Exceto se for a última letra de uma sílaba ou palavra, pois nesse caso ficará como, ki.	K <u>á</u> tia
43.	l	[l]	a. Início de sílaba e palavra; b. Ou seguindo consoante na mesma sílaba (bl,cl,fl,gl,pl,vl); c. Ou posição intervocálica	l <u>a</u> ta, b <u>l</u> oco a <u>l</u> a
44.	l	[ʊ]	a. Final de sílaba; b. Ou final de palavra; c. Ou se está em final de sílaba e a sílaba seguinte é iniciada por consoante	cal <u>d</u> o Bras <u>il</u> mal <u>d</u> ade
45.	lh	[ʎ]	Se ocorrer antes das vogais “a, e, o, u” no começo ou interior das palavras.	l <u>h</u> ama acol <u>h</u> edores
46.	lh	[l]	a. Se ocorrer antes de <e> em final de palavras; b. Ou se ocorrer antes da vogal <i>.	ental <u>h</u> e canal <u>h</u> ice
47.	m	[m]	a. Início de palavra + vogal; b. Ou seguindo consoante em sílaba distinta; c. Ou na sequência Vogal + m + vogal (posição intervocálica);	ma <u>c</u> a ad <u>m</u> inistrar am <u>a</u> relo
48.	m	[mɪ]	Se “m” é a última letra da sílaba e a sílaba seguinte é iniciada por consoante.	am <u>n</u> ésia
49.	n	[n]	a. Início de palavra; b. Ou seguindo consoante; c. Ou Posição intervocálica;	n <u>a</u> da car <u>n</u> e design <u>a</u> do a <u>n</u> o
50.	nh	[ɲ]	Quando não for seguida por -ia	gal <u>inh</u> a
51.	nh	[n]	Quando seguida por -ia	compan <u>h</u> ia
52.	o	[o]	a. Se em posição pretônica não inicial; b. Se precede “nh”; c. Na sequência: “o” vogal átona + “l” em final de palavra; d. Na sequência: “o” vogal tônica + “a”; e. Na sequência: “o” + “so”+ final de palavra; f. Na sequência: “o” + “r”+ final de palavra, exceto nas palavras “maior, menor, melhor, pior, suor”, nas quais “o” tem som de [ɔ]. g. Na sequência: “o” + qualquer letra+ vogal tônica; h. Se “o” inicia a palavra e é seguido por “l” ou “lh”; i. Na sequência “h”+ “o” + (r, ou s ou t); j. Se “o” + nh; k. Ou se “o” é seguido por ‘so’;	od <u>o</u> rífero od <u>o</u> r con <u>o</u> hecer sol <u>o</u> dado Lisbo <u>o</u> gost <u>o</u> so am <u>o</u> r op <u>o</u> r ol <u>o</u> aria hor <u>o</u> ário son <u>o</u> ho amor <u>o</u> so ole <u>o</u> so obrig <u>o</u> ado

			l. Ou se “ol” está no início da palavra; m. Na sequência ‘o’ em início de palavra e seguido por duas consoantes.	
53.	o	[ɔ]	a.Na sequência: “o” vogal tônica + “l” em final de palavra; b.Na sequência: “o” + “sa, sos, sas’+ final de palavra; c.Na sequência: “o” + “z”+ final de palavra, exceto na palavra “arroz”, em que o “o” tem som de [o]. d.Se ‘ol’ está no final da palavra, o = [ɔ] e.Ou se ‘o’ é seguido por ‘ra’	sol saudosa voz algor gostosa futebo agora
54.	o	[ʊ]	Sempre em posição final átona.	sapo
55.	o	[õ]	Nos casos de nasalização, ou seja, em que o <m>, <n> ou o <nh> aparecem na sílaba seguinte ao <o> e esse <o> é <u>tônico</u>	como, sono, sonho
56.	ó	[ɔ]	-	vovó
57.	ô	[o]	-	vovô
58.	oo	[oo]	Na sequência ‘oo’, quando não está em final de palavra.	cooperar
59.	p	[p]	exceto quando sucedido por b, c, ç, f, g, n, s, t, l na sílaba seguinte.	pato, professor
60.	p	[pi]	sucedido por b, c, ç, f, g, n, s, t, l.	pneu, psicologia
61.	qu	[kɔ]	Seguido de a, à, á, â, o, ɔ	quadro, quociente
62.	qu	[k]	Seguido de e, é, ê, i, í	queijo, quilombo
63.	qu	[kɔ]	a.Se “qu” + “e” + (n, s, i, l) e a letra “q” não é a primeira letra da palavra, o identificador fonético pronuncia a vogal “u”, caso contrário, permanece a regra das sílabas “qu”. b. Se a letra <q> for sucedida pela vogal <u>, mais a vogal <i> e uma das consoantes <s, c, r, v, d, q, l>, então <u> é transcrito como [ʊ].	cinquenta, sequência, liquificador, tranquilidade
64.	r	[x]	a. Início de palavras; b. Final de palavras; c. Após as consoantes “s,n,l” d. Antes das consoantes “p, t, c, qu, f, ch”.	rato, israel, honra, abaixar
65.	r	[r]	a. Entre vogais; b. Encontros consonantais (br, dr, gr, pr, tr, cr, fr, vr);	araras, brinco, corda,
66.	rr	[x]	-	carro
67.	r	[ɣ]	Final de sílaba antes de consoante sonora “b, d, g, v, z, j, m, n, nh, l, lh”.	gordo
68.	s	[s]	a. Sempre em início de palavras; b. Ou após as consoantes r,l,p,b,n; c. Ou em final de sílaba seguido por consoante desvozeada (s,t,p,k,f); d. Ou em final de sílaba seguido de consoante alveolar (t,d,n,l); e. Ou em final de palavras, ocorre apenas a fricativa [s].	sala, psicologia, descreve, destaque, dias
69.	s	[z]	a. Entre vogais; b. Ou seguida por uma consoante vozeada (nh,n,r,z,v,g,d,b,m,l,lh).	asa, musgo
70.	ss	[s]	-	assado

71.	sç	[s]	-	nas <u>ç</u> a, cres <u>ç</u> a
72.	sc	[s]	Se for seguida por “e, i” e suas variantes acentuadas	ascen <u>ç</u> ão
73.	sc	[sk]	Se for seguida por “a, o, u” e suas variantes acentuadas.	pis <u>ç</u> ar, afres <u>ç</u> o, bascul <u>ç</u> ante
74.	t	[t]	a. Antes das vogais <a, o, u>; b. Ou quando seguido na mesma sílaba por <r> ou <l>.	tap <u>a</u> , trap <u>o</u> , at <u>l</u> as, tel <u>a</u>
75.	t	[ʃ]	a. Diante de <i>; b. Ou antes de <e> em posição átona de final de sílaba.	tim <u>e</u> , tom <u>a</u> te
76.	t	[ʃi]	Quando não for sucedido por vogal na mesma sílaba.	at <u>m</u> osfera, rit <u>m</u> o
77.	u	[u]	a. Em sílaba átonas, principalmente, finais de palavras; b. Ou segundo elemento de ditongo; c. Ou nas sequências: (“k” ou “g”) + “o” + “vogal”; d. vogal + u + vogal	ócul <u>u</u> s, sa <u>u</u> dade, gu <u>u</u> arda, mau <u>u</u> á
78.	u	[u]	Em posição tônica e pretônica.	talud <u>u</u> e, urub <u>u</u>
79.	u	[ũ]	Nos casos de nasalização, ou seja, em que o <nh> aparece na sílaba seguinte ao <u> e esse <u> é <u>tônico</u> .	pun <u>h</u> a, un <u>h</u> a, um <u>a</u>
80.	ú	[u]	-	saú <u>d</u> e
81.	v	[v]	Exceto quando seguido de “n”.	vac <u>a</u>
82.	v	[vi]	Quando seguido de “n”.	ovni
83.	w	[u]	-	whis <u>k</u> y
84.	x	[ʃ]	a. Sempre em início de palavras; b. Ou se ocorrer após <en>; c. Ou se ocorrer após os ditongos <ai>, <ei>, <ou>; d. Ou se a palavra começa com consoante (exceto as consoantes “f” e “m”) + i + x.	x <u>í</u> cara, en <u>x</u> ada, caix <u>a</u> , peix <u>e</u>
85.	x	[ks]	a. Sempre em finais de palavras; b. Ou se a palavra começa com “f” ou “m” + i + x. c. Ou se (fl) + (e, u) + (x).	tórax, fixo
86.	x	[z]	a. Se a palavra é iniciada por “e”, “ê” ou “ine” + x + vogal e consoante; b. ou se a palavra é iniciada por “e”, “ê” ou “ine” + x + consoante exceto “v”. c. Ou se precede consoante vozeada (b,d,g,j,l,m,n,r,v,z); d. Ou se <e> inicia palavra + <x> + <a> ou <e> ou <o> ou <u>	ex <u>x</u> emplo, in <u>x</u> istente, ê <u>x</u> odo, ex <u>x</u> ame
87.	x	[s]	a. Seguido de consoante desvozeada (f, c, p, q, t, s); b. Ou se ocorrer <e, ê> em início de palavra + x + (c, f, p, t). c. Se ‘trou’ + ‘x’, então x = s. d. Ou se (‘m’, ‘pr’) + (o, ó, a, á) + (x) + (im). e. Ou se (au) + (x) + (íl, il).	in <u>x</u> plicável, ex <u>x</u> cluir, troux <u>e</u> , pr <u>x</u> imo, aux <u>x</u> ílio
88.	xc	[s]	Se “xc” for seguida por “e,i” e suas variantes acentuadas.	ex <u>ç</u> eção
89.	y	[i]	-	Y <u>u</u> ri

90.	z	[z]	Exceto final de palavra.	<u>z</u> ebra, agiliz <u>ar</u>
91.	z	[s]	Em finais de palavras.	rap <u>az</u>
92.	ua	[ua]	Quando não suceder (c,g,q).	estát <u>ua</u>
93.	(c,g,q) ua	[uɑ]	Após uma consoante oclusiva velar (c,g,q).	agu <u>ar</u>
94.	(c,g,q) ue	[uɐ]	Após uma consoante oclusiva velar (c,g,q).	enxag <u>ue</u>
95.	ue	[ui]	Quando não suceder (c,g,q).	tênu <u>e</u>
96.	ui	[ui]	Exceção: Quando a vogal <u> é seguida por consoantes nasais <m, n, nh> ou por <r>, deve se transcrita como hiato [i, u]. Exemplos: ra[i]nha, conclu[i]r, sa[i]rmos.	<u>c</u> uidar, <u>fl</u> uido
97.	muito	[muĩto]	-	<u>m</u> uito
98.	uo	[uɔ]	Se ‘uo’ suceder ‘q’.	qu <u>o</u> tista
99.	uo	[uɔ]	Em final de palavra, sucedendo sílaba tônica.	contínu <u>o</u>
0.	ul	[uɔ]	-	azu <u>l</u>
1.	ia, ea	[ia]	-	sé <u>ri</u> a, á <u>re</u> a
2.	ie	[ii]	-	sé <u>ri</u> e, cá <u>ri</u> e
3.	io	[io]	Precedido de (c, ss) e seguido de (n)	estacion <u>am</u> ento, acion <u>is</u> ta
4.	io	[iɔ]	a. Quando não precedido de (c, ss) e seguido de (n); b. Em final de palavra	auxí <u>l</u> io, demô <u>n</u> io
5.	iu	[iɔ]	-	<u>vi</u> u, sa <u>iu</u>
6.	il	[iɔ]	-	fá <u>ci</u> l
7.	ao, au	[aɔ]	-	ma <u>o</u> metano, <u>fl</u> auta
8.	ai	[ai]	Exceção: Quando a vogal <u> é seguida por consoantes nasais <m, n, nh> ou por <r>, deve se transcrita como hiato [i, u]. Exemplos: ra[i]nha, conclu[i]r, sa[i]rmos.	<u>b</u> aixa, <u>ca</u> ixa
9.	ái	[ai]	Se <i> for acentuada, as vogais passam a caracterizar um hiato e devem aparecer em sílabas separadas	saí <u>da</u>
0.	al	[aɔ]	Sempre que for seguido de consoante.	<u>al</u> moço
1.	al	[al]	Quando for seguido de <h>.	<u>al</u> ho
2.	aú	[au]	-	saú <u>da</u> , saú <u>va</u>
3.	ei	[ei]	-	are <u>ia</u> , fe <u>ir</u> a
4.	eí	[ei]	-	ve <u>í</u> culo
5.	éi	[ɛi]	-	pap <u>é</u> is
6.	io, eo	[iɔ]	-	armá <u>ri</u> o, subterrâ <u>ne</u> o
7.	éo	[ɛɔ]	-	auré <u>o</u> la
8.	éu	[ɛɔ]	-	c <u>é</u> u
9.	eú	[eu]	-	cont <u>eu</u> do
0.	eu	[eɔ]	-	<u>m</u> eu
1.	el	[ɛɔ]	Se ocorrer em fim de palavras. Caso contrário, [eɔ]	an <u>el</u>
2.	ol	[ɔɔ]	Em final de palavras.	ment <u>o</u> l
3.	oi	[oi]	-	apo <u>iar</u>
4.	oi	[ɔi]	A sequência gráfica <oi> será transcrita como [ɔi] quando sucedida por “a, am, co, ca, de, e, em”.	<u>j</u> ibo <u>ia</u> , apo <u>iam</u> , <u>h</u> ero <u>ic</u> o, <u>pe</u> restro <u>ic</u> a, <u>as</u> teroi <u>de</u> , apo <u>ie</u> , apo <u>iem</u>
5.	oí	[oi]	-	ego <u>í</u> smo
6.	oe	[oi]	-	abenç <u>oe</u>
7.	oa	[ua]	-	nó <u>do</u> a
8.	oo	[ou]	Em final de palavra.	<u>vo</u> o

9.	oó	[ɔ]	-	alcoó <u>l</u> ico
0.	ou	[ou]	-	sou <u>ber</u>
1.	oou	[ooɔ]	-	vo <u>u</u>
2.	ãe, ãi	[ẽĩ]	-	mã <u>e</u> , câ <u>i</u> bra
3.	ão	[õ]	-	mã <u>o</u>
4.	am	[õ]	-	for <u>am</u>
5.	õe	[õĩ]	-	coraç <u>õ</u> es
6.	uim	[ũ]	-	ru <u>im</u>
7.	quai	[qɔai]	-	qu <u>ai</u> s
8.	guai	[gɔai]	-	igu <u>ai</u> s
9.	guão	[gɔõ]	-	sagu <u>ã</u> o
0.	quan	[kɔẽ]	-	qu <u>an</u> do
1.	guei	[gɔei]	-	enxagu <u>e</u> i
2.	am, an, âm, ân	[ẽ]	-	m <u>am</u> bo, cant <u>a</u> , âmb <u>i</u> to, ambulâ <u>n</u> cia
3.	am	[õ]	Apenas em final de palavra	fal <u>am</u>
4.	ão	[õ]	-	naç <u>ã</u> o
5.	em, en, ên, êm, ém, éns	[ẽĩ]	-	sem <u>pre</u> , at <u>en</u> ção, b <u>ên</u> ção, ret <u>ê</u> m, tamb <u>é</u> m, parab <u>é</u> ns
6.	im, ím, in, ín	[ĩ]	-	im <u>pr</u> ovável, í <u>mp</u> ar, in <u>s</u> ano, í <u>n</u> timo
7.	õe	[õĩ]	-	p <u>õ</u> e
8.	om, on, ôn	[õ]	-	bomb <u>o</u> m, con <u>ch</u> a, cônc <u>av</u> o
9.	um, un, úm	[ũ]	-	com <u>u</u> m, <u>un</u> tar, cúmp <u>l</u> ice
0.	-ear	[iar]	Nos finais de palavras em -ear ocorre uma pronúncia com hiato	gram <u>pe</u> ar
1.	-oar	[oar]	Nos finais de palavras em -oar, ocorre uma pronúncia com hiato	amont <u>o</u> ar
2.	-ário	[ariɔ]	Nos finais de palavras em -ário ocorre uma pronúncia ditongada	arm <u>á</u> rio
3.	-ório	[ɔriɔ]	Nos finais de palavras em -ório ocorre uma pronúncia ditongada	not <u>ó</u> rio

No processo de transcrição realizado pelo PETRUS, o algoritmo associa um dado grafema ao seu respectivo fone, seguindo a ordem esquerda-direita de ocorrência de cada grafema da UL em análise. Uma vez realizada a transcrição fonética, as informações sobre tonicidade, divisão silábica, classe gramatical são organizadas e disponibilizadas ao usuário, como já exemplificado na Figura 12.

2.1.6.2 Módulo para flexão e transcrição fonética automática de verbos do PB⁹⁴

A velocidade com que a tecnologia é desenvolvida tem provocado mudanças na sociedade cada vez mais profundas, impactando o dia-a-dia das pessoas. Exemplos não faltam. Sistemas computacionais que utilizam aplicações de voz estão presentes em nosso cotidiano, por exemplo, nos permitindo o acesso a serviços bancários e de compra de produtos via telefone, sistemas de ajuda à pessoas com problemas motores, visuais ou mesmo de fala, sistemas de localização (GPS), interação com celulares, televisores e *videogames* via voz, entre outros. Em uma atualidade motivada para o uso de várias aplicações tecnológicas ativadas pela fala, tanto a indústria quanto pesquisadores da área de tecnologia de fala têm se dedicado, por exemplo, a minimização dos erros causados por um módulo específico contido em sistemas dedicados à síntese de fala: o módulo de conversão grafema-fone.⁹⁵

A dificuldade que os atuais sistemas computacionais que realizam a conversão grafema-fone em PB têm ao converter as vogais médias em posição tônica de verbos nos levou a conduzir a pesquisa ora descrita que visou sistematizar a pronúncia de verbos regulares e irregulares do PB. O resultado desta investigação tornará possível implementar futuramente no sistema PETRUS uma função adicional: a de, a partir de um verbo no infinitivo, derivar automaticamente a flexão verbal bem como a transcrição fonética dessa flexão.

Diante desse desafio, apresentamos, a seguir, os resultados obtidos.

2.1.6.2.1 O sistema verbal do português: características morfofonológicas

A morfologia verbal do português organiza-se não apenas em função da concatenação de afixos (fal.á.va.mos), mas também por alterações de ordem fonológica como, por exemplo, alternância consonantal no final do radical (medir > meço), acréscimo de consoante ao radical (ver > vejo), ditongação no radical (caber > coube), nasalidade (por > ponho), alteração de vogal temática (ver > visse) e alternância vocálica morfemática de número-pessoa (ferir > firo

⁹⁴ Este estudo é resultado do estágio de pesquisa realizado no Instituto de Tecnologia da Universidade de Coimbra, Portugal, sob orientação do professor Fernando Perdigão e da pesquisadora da Microsoft Sara Candeias (Bolsa BEPE-FAPESP Processo: 2014/00613-5).

⁹⁵ Módulo responsável pelo processo de conversão de um texto de entrada representado em caracteres ortográficos (ou *grafemas*) para uma representação de um texto de saída constituído por seus correlatos sonoros (normalmente caracteres fonéticos ou fonológicos), de forma inequívoca. Através da transcrição fonética de um texto escrito é possível inferir a pronúncia das palavras nele contidas. O sucesso dessa etapa é importante para a implementação de sistemas computacionais visando à produção de sons corretamente.

> fere). Nesta seção, trataremos desse último caso da chamada alomorfia de vogal da raiz verbal.

A alternância vocálica afeta as formas verbais do PB, cujos radicais possuem vogal média alta [e, o] ou média baixa [ɛ, ɔ]. Segundo Bisol (1996) e Mateus (1982), um caso restrito de alternância vocálica é a harmonia verbal que ocorre em verbos da 2^a e 3^a conjugações. Na harmonia verbal, a última vogal do radical (com exceção de /a/) se harmoniza em altura com a vogal temática do verbo considerado. Por exemplo, podemos observar a harmonização vocálica dos traços da vogal do radical do verbo <sentir> condicionada por sua vogal temática [i] que apresenta traço fonológico [+alto], o que difere do traço [-alto] da vogal [e] do radical “sent-”. Dessa maneira, a vogal do radical altera seu status de [-alta] para [+alta], resultando na forma “sint-”.

O problema para a conversão automática de grafia em fone ocorre, no entanto, nos casos em que se aplica a regra de abaixamento vocálico, que é aplicada apenas em vogais com o traço [+acento] e se a regra de harmonização vocálica não for aplicada. A regra de abaixamento determina que uma vogal com traços [-alto] e [+acento] se torne [+baixa]. Aplicada após a regra de acento,⁹⁶ essa regra de abaixamento é responsável por transformar o primeiro [e] de [be'bex] em [ɛ] na forma flexionada ['bɛbis], da 2^a pessoa do singular, do Presente do Indicativo, por exemplo, pois estabelece que uma vogal com os traços [-alto] e [+acento] se torne [+baixa].

Vários autores têm descrito a alternância vocálica em certas raízes verbais com determinados sufixos flexionais, em diferentes quadros teóricos. Harris (1974) afirma que o fenômeno deve ser pensado como prova da *elsewhere condition*, por exemplo. Galvão (1993) fornece uma revisão das análises lineares do fenômeno da alternância vocálica realizadas por Harris (1974), Redenbarger (1981) e Mateus (1982) e o reanalisa considerando a perspectiva da geometria de traços. Wetzels (1995), por sua vez, propõe testar especificamente a adequação de descrever a altura da vogal, considerando a proposta de Clements (1991). Há ainda trabalhos como o de Cristófar-Silva e Campos (2008) que analisam a variação existente entre vogais médias e altas relacionadas com formas verbais irregulares do PB que apresentam terminações infinitivas em -ear (chantagear) e -iar (variar).

Especificamente sobre a pronúncia das vogais médias pretônicas, sem dúvida, uma das riquezas do PB é a gama de diferentes realizações do sistema vocálico em pauta pretônica, com predominância das realizações abertas ou fechadas de acordo com cada região do país.

⁹⁶ Regra de acento: determina que a vogal de um radical verbal recebe o traço [+acento] quando esta é sucedida por CVC ou CV em final de palavra.

Trata-se, portanto, de um tema recorrente de pesquisa, como podemos constatar a partir dos trabalhos de Bisol (1981), sobre a regra de harmonia vocálica no falar gaúcho;⁹⁷ Callou & Leite (1986), sobre a ação da regra de harmonia vocálica na norma culta do Rio de Janeiro; Maia (1986), sobre o comportamento das vogais pretônicas médias na fala de Natal; Viegas (1987), sobre o alçamento das vogais médias pretônicas na região de Belo Horizonte, em Minas Gerais; Silva (1989), sobre as vogais pretônicas na variedade culta de Salvador, na Bahia; Bortoni et al. (1992), sobre os condicionamentos das regras de elevação e abaixamento no dialeto de Brasília; Freitas (2001), sobre as vogais médias pretônicas na cidade de Bragança, no Pará; Schwindt (2002), sobre a regra de harmonização vocálica no Rio Grande do Sul; Marques (2006), sobre as vogais médias pretônicas em situação de contato dialetal entre paraibanos e cariocas no Rio de Janeiro e entre brasileiros e portugueses na cidade de Lisboa e, por fim, podemos citar Carmo (2009) que estudou as vogais pretônicas de verbos na fala culta do interior paulista. Além de serem consideradas um fator que diferencia os dialetos brasileiros, as realizações do sistema vocálico em pauta pretônica também caracterizam diferenças entre o PB e o PE;⁹⁸ sendo, talvez, uma das características que mais diferenciam as duas variedades, conforme ressaltou Teyssier (1997 apud Viegas, 2001).

Como podemos depreender desse panorama não exaustivo, a pronúncia das vogais das formas verbais em PB é um assunto amplo e complexo que ainda necessita ser abordado de um ponto de vista sistemático para ser aproveitado pelos sistemas computacionais desenvolvidos na área de tecnologia da linguagem, eixo condutor de nossa investigação. Neste estudo, restringimos nosso objetivo a identificar os tipos de irregularidades possíveis na raiz verbal, considerando-se o sistema flexional dos verbos do PB. Portanto, decidimos também investigar primeiro a pronúncia verbal em obras de referência do PB como língua materna. Para tanto, consultamos as gramáticas normativas de Bechara (2009) e Cunha; Cintra (2002). Investigamos o tema da flexão verbal irregular (chamada, nas gramáticas, de conjugação) e também não observamos um estudo sistemático que pudesse ser prontamente aproveitado para o desenvolvimento do sistema ora proposto. Isso porque as explicações e exemplificações para o fenômeno da alternância vocálica ou consonantal não tiveram como

⁹⁷ Especificamente em relação à harmonização vocálica nas vogais pretônicas dos verbos, Bisol (1981) destaca a influência dos sufixos verbais, lembrando que, em verbos de 2ª e de 3ª conjugações, há uma abundância de condicionadores na flexão verbal, que cria vogais altas e que se espalha pelo paradigma,

⁹⁸ Conforme aponta Bisol (1981), nossa herança tomou rumos diferentes daqueles tomados pelas pretônicas em Portugal em cujo dialeto as posteriores tornam-se altas enquanto as anteriores são produzidas com /e/ mudo. Nesse sentido, segundo Teyssier, (1997 apud VIEGAS, 2001) a pronúncia brasileira perpetua a pronúncia do português europeu antes de mutações que nele ocorreram no século XVIII.

base um estudo que englobasse frequências derivadas de córpus extenso, nem descrevem de forma exaustiva e representativa o sistema de pronúncia verbal do PB, considerando todas as flexões verbais irregulares possíveis desse idioma e paradigmas verbais para sua conjugação.

Enfim, ao analisar essas e outras obras de referência sobre o tema, acabamos chegando a mesma constatação de que o que nossas gramáticas alinham, em ordem alfabética, como verbos irregulares, deveria ser entendido como um desvio do padrão geral morfológico, que não deixa de ser regular, no sentido de que é suscetível também a uma padronização. Assim, para os propósitos deste trabalho, notamos ser necessária a realização de uma investigação mais abrangente, para que, assim, pudéssemos, com base em córpus, ter um melhor conhecimento da sistematicidade da língua e encontrar padrões recorrentes (regularidades) nas irregularidades. Só então, de posse desse conhecimento, é que foi possível construir a parte do sistema dedicada a realizar a transcrição fonética de verbos do PB.

2.1.6.2.2 Metodologia

A metodologia adotada para a construção das regras linguísticas que serão inseridas no sistema de flexão verbal do PB é a mesma descrita em Candeias et al. (2013) para o PE. Consistiu em uma análise sistemática da pronúncia das formas verbais flexionadas do PB, que englobou, de forma geral: a análise da vogal temática dos verbos (que faz parte dos sufixos), a identificação do paradigma verbal, bem como a alternância do timbre vocálico da última vogal do radical verbal. Além disso, teve como aporte para sua execução bibliografia existente sobre o tema da flexão verbal no PB (citadas na seção anterior), e suportado com 1.000 verbos provenientes de um córpus de estudo, derivado a partir do CETENFolha.

Neste estudo, tratamos apenas das irregularidades referentes à pronúncia, ou seja, a parte do sistema computacional dedicada à transcrição fonética das formas verbais flexionadas, dada as dimensões limites que um assunto como este poderia adquirir se tomado em sua totalidade.

No nível ortográfico, alguns sistemas dedicados à flexão automática de verbos foram construídos para o português e estão disponíveis gratuitamente para uso na Web. São eles: o Lxconjugator,⁹⁹ o PRIBERAM dicionário¹⁰⁰ e o Conjuga-me¹⁰¹.

⁹⁹ LxConjugator <http://lxcenter.di.fc.ul.pt/services/en/LXServicesConjugator.html>.

¹⁰⁰ PRIBERAM dicionário: <http://www.priberam.pt/dlpo/Conjugar>.

¹⁰¹ Conjuga-me. <http://www.conjuga-me.net>.

Em relação ao *cópus* de estudo, o CETENFolha foi adotado para ranquear os 1.000 verbos mais frequentes. Esse *cópus* foi escolhido por ser extenso o suficiente para a obtenção de uma frequência de ocorrências representativa do uso de verbos no PB (pelo menos em textos jornalísticos) e também porque todas as sentenças estão anotadas morfossintaticamente, o que facilita a obtenção de formas verbais apenas. No processo de ranqueamento dos verbos mais frequentes, todas as formas flexionadas foram consideradas, exceto os participípios passados, dada a ambiguidade gramatical que possibilita que sejam classificados tanto como adjetivos quanto como verbos.

2.1.6.2.2.1 Metodologia para marcação automática das irregularidades verbais

Admitimos, neste trabalho, que um verbo regular é definido pela concatenação de seu radical com um dado sufixo sem qualquer tipo de alteração no radical ou sufixos. Assim, uma forma verbal é regular, se e apenas se, seguir este padrão de flexão. Um verbo será considerado menos regular que outro se tal verbo apresentar um número maior de irregularidades, ou seja, se houver alterações no radical e nos sufixos em relação à forma verbal em comparação.

O processo de comparação de duas formas uma forma verbal irregular (correta) e de sua forma derivada (como se fosse regular) consiste na contabilidade de inserções, substituições e apagamentos de letras ou símbolos fonéticos na forma derivada, para se chegar à forma correta. Esta comparação permite observar a existência de irregularidades e o local onde ocorrem, se no radical, nos sufixos ou em ambos. Muitas vezes, as irregularidades recaem no radical, sendo necessário mais de um radical para concluir a conjugação dos verbos, como acontece com o verbo <estar>, que admite 12 tipos diferentes de radicais para conjugá-lo em sua forma fonética ([est], [esto], [esta], [estʃivɪ], [estʃivɛ], [estevɪ], [estʃive], [esteʒv], [esteʒẽ], [esteʒa], [esteʒẽw̃], [estʃivesɪ]), enquanto que para sua forma ortográfica são necessários 8 tipos (<est>, <estou>, <está>, <estão>, <estive>, <estive>, <estivé>, <esteja>).

Quanto mais irregularidades um verbo apresentar, envolvendo número de radicais diferentes, de sufixos e de irregularidades total, maior será seu índice de irregularidade. Para contabilizar computacionalmente essa irregularidade, utilizamos a distância de Levenshtein (1999) ou distância de edição entre duas *strings* (sequências de caracteres), dada pelo número mínimo de operações necessárias para transformar uma *string* em outra, para realizar

contagem e identificação das irregularidades verbais, seguindo o mesmo raciocínio adotado para a construção do sistema de pronúncia verbal para a vertente europeia do português (para mais detalhes, conferir Candeias et al., 2013). Para mais informações sobre o cálculo da distância de Levenshtein, vide Capítulo 3.

Tal algoritmo realiza o alinhamento ótimo entre as duas formas de pronúncia possíveis, a real (irregular) e a idealizada, como se o verbo fosse conjugado de forma regular. Se a medida de distância entre essas duas formas não é nula, então dá-se início à realização das alterações necessárias: se o sufixo regular está presente no verbo, as alterações ocorrem em seu radical; caso contrário, há alteração no sufixo; ou ainda no sufixo e no radical. Nesse processo, de partir da forma regular idealizada para o verbo e conjugá-lo de forma correta, irregular, todos os radicais e sufixos alterados são contabilizados.

A seguir, exemplifica-se essa metodologia de anotação automática utilizada para a verificação de irregularidades. Considere a pronúncia do verbo <fazer> flexionado na primeira pessoa do Futuro do Indicativo [fa·'rei] e a sua forma pronunciada como se fosse um verbo regular [faze'rei]. Neste exemplo, observa-se que a forma correta é irregular, no sentido em que se nota a ausência de dois fones (indicados pelos pontos). Como o sufixo regular desta forma é [erei] existe, então, uma alteração no sufixo ([rei] em vez de [erei]), mas também no radical ([fa] em vez de [faz]).

Por meio deste método, as flexões verbais são comparadas e as irregularidades contabilizadas. A partir da contagem de todas as ocorrências de irregularidades em todas as formas flexionadas de um verbo, foi possível estabelecer um índice de irregularidade para os verbos paradigma deste estudo, que são apresentados mais adiante, na Tabela 27.

Esse mesmo algoritmo foi utilizado tanto para a análise de nossos dados, quanto foi inserido no sistema de pronúncia verbal construído para o PB. Todos os tipos de alterações realizados nos verbos para podermos conjugá-los de forma correta, como sendo verbos irregulares do PB, serão mostrados ao usuário do sistema flexionador e transcritor de verbos. Para tanto, vide Anexo B, um exemplo de saída gerada pelo sistema computacional aqui descrito.

Passamos, a seguir, à descrição de resultados alcançados por meio desta pesquisa.

As formas flexionadas de grande parte dos verbos do PB podem ser deduzidas a partir da junção do radical verbal aos sufixos de tempo/modo/pessoa da flexão. Entretanto, existem algumas irregularidades nesse processo, relacionadas tanto à grafia quanto à pronúncia de algumas vogais e consoantes, como anteriormente mencionado. Neste estudo, tratamos apenas das irregularidades referentes à pronúncia, focalizando as vogais.

Retomamos que, em relação à pronúncia das formas verbais, muitos verbos do PB apresentam alterações de timbre ou altura da última vogal do radical, conforme nela recaia, ou não, o acento tônico. Em Cunha & Cintra (2002) e Bechara (2009), por exemplo, alguns casos relacionados a alterações de pronúncia verbal já foram identificados, porém não de forma exhaustiva para toda a língua, tampouco de forma sistematizada, em que todas as situações de interesse para definir a pronúncia de todos – ou pelo menos a maioria – dos verbos existentes foram apresentadas.

Cabe ponderar, entretanto, que alguns verbos considerados irregulares no PB constituem um conjunto de itens lexicais com características fonológicas próprias comuns, reguladas por regras. E que, ao sistematizarmos as mudanças possíveis de ocorrer em tais verbos, quanto às possibilidades de pronúncia, foi possível agrupá-los e organizá-los segundo o tipo e a quantidade de alterações de pronúncia possíveis, conforme o Tempo, o Modo e a Pessoa das formas verbais flexionadas. Para isso, cada um dos 1.000 verbos do corpus de estudo foi analisado quanto a sua irregularidade (se é regular ou não seguindo metodologia adotada), e quando irregular, quais tipos de alterações consonantais ou vocálicas ocorrem, quantas vezes ocorrem e em que tempos, modos e pessoas verbais se dão essas alterações. Após essa análise, foi possível identificar conjuntos de alterações possíveis e sistematizá-las. Vejamos, a seguir, exemplo do trabalho realizado o qual abrange aspectos de alomorfia verbal que não se restringem às vogais verbais. Optamos por uma apresentação mais geral do sistema para, assim, permitir melhor avaliação das decisões metodológicas que tomamos no que diz respeito às vogais em particular.

Tabela 15. Padrões de alteração de pronúncia conforme o tempo, modo e pessoas verbais, relevantes para a flexão dos verbos em PB.

<i>TMP</i>	<i>Descrição</i>	<i>Tempo/Modo/Pessoa</i>	<i>Mudanças</i>
1	Tipo de alteração em que a vogal tônica muda de [e] → [ɛ], [o] → [ɔ] ou [a] → [ɐ̃].	Acontece no Presente do Indicativo (1s,2s,3s,3p), Imperativo (2s) e Presente do Subjuntivo (1s,2s,3s,3p). Ex.: <errar>, <tocar>, <amar>.	9
1a	A vogal tônica do radical torna-se semiaberta [ɛ].	Presente do Indicativo (2s,3s,3p) e Imperativo (2s). Ex.: <beber>, <mover>.	5
1b	A vogal tônica do radical torna-se semiaberta [ɔ].	Presente do Indicativo (2s,3s,3p) e Imperativo (2s). Ex.: <subir>.	4
2a	Alteração da consoante do radical: [d] → [dʒ]; de [t] → [tʃ].	Presente do Subjuntivo (1s,2s,3s). Ex.: <mudar>, <cantar>.	3
2b	A combinação vogal + glide é desfeita.	Presente do Indicativo (3p). Ex.: <construir>.	1
3	Mudança na vogal tônica ou consoante do radical.	Presente do Indicativo (1s) e Presente do Subjuntivo (1s,2s,3s,1p,2p,3p). Ex.: <ouvir>, <ferir>.	7
4	Mudança de vogal para vogal + glide.	Presente do Indicativo (2s,3s) e Imperativo (2s). Ex.: <construir>, <trair>.	3
4a	Mudança de [zi] para [s].	Presente do Indicativo (3s) e Imperativo (2s). Ex.: <induzir>, <jazer>.	2
5	Mudança de consoante no radical: (dʒ→d) and (tʃ→t).	Presente do Indicativo (1s,3p) e Presente do Subjuntivo (1s,2s,3s,1p,2p,3p). Ex.: <decidir>, <partir>.	8
6	Mudança de consoante no radical: (t→tʃ) e (d→dʒ)	Presente do Indicativo (2s,3s); Passado Imperfeito do Indicativo (1s,2s,3s,1p,2p,3p); Passado Perfeito do Indicativo (1s); Imperativo (2s); particípio. Ex.: <debater>, <aprender>.	11
7	Mudança de [e] to [i] em verbos terminados em <-edir>.	Presente do Indicativo (1s,2s,3s,3p); Presente do Subjuntivo (1s,2s,3s,1p,2p,3p); Imperativo (2s). Ex.: <agredir>.	11
8	Inserção de [id] ou [ed] nas classes dos verbos <rir>, <ler> e <prover>.	Presente do Indicativo (2p); Imperativo (2p).	2

Fonte: Adaptada de Marquifável et al. (2014, p. 40).

Conforme podemos observar na Tabela 15, a coluna TMP refere-se ao conjunto de flexões onde a alternância (vocálica ou consonantal) acontece. No caso de considerarmos a linha referente a TMP1, podemos observar nesse primeiro grupo alteração da vogal tônica no radical verbal. Tal alteração ocorre no tempo Presente do modo Indicativo nas 1^a, 2^a e 3^a pessoas do singular e na 3^a pessoa do plural, no modo Imperativo da 2^a pessoa do singular e também, no tempo Presente do modo Subjuntivo das 3^{as} pessoas do singular e da 3^a pessoa do plural. Esse é o mesmo raciocínio adotado para caracterizar e agrupar as alterações presentes nos outros verbos analisados e que foram agrupadas dentro dos 8 conjuntos possíveis de alterações .

Interessante notar, que TMP1a e TMP1b são subgrupos de TMP1, ou seja, os tipos de alterações presentes nos grupos TMP1a e TMP1b estão contidas no grupo maior, o TMP1. O mesmo se dá com o grupo TMP4a, que é subgrupo de TMP4. No grupo TMP4a acontece a mudança dos fones [zi], presentes nos sufixos das formas verbais <induzir> e <jazer>, por exemplo, conjugadas no Presente do Indicativo (3s) e no Imperativo (2s), que se reduzem para [s]: <induzo>, <induz> ([ĩ'duzʊ], [ĩ'dus]); <jazes>, <jaz> (['zazis], ['zas]). Se tais verbos fossem conjugados de forma regular seriam, respectivamente: [ĩ'duzi] e ['zazi].

Os TMPs 2a, 5 e 6 estão relacionados com um processo fonológico típico da variedade escolhida, a palatalização das consoantes oclusivas dentais /t/ e /d/ diante de [i], [ɪ] ou [ĩ].

O conjunto de alterações caracterizado em TMP3 contempla alterações vocálicas ou consonantais que podem ocorrer em radicais verbais do PB, conjugados no Presente do Indicativo (1s) e Presente do Subjuntivo (1s,2s,3s,1p,2p,3p). Para exemplificar, podemos citar o verbo <ouvir>, que na 1ª pessoa do Presente do Indicativo é conjugado como <eu ouço>. A consoante [v] desse radical verbal, se converteu em [s]. Já no caso do verbo <ferir>, a alteração acontece nos mesmos modo, tempo e pessoas das formas flexionadas do verbo <ouvir>, entretanto a alteração de pronúncia que se dá é com a vogal tônica presente no radical verbal: [e] → [i] e [e] → [ɛ].

Em TMP7 observa-se apenas alterações do tipo [e] → [i] que pode acontecer em verbos como o <agredir>. Como podemos observar, esse tipo de alteração acontece em 3 tempos verbais diferentes: Presente do Indicativo (1s, 2s, 3s, 3p); Presente do Subjuntivo (1s, 2s, 3s, 1p, 2p, 3p); Imperativo (2s). Já no TMP8 acontece inserção das vogais [i] ou [e] + [d] nos verbos <rir>, <ler> e <prover> e suas variações <sorrir>, <reler>, etc.

Os conjuntos de alterações que apresentam o maior número de mudanças são o TMP6 e o TMP7, com 11 alterações cada, sendo o primeiro constituído por alterações do tipo consonantal e o segundo, vocálica. Em contrapartida, o grupo com menor quantidade de alterações em suas formas flexionadas é o TMP 2b, que tem como paradigma o verbo <construir>.

Há no PB verbos que apresentam flexões alternativas no que diz respeito a vogais, como o verbo <explodir>, conjugado na 1ª pessoa do singular, no Presente do Indicativo: <eu explodo> ou <eu expludo>. Nesses casos, onde se nota a existência de duas formas possíveis, adotamos como critério de escolha a opção pelas formas mais usuais. Assim sendo, o verbo <explodir> segue, neste trabalho, o paradigma de flexão do verbo <dormir>. O mesmo se deu com o verbo <convergir>, que foi conjugado pelo paradigma do verbo <ferir>. Uma exceção

houve apenas em relação a forma flexionada do verbo <negociar>, conjugada também na 1ª pessoa do singular, no Presente do Indicativo. Optamos, neste caso, pela forma <negocio> e não <negoceio>, apesar desta última seguir o padrão regular de flexão verbal.

Após analisar o tipo de alteração (vocálica ou consonantal) em cada um dos 1.000 verbos de nosso *córpus* de estudo, manualmente, considerando a variedade da Rede Globo aqui adotada e os tipos de mudanças possíveis de ocorrerem (Cf. Tabela 15) apresentamos, na Tabela 16, os 56 paradigmas verbais a partir dos quais, acreditamos ser possível flexionar todos os verbos do PB. Cada verbo paradigma é representante de um conjunto de verbos em que se observam alterações idênticas.

Além dos paradigmas verbais da primeira coluna, é mostrado nas quatro primeiras colunas: a pronúncia de sua forma verbal no infinitivo, conforme a variedade linguística adotada, o(s) TMP(s) (conjunto de alterações) a que cada verbo paradigma é representativo e os tipos de fones alterados. Nas quatro colunas subsequentes, temos a coluna *IP*, onde é indicado o número de flexões que apresentam alterações no radical e/ou sufixos face à forma regular derivada; a indicação do número de radicais (fonológicos) diferentes que são necessários contabilizar para ser possível conjugar todas as formas flexionadas de cada verbo, as alterações ocorridas nos sufixos e, por fim, o índice de irregularidade verbal (Ind). O asterisco presente em alguns verbos da Tabela 16 indica *participio pasado irregular*.

Tabela 16. Verbos paradigmas para o PB em ordem crescente de irregularidade.

	Verbo	Pronúncia	TMP	Mudança de fone	IP	Radical	Sufixo	Ind
	<i>pensar</i>	pẽ'sar	-	-	0	1	0	0
2.	<i>viver</i>	vi'ver	-	-	0	1	0	0
3.	<i>unir</i>	u'nir	-	-	0	1	0	0
4.	<i>pôr</i>	'por	-	-	0	1	0	0
5.	<i>abrir</i>	a'brir	-	(*)	1	2	1	3
6.	<i>mudar</i>	mu'dar	2a	d→dʒ	3	2	0	4
7.	<i>cantar</i>	kẽ'tar	2a	t→tʃ	3	2	0	4
8.	<i>induzir</i>	ĩdu'zir	4a	zi→s	2	2	2	5
9.	<i>jazer</i>	ʒa'zer	4a	zi→s	2	2	2	5
10.	<i>beber</i>	be'ber	1b	e→ɛ	4	2	0	5
11.	<i>mover</i>	mo'ver	1b	o→ɔ	4	2	0	5
12.	<i>frigor</i>	fri'zor	1b	i→ɛ	4	2	0	5
13.	<i>subir</i>	su'bir	1b	u→ɔ	4	2	0	5
14.	<i>ouvir</i>	ow'vir	3	v→s	7	2	0	8
15.	<i>valer</i>	va'ler	3	l→ʎ	7	2	0	8
16.	<i>escrever</i>	iskre'ver	1b	e→ɛ(*)	5	3	1	8
17.	<i>decidir</i>	desi'dzir	5	dʒ→d	8	2	0	9
18.	<i>partir</i>	par'tʃir	5	tʃ→t	8	2	0	9
19.	<i>construir</i>	kõstru'ir	4/2b	u→ɔj/ɔ	4	3	3	9
20.	<i>roer</i>	xo'er	1b	o→ɔj/ɔ	4	3	3	9
21.	<i>errar</i>	e'xar	1	e→ɛ	9	2	0	10
22.	<i>amar</i>	a'mar	1	a→ɸ	9	2	0	10
23.	<i>sentir</i>	sẽ'tʃ'ir	3/5	ẽj→i / tʃ→t	8	3	0	10
24.	<i>tocar</i>	to'kar	1	o→ɔ	9	2	0	10
25.	<i>apojar</i>	apoj'ar	1	o→ɔ	9	2	0	10
26.	<i>aprender</i>	aprẽ'der	6	d→dʒ	11	2	0	12
27.	<i>debater</i>	deba'ter	6	t→tʃ	11	2	0	12
28.	<i>prevenir</i>	preve'nir	7	e→i	11	2	0	12
29.	<i>polir</i>	po'rir	7	o→u	11	2	0	12
30.	<i>dormir</i>	dor'mir	3/1b	o→u/ɔ	11	3	0	13
31.	<i>ansiar</i>	ẽsi'ar	1	i→ej	9	2	3	13
32.	<i>ferir</i>	fe'rir	1b/3	e→ɛ / e→i	11	3	0	13
33.	<i>trair</i>	tra'ir	3+4	a→aj	10	2	3	14
34.	<i>requerer</i>	xeke'rer	1b/3	e→ɛ / e→ej	11	3	2	15
35.	<i>medir</i>	me'dzir	1/3/2b	e→ɛ/dʒ→s/dʒ→d	11	5	0	15
36.	<i>cobrir</i>	ko'brir	3/1b	o→u/ɔ(*)	12	4	1	16
37.	<i>rir</i>	'xir	7/8	x→x'i/'idʒi	13	3	5	20
38.	<i>ler</i>	'ler	3/1b/8	e→ej/l→l'e/l'edʒi	13	4	5	21
39.	<i>prover</i>	pro'ver	3/1b/8	eʒ / e / edʒi	13	4	5	21
40.	<i>perder</i>	per'der	1b/3/6	e→ɛ/d→k/d→dʒ	19	5	0	23
41.	<i>aprazer</i>	apra'zer			26	5	25	55
42.	<i>caber</i>	ka'ber			31	5	23	58
43.	<i>saber</i>	sa'ber			31	6	24	60
44.	<i>querer</i>	ke'rer			35	6	25	65
45.	<i>dar</i>	'dar			31	5	30	65
46.	<i>haver</i>	a'ver			35	9	28	71
47.	<i>ver</i>	'ver			37	6	29	71
48.	<i>poder</i>	po'der			42	10	23	74
49.	<i>estar</i>	is'tar			34	12	34	79
50.	<i>ir</i>	'ir			38	12	36	85
51.	<i>ter</i>	'ter			43	11	34	87
52.	<i>trazer</i>	tra'zer			45	7	37	88
53.	<i>dizer</i>	dʒi'zer			46	7	38	90
54.	<i>fazer</i>	fa'zer			46	8	38	91
55.	<i>vir</i>	'vir			43	11	39	92
56.	<i>ser</i>	'ser			44	12	38	93

Fonte: Marquiefável et al. (2014, p.42).

As alterações que ocorrem na forma ortográfica, mas que mantêm regularidade fonética, como as que ocorrem com as letras <g> e <j> do verbo <fingir>, por exemplo, <fingir>/<finjo>, não são consideradas irregularidades (MARQUIAFAVEL et al., 2014) neste trabalho que trata apenas das alterações de pronúncia, portanto, não são marcadas, nem contabilizadas nessa equação.

Considerando a coluna do índice de irregularidade (Ind) apresentada na Tabela 16, podemos constatar a organização de três grupos verbais de natureza distinta. Nas quatro primeiras linhas estão indicados os paradigmas regulares <pensar>, <viver>, <unir> e <pôr>, por isso apresentam TMP 0, ou seja, não há qualquer tipo de alteração na pronúncia de suas formas flexionadas. Ainda em relação às formas verbais regulares, vale lembrar que, em geral, assume-se que os verbos do PB estão agrupados sob três tipos possíveis de conjugação verbal: 1ª conjugação com vogal temática -a-, 2ª conjugação com vogal temática -e- e a 3ª conjugação, cuja vogal temática é -i-. Entretanto, adotamos aqui uma 4ª conjugação. Isso porque, o verbo <pôr> e seus derivados são considerados, em algumas obras, como sendo verbos irregulares da 2ª conjugação, devido, principalmente à sua etimologia latina. Uma vez que se flexionam da mesma forma, podendo, portanto, constituir um grupo flexional por si só, consideramos adequado classificá-los como verbos regulares da 4ª conjugação, com vogal temática -o-. Essa escolha também se justifica pelos objetivos desta pesquisa que visa à implementação de sistema computacional que, ao final, apresente o *output* correto das formas verbais.

Com base nesse mesmo índice de irregularidade é possível organizar o restante dos verbos em dois grupos: os que apresentam poucas irregularidades e os que são altamente irregulares. Assim, a partir da posição 5 até a 40, podemos observar uma sequência de verbos, cujas irregularidades apresentam um padrão de ocorrência regular, tornando possível identificar como e onde ocorrem. Da posição 40 para a 41 podemos observar uma divisão, na qual se observa um grande salto no índice de irregularidades de certos verbos, em decorrência de alterações em seus sufixos ou radicais. Podemos notar, portanto, que a partir do verbo <perder>, que apresenta índice de irregularidade 23, o verbo subsequente <aprazer> possui um índice de irregularidade 55.

Vale ressaltar que a maioria das irregularidades relacionadas com a pronúncia das flexões verbais irregulares ocorre da alternância vocálica, conforme se pode deduzir da coluna “Mudança de fone” da Tabela 16.

No que diz respeito ao funcionamento interno do sistema ora em discussão, foi atribuído um paradigma à forma infinitiva de um verbo (ortográfico). Nesse processo, o algoritmo poderá fazê-lo de três formas distintas: (i) considerar a última vogal ou consoante do radical verbal; (ii) procurar um padrão na forma ortográfica do verbo ou (iii) checar uma lista de verbos. Por exemplo, os verbos terminados em -ver, podem ser, a princípio derivados do verbo <ver>, como <prever>, <antever> ou <rever>. Entretanto, apesar de os verbos <escrever>, <reaver>, <prover> apresentarem o mesmo padrão de terminação verbal (-ver), eles são conjugados de maneiras distintas: <escrever> é conjugado como regular, apresentando apenas particípio passado irregular; <reaver> conjugado irregularmente como o verbo <haver>; enquanto que <prover> possui paradigma próprio.

Tal conjunto foi implementado – no formato de algoritmos – no sistema de flexão e transcrição de verbos do PB proposto. As regras foram testadas com a lista de 1.000 verbos compilada como cópula de estudo para este trabalho e os resultados obtidos indicam que os 56 paradigmas verbais apresentados neste trabalho podem ser representativos dos verbos do PB.

Tabela 17. Exemplos de regras utilizadas no sistema de flexão e transcrição de verbos proposto.

Regras	Descrição da regra
VERBOS DA 1ª CONJUGAÇÃO	
1	[e] -> [e] quando sucedido por <m,n,nh,j,x,ch,lh> em sílaba distinta. Caso contrário, [e] -> [E] e passa a ser conjugado pelo paradigma de <errar>. Exceção: <invejar> e <flechar> são conjugados pelo paradigma de <errar>, apesar de serem sucedidos por <j>.
2	[o] → [o] quando sucedido por <m,n,nh,j,x,ch,lh> ou terminar em -oar (ex.: abençoar). Caso contrário, será conjugado pelo paradigma do verbo <tocar>. Exceções: <aboiar> segue paradigma de <apoiar>. Mas, <alojar>, <arrojar>, <olhar>, <molhar>, <pimpolhar>, <piolhar> e derivados, seguem paradigma de <tocar>.
3	Verbos terminados em -ear são conjugados como -iar, ou seja, [e] → [ej] na forma rizotônica. (ex.: nomear, acarear, golpear, mediar, ansiar, incendiar, odiar, etc.)
VERBOS DA 2ª CONJUGAÇÃO	
4	Verbos terminados em -eb (perceber), -ec (abastecer), -ex (mexer), -eg (eleger), -ev (atrever), -erc (exercer), -erg (absterger, erguer), -esc (crescer), -screv (escrever), -erv (ferver), -et (cometer), -meter (acometer), -ert (verter), -ed (feder), -ced (abceder), são conjugados pelo paradigma <beber>. Exceção: <querer> (já está adicionado na lista de verbos irregulares).
5	Verbos terminados em -ofr (sofrer), -ol (moler), -olh (acolher), -olv (absolver, volver), -orc (contorcer), -orv (absorver), -orr (correr), -ov (mover), -os (coser), -oz (cozer), -ord (morder), são conjugados pelo paradigma <mover>. Exceção: <poder> (já está na lista de irregulares).
6	Os verbos <doer>, <moer> e <soer> são conjugados pelo paradigma <roer>.
7	Verbos terminados em -azer, como <trazer> e <jazer> seguem paradigma de <fazer>.
8	Os verbos <saber> e <caber> são conjugados pelo paradigma <saber>. O verbo <saber> com significado de sabor tem como 1ª pessoa do singular, no Presente do Indicativo, a forma <saíbo>, a qual pronuncia-se, nas outras formas, como o verbo <saber> com sentido de sabedoria.
VERBOS DA 3ª CONJUGAÇÃO	
9	Verbos terminados em -er (ferir, inserir), -ert (divertir), -erv (servir), -erz (cerzir), -esp (despir), -est (vestir), -egu (seguir), -ter (preterir), -el (expelir), -et (repetir, refletir) são conjugados pelo paradigma <ferir>. Exceção: <amareliar>, cujo paradigma é <pedir>.
10	Verbos terminados em -erg (convergir), -ern (concernir) possuem como paradigma <ferir>.
11	Verbos terminados em -ed (agredir, regredir), -egr (denegrir), -rem (fremir, premir), -en (prevenir) são conjugados pelo paradigma <decidir>.
12	O verbo <pedir> e derivado são conjugados como <medir>.
13	Os verbos <polir> e <sortir>, bem como seus derivados, são conjugados pelo paradigma <polir>.
14	Verbos terminados em -um (consumir, sumir) são conjugados pelo paradigma <dormir>. Exceção: <assumir>, <presumir>, <reassumir>, <resumir>, que são regulares.
15	Verbos terminados em -ub (subir), -ud (acudir, sacudir), -ug (fugir), -ul (bulir, escapular), -usp (cuspir), -up (entupir), -u (destruir) seguem o paradigma <subir>.
16	O verbo <construir> possui paradigma próprio: <construir>, assim como o verbo <frigir>.
17	Verbos terminados em -od (eclodir, erodir, explodir, implodir), -odr (apodrir), -for (esbaforir), -obr (cobrir) com pcp irregular, -orqu (extorquir, retorquir), -ort (enfortir, rostir) são conjugados pelo paradigma <mover>. Exceção: <sortir> que é conjugado como <polir>.
18	Verbos terminados em -ol (abolir, demolir, engolir, rebolir), -or (colorir, espavorir, florir), -obr (cobrir) com pcp irregular, -oss (tossir) são conjugados pelo paradigma <dormir>.
19	Verbos terminados em -a (abstrair, arraiar, atrair, cair, contrair, desvair, esvair, jumbair, parir, sair) seguem o paradigma <trair>.
20	Verbos terminados em -uzir (luzir, produzir, introduzir, abduzir, deduzir, conduzir) seguem o paradigma <induzir>.

Em suma, podemos concluir que a dificuldade que os atuais sistemas computacionais que realizam conversão grafia-fone em PB têm ao converter as vogais médias em posição

tônica de verbos nos levou a conduzir a pesquisa ora descrita que visou sistematizar a pronúncia de verbos regulares e irregulares do PB.

O produto desta investigação permitiu elaborar um sistema no qual, a partir de um verbo no infinitivo, é possível derivar automaticamente a flexão verbal bem como a transcrição fonética dessa flexão. Para tanto, um conjunto de 1.000 verbos regulares e irregulares, derivados do CETENFolha foram analisados quanto à pronúncia de suas formas flexionadas. Quanto às formas verbais consideradas pela literatura como irregulares, pudemos realizar uma sistematização das alternâncias vocálicas e/ou consonantais possíveis de modo que se constituísse um conjunto com 56 verbos paradigmas, a partir dos quais é possível derivar as flexões verbais de verbos (ortográficos) no infinitivo, submetidos como entrada para o sistema proposto. Em suma, para atribuir um paradigma à forma infinitiva de um verbo (ortográfico), o algoritmo poderá fazê-lo de três formas distintas: (i) considerar a última vogal ou consoante do radical verbal; (ii) procurar um padrão na forma ortográfica do verbo ou (iii) checar uma lista de verbos.

O sistema conjugador e transcritor de verbos do PB ora proposto poderá ser uma ferramenta interessante para o ensino-aprendizagem da flexão verbal do PB, como língua estrangeira ou não, bem como da pronúncia dessa flexão verbal, por meio dos símbolos fonéticos transcritos. Isso porque a pronúncia das formas verbais flexionadas em PB não é uma informação que se encontra em dicionários ou está exhaustivamente descrita em gramáticas. Como trabalho futuro, é previsto a incorporação do sistema descrito ao sistema PETRUS.

Conforme mencionado anteriormente, a motivação deste capítulo foi mostrar as etapas contidas no processo de conversão grafema-fone do sistema PETRUS. Diante do explanado, cada etapa exigiu recursos linguísticos específicos, ora constituídos por regras ora por listas (por exemplo, dicionário de HHs).

Importante lembrar que o projeto ora proposto tem perfil multidisciplinar, por estar inserido no campo de pesquisa que trata computacionalmente as linguagens humanas. E, como qualquer outro projeto dessa natureza, necessita para seu desenvolvimento pleno um trabalho conjunto de profissionais de diferentes áreas do conhecimento, no caso, profissionais das áreas de linguística e computação. Desta feita, a equipe deste projeto é composta atualmente por 3 linguistas (orientanda, orientadora Profa. Dra. Claudia Zavaglia e o coorientador Prof. Dr. Luiz Carlos Cagliari) e 1 informata para a implementação computacional do sistema proposto, bem como sua disponibilização online na página do

NILC. O referido informata é o mestrando em Ciências da Computação do NILC-ICMC-USP, Alessandro Bokan. As atividades sob sua responsabilidade consistiram na implementação das regras e recursos linguísticos criados e da alocação do sistema PETRUS no portal do NILC.

Após a conclusão da primeira versão do sistema PETRUS, fez-se necessário avaliar o desempenho com que ele realiza a tarefa de transcrição fonética automática. Dessa maneira, apresentamos os resultados preliminares obtidos no próximo capítulo.

CAPÍTULO 3

RESULTADOS

Segundo Oliveira (1996, p. 56), “é importante que o desenvolvimento e avaliação do desempenho das regras para a transcrição automática se fundamentem em córpis que representem de forma fidedigna o léxico comum da língua”.

Para avaliar o processo de transcrição realizado pelo sistema PETRUS, fizemos a comparação entre as transcrições de um dicionário fonético composto por 2.130 palavras (*tokens*) escolhidas aleatoriamente de nosso córpis de estudo e comparamos suas transcrições com as geradas automaticamente pelo sistema. O córpis utilizado para o teste, portanto, é derivado de um córpis de língua escrita, cuja transcrição fonética foi realizada manualmente pela doutoranda, tendo sido adotado como sistema de notação o *Alfabeto Fonético Internacional*, e seguindo-se as especificidades do dialeto adotado. Abreviaturas, siglas e estrangeirismos foram removidos do córpis, pois não são tratáveis pelo PETRUS ainda.

Consideraremos que o nosso sistema gerará automaticamente transcrições fonéticas de qualidade quanto mais próximo o percentual de acerto estiver de 100%.

Para contabilizar de forma automática a taxa de acerto da transcrição gerada pelo sistema PETRUS em relação à transcrição das palavras de nosso córpis de teste, utilizamos a *distância de Levenshtein* (Levenshtein, 2013) ou *distância de edição* entre duas *strings* (sequências de caracteres). Ela é dada pelo número de operações necessárias para que uma *string* fique igual à outra. Entende-se por operações, os procedimentos de inserção, apagamento ou substituição, que são calculados pela seguinte equação:

$$(\text{tamanho}(string1) - \text{distânciaedição}(string1, string2)) / \text{tamanho}(string1)$$

A título de exemplo, considere a palavra <existiam>, com [ezis'ʃiãõ] como string 1 e [exis'ʃiãõ] (como string 2). A distância de edição seria 1 porque só é necessário substituir [z] por [x] para que fiquem idênticas. O tamanho das strings em ambos os casos (contando os caracteres especiais) é o mesmo, 11. Desse modo:

$$(1) \quad (11 - 1)/11 = 0.909$$

O resultado desse cálculo mostra quão parecida uma *string* é da outra, e é indicado por valores que vão de 0 a 1.

O cálculo dessa distância é muito útil para ser usado em aplicações nas quais é preciso determinar o quão semelhantes são duas *strings*, como é o caso do teste que realizamos. Assim, para contabilizar a taxa de acerto da transcrição gerada pelo sistema PETRUS, medimos a distância de edição entre as transcrições do córpus de teste e as transcrições geradas pelo nosso sistema. A seguir, somamos todas as pontuações das transcrições geradas pelo sistema PETRUS e dividimos pelo total de fones para obtermos o valor de similitude entre as duas *strings*, ou seja, a taxa de acerto. Na Tabela 18, é possível observar os resultados da avaliação realizada.

Tabela 18. Taxas de acerto obtidas na avaliação do desempenho do sistema PETRUS

Resultados Preliminares do PETRUS		
<i>Total de palavras</i>	2.126	100%
<i>Palavras certas</i>	1.799	84.6%
<i>Total de fones</i>	16.820	100%
<i>Fones certos</i>	16.400	97.5%

A partir da Tabela 18, podemos observar que, no total de 2.126 palavras, existiam 16.820 fones a serem transcritos. Após a realização da conversão grafema-fone pelo sistema PETRUS, 16.400 destes fones foram convertidos corretamente. Esse valor corresponde a uma taxa de acerto de 97.5%, um resultado bastante aceitável.

Descrições de outros sistemas de conversão grafema-fone apresentam resultados que nos indicam que o desempenho atual do sistema PETRUS está um pouco acima da média geral de acerto publicada, que é de 89.98%. Simões (1999) cita uma taxa de acerto de 96% em

nível de palavra para o sistema Ortofon, que é baseado em regras. Silva (2006) cita uma taxa de 97,44% obtida em nível de fones corretamente transcritos, a partir de seu conversor grafema-fone também construído com base em regras linguísticas. Siravenha et al. (2006) e Neto et al. (2010) desenvolveram um conversor grafema-fonema com determinação de vogal tônica, no âmbito do projeto FalaBrasil. O algoritmo por eles proposto utilizou as regras de conversão de Silva et al. (2006), com algumas adaptações realizadas quanto às regras de transcrição das letras <a, u, x>. Entretanto, Neto et al. (2010, p. 57) ressaltam que o conversor desenvolvido não se baseia em nenhum dialeto específico do PB nem tão pouco apresentam a taxa de acerto do sistema por eles implementado.

A Tabela 19 traz uma comparação-resumo do desempenho dos sistemas de transcrição automática produzidos para o PB, com o uso da abordagem baseada em regras, comparados com os resultados preliminares do sistema PETRUS. Os outros sistemas conversores produzidos para o PB, por meio da abordagem baseada em dados, não são apresentados na Tabela 19 abaixo devido ao fato de tais sistemas representarem um problema de classificação. E, como tal, ao realizarem a classificação, apresentam quatro tipos de resultados para essa ação. Por outro lado, o sistema PETRUS, que é baseado em regras geradas manualmente, tem uma avaliação diferente: deve ser avaliado por taxa de acerto, que pode ser avaliada em nível de fones ou de palavras.

Tabela 19. Comparação entre os sistemas de conversão baseados em regras do PB.

Abordagem	Sistema	Taxa de acerto ao fone
<i>Regras linguísticas</i>	<i>Barbosa et al. (2003)</i>	98.43
<i>Regras linguísticas</i>	<i>PETRUS (2014)</i>	97.5%
<i>Regras linguísticas</i>	<i>Silva (2006)</i>	97.44%

No entanto, vale a ressalva de que apesar de as taxas de acerto/erro publicadas servirem como referência para sabermos se o sistema PETRUS apresenta um desempenho aceitável frente aos demais, é preciso atentar para o fato de que os corpú de teste utilizados pelos sistemas correlatos aqui citados não estão disponíveis para que diferentes sistemas sejam testados com um mesmo corpú e, conseqüentemente, terem seu desempenho comparado de

forma adequada. Além disso, os corpúsculos de teste por eles utilizados não são os mesmos, variam em tamanho, origem, tipos de palavras e frequência de fonemas, o que influencia também na taxa de acerto obtida. Por exemplo, quanto mais fonemas problemáticos contiver um texto, maior será a probabilidade de erros de um sistema. Assim, seria necessário realizar uma comparação do desempenho das ferramentas citadas com base em um mesmo corpúsculo de teste para só então podermos fazer um ranqueamento adequado do desempenho de sistemas de conversão grafema-fonema produzidos para o PB.

Por fim, podemos dizer quanto à taxa de acerto publicada pelos sistemas baseados em regras são apenas citadas, sem a menção sobre o método como foi obtido o resultado, nem tampouco se disponibiliza o resultado do teste realizado. Em relação ao PETRUS, disponibilizamos os erros ocorridos no Anexo C desta tese.

No que diz respeito aos tipos de erros encontrados na lista de palavras transcrita pelo sistema PETRUS, constata-se que a maior parte deles está relacionada com o timbre vocálico dos grafemas <e> (20.3%) e <o> (18.1%) que é um reflexo direto da dificuldade na determinação de sua pronúncia, uma vez que não há marcação gráfica (acento agudo ou circunflexo) para indicar o timbre correto. Trata-se essencialmente de erros de transcrição de vogal tônica de nomes (<alicerce>[ɛ], <bola>[ɔ]), adjetivos (<alegre>[ɛ], <forte>[ɔ]), e de formas verbais em que é frequente haver alternância da vogal tônica ao longo da flexão verbal (<acontece> [ɛ], <adota> [ɔ]).

O restante dos erros são, na grande maioria, casos de exceção à regra do <e> com timbre aberto, tais como “cr[ɛ]do” e “segr[e]do”, cujas tônicas apresentam a mesma categoria gramatical, o mesmo contexto anterior (<r>) e posterior (<do>), entretanto timbres distintos. A existência de exceções também foi observada em relação a vogal tônica <o>, como em “gar[o]to” e “f[ɔ]to”.

Surgem ainda problemas de transcrição de vogais médias presentes em posição pretônica de formas derivadas com os sufixos “-mente, -inh, -zinh, -íssim”, como ocorre, por exemplo, com os derivados da palavra <mole>: “m[ɔ]lemente”, “m[ɔ]linho”, “m[ɔ]lezinho” e “m[ɔ]líssimo”, que não são corretamente transcritos no estado atual do sistema PETRUS. Para esses casos prevemos a necessidade de criação de regras específicas para transcreever adequadamente o timbre dessas vogais e que, por isso, merecerá atenção em trabalho futuro.

Esses são resultados preliminares que podem ser melhorados com a inclusão de mais regras de transcrição para tratamento específico de casos problemáticos.

No próximo capítulo, apresentamos considerações sobre o trabalho desenvolvido e indicação das próximas atividades a serem realizadas.

CONCLUSÕES E ENCAMINHAMENTOS FUTUROS

4.1 Considerações sobre o trabalho desenvolvido

O presente trabalho procurou, em primeiro lugar, apresentar o problema da falta de transcrição fonética em dicionários do PB e as dificuldades envolvidas em uma tarefa como essa. Desse modo, mostramos que um conversor grafema-fone, em geral, desenvolvido para fazer parte de sistema de síntese de voz, poderia ser construído para atender às necessidades de lexicógrafos que desejam adicionar transcrição fonética nas obras lexicográficas que produzem para o PB.

A seguir, realizamos um levantamento bibliográfico na literatura sobre as abordagens existentes para a conversão de um sistema conversor, tanto em PE quanto em PB, para escolher a metodologia que promovesse os melhores resultados, no caso, a baseada em regras linguísticas.

Em virtude da já citada regularidade ortográfica, têm sido desenvolvidos sistemas de conversão grafema-fone(ma) predominantemente com o uso de regras linguísticas para realizar transcrição, muito embora tenham sido propostas novas abordagens para resolver a questão da conversão grafema-fone.

Na primeira fase de nosso trabalho, optamos pelas abordagens mais populares da conversão grafema-fone em PB. Referimo-nos, naturalmente, à aplicação de regras, que no nosso caso específico, foram combinadas com o uso de um pequeno dicionário de exceções para a transcrição de HHs. Atualmente, a lista de exceções não é muito extensa, mas admitimos a possibilidade de adicionar novas unidades lexicais.

De acordo com Braga (2008), o uso de regras linguísticas é, sem dúvida, a melhor estratégia para resolver o problema da conversão grafema-fone, pois se por um lado, é mais econômica em termos de memória computacional, permite, concomitantemente, atingir

desempenhos superiores aos obtidos por métodos automáticos. Essa mesma técnica é também bastante eficaz para lidar com palavras novas, que vão sendo permanentemente associadas ao léxico (TAYLOR, 2005). A utilização de regras linguísticas geradas manualmente foi a estratégia que viabilizou o desenvolvimento do primeiro sistema de conversão grafema-fone para a variedade da Rede Globo, denominado PETRUS, dando assim, seguimento ao objetivo principal enunciado na Introdução deste trabalho.

Em seguida, fizemos uma revisão da literatura em relação aos requisitos necessários para a confecção das regras de transcrição a serem inseridas no sistema aqui proposto. Tal revisão englobou um percurso em temas específicos das áreas de fonética, fonologia e prosódia, os quais tiveram importância específica para o delineamento das regras construídas e das etapas de funcionamento do sistema. A função de cada uma das etapas do processo de conversão grafema-fone foram descritas com detalhes no Capítulo 2.

O sistema implementado, denominado PETRUS, encontra-se disponível gratuitamente para uso, no portal do NILC. O referido sistema é fruto não apenas do trabalho de doutorado ora relatado, mas também da colaboração de um mestrando em Ciências da Computação, também pertencente ao NILC, que ficou responsável pela implementação computacional do sistema.

Entre os dias 6 e 9 de outubro de 2014, participamos do XI PROPOR – *International Conference on Computational Processing of Portuguese Language*, o principal evento na área de Processamento de Linguagem Natural para o Português. Em um de seus eventos satélites, o *Software Demonstration Workshop*, apresentamos o PETRUS e recebemos o prêmio *Best software demonstration*, o que muito nos gratificou.

Atualmente, sua etapa de transcrição ortográfica-fonética está sendo realizada com base em cinco módulos (etapas) distintos: transcritor de morfemas localizados à esquerda de sua base, transcritor de homógrafos heterófonos, divisor silábico, marcador de sílaba tônica e o transcritor, responsável pela aplicação de regras de transcrição. Essa configuração atual do PETRUS permite obter transcrições fonéticas, com base na pronúncia adotada, a uma taxa de acerto de 97.5% ao fone. Entretanto, vislumbramos uma possibilidade de aumento dessa taxa a partir de uma reformulação das regras problemáticas existentes, considerando o tipo de erro que têm provocado, e da adição de regras específicas para a determinação do timbre correto das vogais <e> e <o> em posição tônica de nomes.

Por fim, podemos dizer que a implementação das regras do sistema de conversão grafema-fone apresentado, mais do que representar o fechamento do primeiro ciclo deste trabalho, abre caminhos para a realização de novas tarefas, que darão continuidade ao

trabalho até aqui realizado. Apontaremos na próxima seção algumas sugestões de atividades futuras que podem contribuir para um ajuste fino da qualidade da saída produzida atualmente pelo sistema PETRUS.

4.2 Encaminhamentos Futuros

Como trabalho futuro, pretendemos aumentar o contexto de sua atuação, com a elaboração e implementação de regras relacionadas com a coarticulação entre palavras.

O PETRUS trata apenas dos aspectos relacionados com a fonologia lexical, ou seja, ele se dedica à transcrição de unidades lexicais simples. No entanto, uma obra lexicográfica pode não ser composta apenas por unidades lexicais simples, mas também por combinações de palavras que funcionam como um item lexical único. Tal agrupamento de palavras compreende as expressões idiomáticas (ex: “agarrar com unhas e dentes”), as locuções (ex: “estrada de ferro”), os provérbios (ex: “A necessidade é a mãe das invenções”), binômios (ex: “espinha-santa”), colocações (ex: “ódio mortal”), entre outros, considerados exemplos de um mesmo fenômeno: unidades lexicais maiores que uma palavra.

Tais unidades podem apresentar características fonéticas alteradas em função das palavras que lhes são adjacentes, derivados de fenômenos de coarticulação possíveis entre palavras. Por exemplo, o fone [s] é comum em final de palavras, contudo, se sonoriza quando o primeiro fone da palavra seguinte for uma vogal (ex: “velho[za]migos”). Tal fenômeno de coarticulação é denominado de *sonorização da fricativa*. O fenômeno de *degeminação* inclui casos de sequência de vogais idênticas (ex: “velh[o]fício”) e também constitui um caso de ajuste pós-lexical. Há, ainda, os casos de *elisão*, em que há a eliminação de um fone (ex: “tequil[o]uro”), de *tapping*, processo de reestruturação da sílaba em que o segmento final passa a inicial da primeira sílaba do segundo vocábulo (ex: “arar[i]dosa”), de ditongação (ex: “velh[oa]migo”) e de *haplologia*, ou seja, de elisão de uma das sílabas quando há duas semelhantes (ex: “ida[dzi]vertida”).

A Fonologia Prosódica aborda esses fenômenos que ocorrem nas palavras quando elas estão inseridas em um contexto mais amplo. É preciso, portanto, que o sistema de conversão contenha mais um módulo, posterior ao de conversão grafema-fone, capaz de lidar com esses ajustes. A versão atual do PETRUS não faz nenhum tipo de tratamento pós-lexical.

Portanto, apresentamos a proposta de um trabalho a ser realizado, no sentido de implementar um novo módulo, que será incorporado ao nosso sistema. O objetivo desse novo módulo é exatamente o de prover um tratamento pós-lexical à saída do PETRUS.

As próximas atividades estão relacionadas, de modo geral, à análise dos erros gerados visando reelaboração das regras de transcrição existentes, bem como a implementação de novas regras relacionadas com a determinação do timbre tônico das vogais <e> e <o> em nomes, com vistas à melhoria do desempenho da transcrição fonética gerada pelo PETRUS atualmente.

Problema de solução não trivial em PB é a transcrição fonética automática das vogais <e> e <o> que podem ter pronúncia variada (aberta, fechada, reduzida) conforme o contexto linguístico em que estão inseridas. Diante desse cenário, temos como objetivo determinar as circunstâncias fonéticas (vogais e consoantes adjacentes à tônica em estudo) e possíveis fatores linguísticos (classe gramatical) que condicionam a alternância vocálica de tais vogais em formas nominais do PB. Por meio de um estudo como o que se propõe, será possível identificar os contextos linguísticos em que tais vogais são pronunciadas com timbre aberto na referida classe gramatical, criar regras de transcrição que especifiquem tais contextos, implementá-las no PETRUS e verificar seu desempenho.

A dificuldade que os atuais sistemas computacionais que realizam conversão grafia-fone(ma) em PB têm ao converter as vogais médias em posição tônica e a ausência de estatísticas com base em córpus extensos para o estudo da abertura/fechamento das vogais médias em posição tônica de nomes e verbos do PB nos fazem considerar ser ainda necessário haver novos estudos, como o que se propõe.

Como vimos, as pesquisas fonológicas dedicadas ao PB reconhecem um sistema vocálico constituído por sete vogais /a, i, u, e, ε, o, ɔ/ em posição tônica. Os fonemas /a, i, u/ apresentam uma relação grafia-som inequívoca, pois todo <i> grafado, em posição tônica de palavra, será sempre pronunciado como /i/, (ex.: bronqu/i/te) , assim como todo <a> será sempre /a/ (mac/a/co) e todo <u> será sempre /u/ (mal/u/co). Entretanto, a relação grafia-som não se apresenta de maneira tão objetiva para os outros quatro fonemas /e, ε, o, ɔ/. Essa falta de relação unívoca entre grafia e respectivos sons “se deve às tensões provenientes entre o dinamismo da evolução fonética de todas as línguas naturais do mundo e a natureza conservativa típica de todas as suas respectivas ortografias” (BRAGA, 2008). Esse não-isomorfismo entre grafemas e seus respectivos fone(ma)s em PB causa dificuldades nas situações em que se torna necessário convertê-las da forma gráfica para a sonora. Para ilustrar, podemos citar casos em que <e, o> podem ter pronúncia fechada, como nos

substantivos <p[o]rto> e <p[e]rco>, mas abertas em <p[ɔ]rta> e <f[ɛ]sta>.

Diversos estudos têm sido realizados acerca das vogais médias do PB. De fato, essas vogais têm sido amplamente estudadas em posição pretônica, segundo diferentes linhas teóricas (Câmara Jr. (1970), Leite (1974), Bisol (1981), Quicoli (1990) Viegas (1987), Wetzels (1995), entre outros). Em Lee e Oliveira (2003), por exemplo, as realizações das vogais médias em posição pretônica são consideradas um fator determinante para a diferenciação de dialetos brasileiros. Entretanto, essa diversidade não se dá apenas em posição pretônica. Há estudos que consideram a variação das vogais médias em posições postônica final (Cf. Silva, 2009) e não-final (Cf. Brandão; De Paula, 2007) no PB.

Em posição tônica, entretanto, constata-se um número menor de trabalhos. Dentre eles, podemos citar os de Vieira (1990), Cunha (1991), Wetzels (1991), Cagliari (1997) e Alves (1999). De um modo geral, essas investigações têm em comum o fato de descreverem o comportamento das vogais médias tônicas em formais nominais, dedicando-se ao estudo da alternância do timbre das vogais médias em nomes no masculino singular e nas formas de feminino e plural.

Contudo, não se observa, até o presente momento, um estudo voltado para a realização de levantamento estatístico sobre a ocorrência de vogais médias em posição tônica, considerando um conjunto maior de nomes, adjetivos ou verbos, por exemplo. Uma vez que certos padrões existem e podem dar suporte para a predição da qualidade vocálica em determinados contextos, como por exemplo, o de sílaba tônica em qualquer uma das categorias gramaticais existentes, percebe-se que é relevante investigar a alternância entre as vogais médias.

Portanto, com a presente proposta de investigação, avançamos em relação aos trabalhos anteriores que se propuseram à análise do fenômeno com um pequeno conjunto de dados e não previram uma generalização maior do estudo, cujo conhecimento linguístico então produzido pudesse ser aproveitado na área de Tecnologias da Fala, mais especificamente, na construção de regras para sistema de conversão grafema-fone dedicados ao PB.

Além desses, apontamos de forma sucinta os seguintes temas que podem ser explorados com maiores detalhes, com vistas à melhoria dos resultados de transcrição do sistema desenvolvido:

- Investigar os contextos linguísticos de ocorrência (ou não) de harmonização vocálica em nomes e verbos do PB, pronunciados conforme o dialeto da Rede Globo, com base em

cópus representativo dessa variedade, a fim de criar regras que possam ser implementadas no PETRUS.

- Investigar os contextos linguísticos em que as vogais <e> e <o> postônicas não-finais são reduzidas, visando a elaboração de regras de transcrição para o nosso sistema.
- Investigar a oscilação existente entre a pronúncia aberta (*ideia, boleia*) e a fechada (*aldeia, baleia*) para especificar no formato de regras de transcrição os casos em que a sequência gráfica <ei> deve ser transcrita foneticamente como [ɛi] ou como [ei].
- Implementar regras para transcrição fonética das vogais médias pretônicas em contextos nos quais o alçamento dessas vogais é, de acordo com Bisol (1981), categórico. São eles: o contexto de hiato (“b[u]emia”), prefixo (“d[i]screve”) e da vogal <e> em início de palavra sucedida por <s> ou <x> (“[i]screve”; “[i]xclui”). Além deles, investigar outros possíveis contextos linguísticos que possam alterar sua pronúncia, no dialeto aqui adotado.
- Habilitar a opção de realizar divisão silábica da sequência glide + vogal tanto como ditongo crescente quanto como hiato. Assim, o sistema ganharia em flexibilidade e o usuário do sistema poderia optar pelo forma que fosse mais adequada aos seus interesses.
- Elaborar um dicionário fonético com palavras frequentes do PB contendo a consoante <x>, visando a elaboração de regras mais refinadas para a sua transcrição fonética automática, bem como de um dicionário de exceções para os casos mais específicos.
- Elaborar um conjunto de regras para atribuição de acento secundário em palavras.
- Desenvolver um sintetizador que pronunciará as palavras transcritas pelo PETRUS. O áudio pode ser um recurso interessante para aprendizes de língua, independentemente de sua idade. Disponibilizando-o para *download*, o referido sintetizador poderia ser aproveitado para sintetizar palavras em dicionários *online*.
- Adaptar o sistema de transcrição atual da pronúncia da Rede Globo para outras variantes do PB, bem como para outras línguas.

REFERÊNCIAS

- ABAURRE-GNERRE, Maria Bernadete Marques. Alguns casos de formação de plural em português: uma abordagem natural. **Cadernos de Estudos Linguísticos**, Campinas, n.5, p.127-156, 1983.
- ALBANO, Eleonora Cavalcante; MOREIRA, Agnaldo Antônio. Archsegment-based letter-to-phone conversion for concatenative speech synthesis in portuguese. In: INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 1996, Philadelphia. **Proceedings...**, Philadelphia, 1996. p.1708-1711. Disponível em: <<http://www.asel.udel.edu/icslp/cdrom/vol3/612/a612.pdf>>. Acesso em: 20 out. 2014.
- ALLEN, Jonathan; HUNNICUT, Sharon; KLATT, Dennis. **From text to speech: the MITALK system**. Cambridge: Cambridge University Press, 1987.
- ALMEIDA, Antonio. Portuguese nasal vowels: phonetics and phonemics. In: SCHMIDT-RADEFELT, Juergen (Org.). **Readings in portuguese linguistics**. New York: North Holland, 1976. p. 348-396.
- ALMEIDA, Leonardo Silva de. **Um estudo sobre síntese de fala para o português brasileiro**. 2005. 119f. Dissertação (Mestrado em Engenharia Elétrica)–Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.
- ALVES, Marlúcia Maria. **As vogais médias em posição tônica nos nomes do português brasileiro**. 1999. 136f. Dissertação (Mestrado em Estudos Linguísticos)-Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 1999.
- BACK, Eurico. São fonemas as vogais nasais do português? **Construtura**, São Paulo, n.4, p. 297-317, 1973.
- BAGNO, Marcos. **Gramática pedagógica do português brasileiro**. São Paulo: Parábola, 2011.
- BARBOSA, Filipe et al. Grapheme-phone transcription algorithm for a brazilian portuguese TTS. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE (PROPOR'03), 6, 2003, Faro. **Proceedings...** Berlin, Heidelberg: Springer-Verlag, 2003. p.23-30.
- BARBOSA, Maria Aparecida. Dicionário, vocabulário, glossário: concepções. In: ALVES, Ieda Maria (Org.). **A constituição da normalização terminológica no Brasil**. 2 ed. São Paulo: FFLCH/CITRAT, 2001.
- _____. Dicionário, vocabulário, glossário: concepções. In: ALVES, Ieda Maria (Org.). **A constituição da normalização terminológica no Brasil**. São Paulo: FFLCH/CITRAT, 1996. p.23-45. (*Cadernos de Terminologia*, v. 1).

BARBOSA, Plínio Almeida. Do grau de não perifericidade da vogal /a/ pós-tônica final. **Revista Diadorim: Revista de Estudos Linguísticos e Literários do Programa de Pós-Graduação em Letras Vernáculas da Universidade Federal do Rio de Janeiro**. v. 12, 2012. Disponível em: <<http://www.revistadiadorim.letras.ufrj.br>>. Acesso em: 12 dez. 2014.

_____.; ALBANO, Eleonora. Brazilian portuguese. illustrations of the IPA. **Journal of the International Phonetic Association**, v. 2, n.34, p. 227-232, 2004.

_____. et al. Aiuruete: a high-quality concatenative text-to-speech system for brazilian portuguese with demisyllabic analysis-based units and a hierarchical model of rhythm production. In: EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY (EUROSPEECH 99), 6., 1999, Budapest. **Proceedings...**, Budapest. 1999. p. 259-262.

BARROS, Maria João; WEISS, Christian. Maximum entropy motivated grapheme-to-phoneme, stress and syllable boundary prediction for Portuguese text-to-speech. In: JORNADAS EN TECNOLOGIA DEL HABLA, 4., 2006. Zaragoza. **Proceedings...**, Zaragoza. 2006. p.177-182. Disponível em: <http://lorien.die.upm.es/~lapiz/rtth/JORNADAS/IV/finals/4jth_127.pdf>. Acesso em: 20 out. 2014.

BARTLETT, Susan; KONDRAK, Grzegorz; CHERRY, Colin. Automatic syllabification with structured SVMs for letter-to-phoneme conversion. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS: HUMAN LANGUAGE TECHNOLOGIES, 2008. Columbus. **Proceedings...**, Columbus. Association for Computational Linguistics, 2008. p.568-576.

BATTISTI, Elisa. **Elevação das vogais médias pretônicas em sílaba inicial de vocábulo na fala gaúcha**. 1993. 125 f. Dissertação (Mestrado em Letras)- Universidade Federal do Rio Grande do Sul, Porto Alegre-RS, 1993.

BECHARA, Evanildo. Divisão de sílabas em fonética e em ortografia. **O Dia**, Rio de Janeiro, 15 maio 2011. Disponível em: <<http://www.academia.org.br/abl/cgi/cgilua.exe/sys/start.htm?infoid=11762&sid=78>>. Acesso em: 20 out. 2014.

_____. **Moderna gramática portuguesa**: revista, ampliada e atualizada conforme o novo Acordo Ortográfico. Rio de Janeiro: Nova Fronteira, 2009.

BERINGER, Nicole. How to integrate phonetic and linguistic knowledge in a text-to-phoneme conversion task: a syllabic TPC tool for French. In: INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 8., 2004. Jeju Island. **Proceedings...**, 2004. p. 1941-1944.

BIDERMAN, Maria Tereza Camargo. A ciência da lexicografia. **Alfa: Revista de Linguística**, São Paulo, v.28, n.1, p.1-26, 1984.

BINNENPOORTE, Diana Maria. **Phonetic transcriptions of large speech corpora**. 2006. 169f. Dissertation (PhD Thesis), Radboud University Nijmegen, the Netherlands, 2006. Disponível em: <<http://repository.ubn.ru.nl/handle/2066/27415>>. Acesso em: 20 out. 2014.

BISANI, Maximilian; NEY, Hermann. Joint-sequence models for grapheme-to-phoneme conversion. **Journal of Speech Communication**, Amsterdam, v.50, n.5, p.434-451, 2008.

BISOL, Leda. **Introdução a estudos de fonologia do português brasileiro**. Porto Alegre: EDIPUCRS, 1996.

_____. O acento e o pé métrico binário. **Cadernos de Estudos Linguísticos**, Campinas, n.22, p.69-80, 1992.

_____. O ditongo na perspectiva da fonologia atual. **D.E.L.T.A**, São Paulo, v.5, n.2, p.185-224, 1989.

_____. **Harmonização vocálica**. 1981. 332f. Tese (Doutorado em Estudos Linguísticos) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1981.

BIZZOCCHI, Aldo. O be-a-bá da sílaba. **Língua Portuguesa**, ano 7, n.78, abr. 2012. Disponível em: <<http://www.aldobizzocchi.com.br/divulgacao.asp>>. Acesso em: 20 out. 2014.

BORBA, Francisco da Silva. **Dicionário de usos do português do Brasil**. São Paulo: Ática, 2002.

_____. **Introdução aos estudos linguísticos**. São Paulo: Editora Nacional, 1970.

BORTONI, Stela; GOMES, Christina; MALVAR, Elisabete. A variação das vogais médias pretônicas no português de Brasília: um fenômeno neogramático ou de difusão lexical? **Revista Estudos da Linguagem**, Belo Horizonte, v.1, n.1, p.9-30, 1992.

BRAGA, Daniela. **Algoritmos de processamento da linguagem natural para sistemas de conversão texto fala em português**. 2008. 228f. Tese (Doutorado em Síntese de Fala)- Departamento de Galego-Português, Francés e Linguística, Universidade da Coruña, A Coruña, 2008.

_____; FREITAS, Diamantino; FERREIRA, Helder. Processamento linguístico aplicado à síntese da fala. In: 3º CONGRESSO LUSO-MOÇAMBICANO DE ENGENHARIA, 2003. Maputo. Anais..., Maputo, 2003. Disponível em: <<http://hfilipe.no.sapo.pt/publicacoes/CLME2003.pdf>>. Acesso em: 20 out. 2014.

_____; MARQUES, Maria Aldina. Desambiguação de homógrafos para sistemas de conversão texto-fala em português. **Diacrítica**, v.21, n.1, p.25-50, 2007. Disponível em: <http://ceh.ilch.uminho.pt/publicacoes/Diacritica_21-1.pdf>. Acesso em: 20 out. 2014.

_____; RESENDE JÚNIOR, Fernando Gil Vianna. Módulos de processamento de texto baseados em regras para sistemas de conversão texto-fala em português europeu. In: ENCONTRO DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA. 21., Porto, 2007. **Anais...**, Lisboa, APL, 2007, p.141-156.

BRANDÃO, Silvia Figueiredo; DE PAULA, Alessandra. O comportamento das vogais médias postônicas não finais na fala fluminense. In: SIMPÓSIO SOBRE AS VOGAIS (SIS-VOGAIS), 2007, João Pessoa. **Anais...**, João Pessoa: Universidade Federal da Paraíba, 2007. p.15-17.

CAGLIARI, Luiz Carlos. Um modelo de transcrição fonética para um dicionário. In: PETROV, Petar et al. (Org.). **Avanços em ciências da linguagem**. Santiago de Compostela: Através, 2012. p. 475-490.

_____. Integrating phonetic transcription in a Brazilian Portuguese dictionary. **Kernerman Dictionary News**, Tel Aviv, v.19, p.1-6, 2011.

_____. **Elementos de fonética do português brasileiro**. São Paulo: Paulistana, 2009.

_____. **Análise fonológica**: introdução à teoria e à prática com especial destaque para o modelo fonêmico. Campinas: Mercado de Letras, 2002.

_____. A ortografia na escola e na vida. In: CAGLIARI, Luiz Carlos; MASSINI-CAGLIARI, Gladis. **Diante das letras**: a escrita na alfabetização. Campinas: Mercado das Letras, 1999a.

_____. **Acento em português**: estudos sobre as regras de atribuição de acento em português. Campinas: L. C. Cagliari, 1999b.

_____. **Elementos de fonética do português brasileiro**. 1981. 194 f. Tese (Livre-Docência em Linguística)–Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1981.

_____. **An experimental study of nasality with particular reference to brazilian portuguese**. 1977. 319f. Tese (Doutorado em Filosofia)- University of Edinburgh, Edinburgh, Scotland, 1977.

_____.; MASSINI-CAGLIARI, Gladis. A epêntese consonantal em português e sua interpretação na teoria da otimalidade. **Revista de Estudos da Linguagem**, Belo Horizonte, v.9, n.1, p.163-192, 2000.

CALLOU, Dinah; LEITE, Yonne. **Iniciação à fonética e à fonologia**. Rio de Janeiro: Zahar, 2003.

_____. Variação das vogais pretônicas. In: SIMPÓSIO-DIVERSIDADE LINGUÍSTICA NO BRASIL, 1., 1986, Salvador. **Atas...**, Salvador: UFBA, 1986. p. 157-169.

CALLOU, Dinah et al. Variação e diferenciação dialetal: a pronúncia do /R/ no português do Brasil. In: KOCH, Ingedore Villaça (Org.). **Gramática do português falado**, desenvolvimentos. Campinas: Ed. da UNICAMP; 1997.v. 6.

CÂMARA JUNIOR, Joaquim Mattoso. **História da linguística**, 4. ed. Petrópolis: Vozes, 1986.

_____. **Estrutura da língua portuguesa**. Petrópolis: Vozes, 1972.

CÂMARA JUNIOR, Joaquim Mattoso. **Estrutura da língua portuguesa**. Petrópolis: Vozes, 1970.

_____. **Problemas de linguística descritiva**. Petrópolis: Vozes, 1969.

_____. **Dicionário de filologia e gramática**: referente à língua portuguesa. 2.ed. ref. Rio de Janeiro: J. Ozon, 1964.

_____. **Para o estudo da fonêmica portuguesa**. Rio de Janeiro: Organização Simões, 1953.

CANDEIAS, Sara; VEIGA, Arlindo; PERDIGÃO, Fernando. *Sistema automático de pronúncia de verbos*. In: III JORNADA DE DESCRIÇÃO DO PORTUGUÊS, 3., 2013, Fortaleza. **Jornada de descrição do português**. Fortaleza. 2013, p. 28-35.

CARDOSO, Suzana Alice Marcelino. Projeto Atlas Linguístico do Brasil-Projeto ALiB: descrição e estágio atual. **Revista da ABRALIN**, Belém-PA, v. 8, n.1, p.185-198, 2009. Disponível em: <<http://www.abralin.org/site/data/uploads/revistas/2009-vol-8-n-1/suzanaalice.pdf>>. Acesso em: 20 out. 2014.

CARMO, Márcia Cristina do. **As vogais médias pretônicas dos verbos na fala culta do interior paulista**. 2009. 119f. Dissertação (Mestrado em Estudos Linguísticos)-Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto-SP, 2009.

CASEIRO, Diamantino António. **Finite-state methods in automatic speech recognition**. 2003. 199 f. Tese (Doutorado em Engenharia Informática e Computadores) - Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisboa, 2003.

CASEIRO, Diamantino António et al. Grapheme-to-phone using finite-state transducers. In: WORKSHOP ON SPEECH SYNTHESIS, 1., 2002, Santa Monica. **Proceedings...**, 2002. Santa Monica: Institute of Electrical and Electronics Engineers, 2002. p.215-218.

CASTELEIRO, João Malaca (Coord.). **Dicionário da língua portuguesa contemporânea**. Lisboa: Academia das Ciências de Lisboa: Editorial Verbo, 2001. 2v.

CASTILHO, Ataliba Teixeira de. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2012.

_____. Variação dialetal e ensino institucionalizado da língua portuguesa. **Cadernos de Estudos Linguísticos**, n. 1, Campinas, 1978, p. 13-20.

CAVALIERE, Ricardo. **Pontos essenciais em fonética e fonologia**. Rio de Janeiro: Lucerna, 2005.

CELIA, Gianni Fontis. **Variação das vogais médias pretônicas no português de Nova Venécia-ES**. 2004. 113f. Dissertação (Mestrado em Estudos Linguísticos)-Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2004.

CHBANE, Dimas Trevizan. **Desenvolvimento de sistema para conversão de textos em fonemas no idioma português**. 1994. 124f. Dissertação (Mestrado em Engenharia)- Escola Politécnica, Universidade de São Paulo, 1994.

CHOMSKY, Noam; HALLE, Morris. **The sound pattern of english**. New York: Harper and Row, 1968.

_____. **Aspects of the theory of syntax**. Cambridge: MIT press, 1965.

CLEMENTS, George N. *Vowel height assimilation in bantu languages*. In: ANNUAL MEETING OF THE BERKELEY LINGUISTICS SOCIETY, 2., Berkeley. **7th Annual Meeting of the Berkeley Linguistics Society: Special Session on African Language Structures**, Berkeley: eLanguage, 1991. p. 25-64. Disponível em: <<http://elanguage.net/journals/bls/article/view/2804>>. Acesso em: 18 dez. 2014.

COLLISCHONN, Gisela. A epêntese vocálica no português do sul do Brasil. In: BISOL, Leda; BRESCANCINI, C. (Org.). **Fonologia e variação: recortes do português brasileiro (VARSUL)**. Porto Alegre: EDIPUCRS, 2002. p.205-230.

_____. Um estudo da epêntese à luz da teoria da sílaba de Junko Ito (1986). **Letras de Hoje**, Porto Alegre, v.31, n.2, p.149-158, 1996.

CORRÊA, Leda; JESUS, Manuela Oliveira de; ANJOS, Daniele Oliveira dos. Pronúncia do português brasileiro e sua importância na elaboração de dicionário de português para estrangeiros. **Interdisciplinar: Revista de Estudos em Língua e Literatura**, Itabaiana, ano 5, v.12, p.49-61, 2010.

CORREIA, Margarita. **Os dicionários portugueses**. Lisboa: Editorial Caminho, 2009.

_____. A escola e norma padrão do português. Expresso; caderno principal, Lisboa, 8 set. 2001. p.23.

COSTA, Iara Bemquerer. **O acento em português: estudo de algumas mudanças no modelo da fonologia gerativa**. 1978. 182 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade de Campinas, Campinas, 1978.

COUTO, Igor Costa do. **Desenvolvimento de recursos para a construção de um sistema texto-fala para o português brasileiro**. 2010. 88 f. Dissertação (Mestrado em Engenharia Elétrica)-Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2010.

CRISTÓFARO-SILVA, Thaïs. **Fonética e fonologia dos portugueses: roteiro de estudos e guia de exercícios**. 3.ed. São Paulo: Contexto, 2000.

_____. O método das vogais cardeais e as vogais do português brasileiro. **Revista de Estudos Linguísticos**, Belo Horizonte, v.8, n.2, p.128-153, 1999.

_____. **Fonética e fonologia dos portugueses: roteiro de estudos e guia de exercícios**. 3.ed. São Paulo: Contexto, 1998.

CRISTÓFARO-SILVA, Thaís; CAMPOS, Carlo Alessandro de Oliveira. Abertura vocálica em verbos irregulares do português. *Alfa*, São Paulo, n. 52, v. 2, p. 371-390, 2008.

_____.; YEHIA, Hani Camille. **Sonoridade em artes, saúde e tecnologia**. Belo Horizonte: Faculdade de Letras, 2009.

CRYSTAL, David. **A dictionary of linguistics and phonetics**. New York: John Wiley & Sons, 2008.

_____. **The Cambridge encyclopedia of the english language**. Cambridge: Ernst Klett Sprachen, 1997.

CUCCHIARINI, Catia. **Phonetic transcription: a methodological and empirical study**. 1993. 213 f. Tese (Doutorado em Linguística). University of Nijmegen, the Netherlands, 1993.

CUNHA, Celso; CINTRA, Luís F. Lindley. **Nova gramática do português contemporâneo**. Lisboa: Sá da Costa, 2002.

CUNHA, Celso; CINTRA, Luís F. Lindley. **Nova gramática do português contemporâneo**. Lisboa: Sá da Costa, 1984.

CUNHA, Viviane. **Um traço do vocalismo português: a metafoia**. 1991. 154f. Dissertação (Mestrado em Língua Portuguesa)-Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 1991.

DAMPER, Robert I. et al. Aligning letters and phonemes for speech synthesis. In: ISCA SPEECH SYNTHESIS WORKSHOP, 5., 2004, Pittsburgh. **Proceedings...**, Pittsburgh, 2004, p.209-214. Disponível em: <http://www.isca-speech.org/archive_open/ssw5/ssw5_209.html>. Acesso em: 21 out. 2014.

DESHPANDE, Madhav M. Phonetics, ancient Indian. In: ASHER, Ronald; SIMPSON, James M. Y. (Ed.). **The encyclopedia of language and linguistics**. Oxford: Pergamon,-1994. v.6, p.3053-3058.

DIVAY, Michel; VITALE Anthony. Algorithms for grapheme-phoneme translation for english and french: applications for database searches and speech synthesis. **Journal of Computational Linguistics**, Cambridge, v.23, n.4, p.495-523, 1997.

DONG, Honghui; TAO, Jianhua; XU, Bo. *Grapheme-to-phoneme conversion in chinese TTS system*. In: INTERNATIONAL SYMPOSIUM ON CHINESE SPOKEN LANGUAGE PROCESSING, 1., 2004, Hong Kong. **4th International Symposium on Chinese Spoken Language Processing**. Hong Kong: INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, 2004. p. 165-168.

DUBOIS, Jean et al. **Dicionário de linguística**. São Paulo: Cultrix, 1993.

DURAN, Magali Sanches; XATARA, Claudia Maria. Lexicografia pedagógica: atores e interfaces. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, São Paulo, v.23, n.2, p.203-222, 2007.

DUTOIT, Thierry. **An introduction to text-to-speech synthesis**. Dordrecht: Kluwer Academic Publishers, 2001.

_____. *High quality text-to-speech synthesis: a comparison of four candidate algorithms*. In: INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 1994, Adelaide. **International Conference on Acoustics, Speech and Signal Processing**. Adelaide: INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, 1994. p.565-568.

FERRARI, Lilian ; BARBOSA, Filipe ; RESENDE JÚNIOR, Fernando. *Construções gramaticais e sistemas de conversão texto-fala: o caso dos homógrafos*. In: INTERNATIONAL CONFERENCE ON COGNITIVE LINGUISTICS, 8, 2003, Braga. **International Conference on Cognitive Linguistics**. Logroño: La Rioja, 2003.

FERREIRA, Aurélio B. H. **Novo dicionário eletrônico Aurélio**. versão 5.0. Curitiba: Positivo, 2004. 1 CD-ROM.

_____.; ANJOS, Margarida dos; FERREIRA, Marina B. (Ed.) **Aurélio século XXI**. 3.ed. Rio de Janeiro: Nova Fronteira, 2000.

FONSECA, Erick Rosa; ROSA, João Luís Garcia. *A two-step convolutional neural network approach for semantic role labeling*. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 1., 2013, Dallas. **International Joint Conference on Neural Networks**. Dallas: INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, 2013. p. 2955-2961.

FRANZEN, Evandro; BARONE, Dante Augusto Couto. *Descoberta automática de regras para conversão de letras para fonema em processos de síntese de texto para fala*. In: CONGRESSO BRASILEIRO DE COMPUTAÇÃO, 2002, Itajaí. **Congresso Brasileiro de Computação**. Itajaí, 2002.

FREITAS, Simone. **As vogais médias pretônicas no falar da cidade de Bragança**. 2001. 125f. Dissertação (Mestrado em Linguística)-Centro de Letras e Artes, Universidade Federal do Pará, Belém, 2001.

GALVÃO, Maria João Chaves. **Root vowel alternation in iberian portuguese: a feature geometry approach**. 1993. 152f. Dissertação (Master's Thesis), University of Washington, Washington, 1993.

GIMSON, Alan Cruttenden. Pronunciation in EFL dictionaries. **Applied Linguistics**, Oxford, v.2, n.3, p.250-262, 1981.

GOMES, Leandro. **Sistema de conversão texto-fala para a língua portuguesa utilizando a abordagem de síntese por regras**. 1998. 114f. Dissertação (Mestrado em Engenharia Elétrica e de Computação)-Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 1998.

GOUVEIA, Paulo; TEIXEIRA, João Paulo; FREITAS, Diamantino. *Divisão silábica automática do texto escrito e falado*. In: ENCONTRO PARA O PROCESSAMENTO COMPUTACIONAL DA LÍNGUA PORTUGUESA ESCRITA E FALADA, 2000, Atibaia. **Encontro para o Processamento Computacional da Língua Portuguesa Escrita e Falada**. Atibaia, 2000. p.65-74.

GREENBERG, Steven. Speaking in shorthand: a syllable-centric perspective for understanding pronunciation variation. **Speech Communication**, v.29, n.2, p.59-176, 1999.

HAENSCH, Günther. **La lexicografía de la lingüística teórica a la lexicografía práctica**. Madrid: Gredos, 1982.

HARRIS, James W. Evidence from portuguese for the 'elsewhere condition'. **Linguistic Inquiry**, Massachussets, v. 5, n. 1, p. 61-80, 1974.

HARTMANN, Reinhard Rudolf Karl. **Teaching and researching lexicography**. London: Longman, 2001.

_____.; JAMES, Gregory. **Dictionary of lexicography**. California: Psychology Press, 1998.

HEAD, Brian Franklin. *A comparison of the segmental phonology of Lisbon and Rio de Janeiro*. 1964. Tese (Doutorado em Linguística), University of Texas, Austin, 1964.

HOSN, Chadia Nadim Aboul. **Conversão grafema-fone para um sistema de reconhecimento de voz com suporte a grandes vocabulários para o português brasileiro**. 2006. 90f. Dissertação (Mestrado em Engenharia Elétrica)-Centro Tecnológico, Universidade Federal do Pará, Belém, 2006.

HOUAISS, Antonio; VILLAR, Mauro Salles; FRANCO, Francisco Manoel de Mello (Ed.). **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2004.

HOWARD, Sara J.; HESELWOOD, Barry C. Learning and teaching phonetic transcription for clinical purposes. **Clinical Linguistics and Phonetics**, v.16, n.5, p.371-401, 2002.

INTERNATIONAL PHONETIC ASSOCIATION. **Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet**. 8th. ed. Cambridge: Cambridge University Press, 2007.

IRIARTE SANROMÁN, Álvaro. **A unidade lexicográfica: palavras, colocações, frasesmas, pragmatemas**. 2000. 393f. Dissertação. (Doutoramento em Ciências da Linguagem), Universidade do Minho, Braga, 2000. Disponível em: <http://repositorium.sdum.uminho.pt/bitstream/1822/4573/1/A_Unidade_Lexicografica.pdf>. Acesso em: 14 jul. 2014.

KATO, Mary Aizawa. **No mundo da escrita: uma perspectiva psicolinguística**. 7.ed. São Paulo: Ática, 2002.

KESSENS, Judith; CUCCHIARINI, Catia; STRIK, Helmer. A data-driven method for modeling pronunciation variation. **Speech Communication**, v.40, n.4, p.17-534, 2003.

LABOV, William. *Principles of linguistic change*. Cambridge: Blackwell, 1994.

LADEFOGED, Peter. **Phonetic data analysis: an introduction to fieldwork and instrumental techniques**. Malden: Blackwell, 2003.

_____.; MADDIESON, Ian. **The sounds of the world's languages**. Oxford: Blackwell, 1996.

_____. **A course in phonetics**. 5th ed. Boston: Thomson Wadsworth, 2006.

LATSCH, Vagner Luis. **Desenvolvimento de um sistema de conversão texto-fala com modelagem de prosódia**. 2011. 156f. Tese (Doutorado em Engenharia Elétrica)-Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

LEE, Seung Hwa. **Morfologia e fonologia lexical do português do Brasil**. 1995. 200f. Tese (Doutorado em Linguística)-Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1995.

_____. Epêntese no português. **Estudos Linguísticos**. Anais dos Seminários do Grupo de Estudos Linguísticos, Ribeirão Preto, v. 2. 1993, p. 847-854.

_____.; OLIVEIRA, Marco Antônio de. Variação inter e intra dialetal no português brasileiro: um problema para a teoria fonológica. In: OLIVEIRA, Dermeval da Hora; COLLISCHONN, Gisela (Org.). **Teoria linguística: fonologia e outros temas**. João Pessoa: Editora Universitária, 2003. p.67-91.

LEITE, Yone. **Portuguese stress and related rules**. 1974. 304f. Tese de Doutorado (Doutorado em Linguística), The University of Texas, Austin, 1974.

LEVENSHTEIN distance. Algorithms and theory of computation handbook, CRC Press LLC, 1999, "Levenshtein distance". In: Dictionary of Algorithms and Data Structures [online], Vreda Pieterse and Paul E. Black, eds. 22 August 2013. Disponível em: <<http://www.nist.gov/dads/HTML/Levenshtein.html>>. Acesso em: 18 dez. 2014.

LIBOSSEK, Marion; SCHIEL, Florian. *Syllable-based text-to-phoneme conversion for German*. In: INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 2., 2000, Beijing. **International Conference on Spoken Language Processing**, Beijing: Chinese Friendship, 2000, p.283-286.

LUFT, Celso Pedro. **Minidicionário Luft**. 20.ed. São Paulo: Ática, 2001.

MADUREIRA, Sandra et al. *Post-stressed syllables in brazilian portuguese as markers*. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 2., 1999, San Francisco. **XIV International Congress of Phonetic Sciences**, San Francisco: University of California, 1999, p. 917-920.

MAIA, Eleonora Albano. **Phonological and lexical processes in a generative grammar of portuguese**. 1981. 172f. Tese (Doutorado), Brown University, Providence, 1981.

MAIA, Vera Lúcia M. Vogais pretônicas médias na fala de Natal. **Estudos Linguísticos e Literários**, Salvador, v.5, p.209-226, 1986.

MALMBERG, Bertil. **A fonética: no mundo dos sons da linguagem**. Lisboa: Livros do Brasil, 1954.

MANA, Franco; MASSIMINO, Paolo; PACCHIOTTI, Alberto. *Using machine learning techniques for grapheme to phoneme transcription*. In: 7th EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY, 1., 2001, Aalborg. **7th European Conference on Speech Communication and Technology**, Aalborg: Institute of Electrical and Electronics Engineers, 2001. p. 1915-1918.

_____.; QUAZZA, Silvia. *Text-to-speech oriented automatic learning of Italian prosody*. In: 4th EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY, 1, 1995, Madrid. **European Conference on Speech Communication and Technology - EUROSPEECH**, Madrid, 1995. p. 589-592.

MARCATO, Fernanda. **A realização das vogais mediais de prefixos na variedade do noroeste paulista**: relatório técnico. São José do Rio Preto: UNESP, 2010. 101 p.

MARCHAND, Yannick; DAMPER, Robert I. Can syllabification improve pronunciation by analogy of eEnglish? **Natural Language Engineering**, v.13, n.1, p.1-24, 2007. Disponível em: <<https://www.mysciencework.com/publication/read/1460418/can-syllabification-improve-pronunciation-by-analogy>>. Acesso em: 23 out. 2014.

MARQUES, Sandra Maria Oliveira. **As vogais médias pretônicas em situação de contato dialetal**. 2006. 159f. Tese (Doutorado em Língua Portuguesa)-Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

MARQUIAFÁVEL, Vanessa et al. Rule-based algorithms for automatic pronunciation of portuguese verbal inflections. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF PORTUGUESE LANGUAGE - PROPOR 2014, 2014, São Carlos. **International Conference on Computational Processing of Portuguese Language**, São Carlos: Springer, Heidelberg, 2014. p. 36-47.

MARTINS, Ana Maria. A posição dos pronomes pessoais clíticos. In: RAPOSO, Eduardo Paiva et al. (Ed.). **Gramática do português**. Lisboa: Fundação Calouste Gulbenkian, 2013. p.2231-2302.

MARTINS, Evilazia Ferreira. **Os glides no português brasileiro**. 2011. 158f. Dissertação (Mestrado em Linguística Teórica e Descritiva)-Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

MASSINI-CAGLIARI, Gladis. *Epêntese e paragoge: processos fonológicos distintos*. In: II CONGRESSO NACIONAL DA ABRALIN, 25, 2000, Florianópolis. **Congresso Internacional da Abralín e Instituto Linguístico Associação Brasileira de Linguística**, Florianópolis: Universidade Federal de Santa Catarina, 2000. p. 400-410.

_____. **O texto na alfabetização: coesão e coerência**. Campinas: Ed. da Autora, 1997. (Espiral. Série alfabetização, v.1)

_____. **Acento e ritmo**. São Paulo, Contexto, 1992.

_____.; CAGLIARI, Luiz Carlos. Fonética. In: MUSSALIM, Fernanda; BENTES, Anna Christina (Org.). **Introdução à linguística: domínios e fronteiras**. São Paulo: Cortez, 2001. p.105-142.

MATEUS, Maria Helena Mira. **O acento de palavra em português: uma nova proposta**. Lisboa: Universidade de Lisboa, Centro de Linguística, 1983. Separata de Boletim de Filologia, t.28, p.211-229.

_____. **Aspectos da fonologia portuguesa**. 2.ed. rev. Lisboa: Instituto de Alta Cultura, 1982.

_____. **Aspectos da fonologia portuguesa**. Lisboa: Instituto de Alta Cultura, 1975.

_____.; VILLALVA, Alina. **O essencial sobre linguística**. Lisboa: Editorial Caminho, 2006.

_____.; D'ANDRADE, Ernesto. **The phonology of portuguese**. Oxford: Oxford University Press, 2003.

_____.; DUARTE, Inês; FARIA, Isabel Hub. **Gramática da língua portuguesa**. Lisboa: Editorial Caminho, 2003. 1130p.

MESTRE, Antonio Ríos. La transcripción fonética automática del diccionario electrónico de formas simples flexivas del español: un estudio fonológico en el léxico. **Estudios de Lingüística del Español**, n.4, p.1, 1999.

MITCHELL, Tom. **Machine learning**. New York: McGraw-Hill, 1997.

MONARETTO, Valéria; QUEDNAU, Laura R.; HORA, Dermeval. As consoantes do português. In: BISOL, Leda. (Org.). **Introdução a estudos de fonologia do português brasileiro**. Porto Alegre: EDIPUCRS, 1996. p. 205-246, 1996.

MONTE, Anderson et al. *A rule-based syllabification algorithm with stress determination for brazilian portuguese natural language processing*. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 17, 2011, Hong Kong. **17th International Congress of Phonetic Sciences**. Hong Kong: Universidade de Hong Kong, 2011. p. 1418-1421.

MORAES, João Antônio; WETZELS, W. Leo. Sobre a duração dos segmentos vocálicos nasais e nasalizados em português: um exercício de fonologia experimental. **Cadernos de Estudos Linguísticos**, Campinas, v.23, p.153-66, 1992. Disponível em: <<http://revistas.iel.unicamp.br/index.php/cel/article/view/2902/2346>>. Acesso em: 23 out. 2014.

MORAIS, José. **A arte de ler**. Tradução Álvaro Lorencini. São Paulo: Ed. da UNESP, 1995.

MUNIZ, Marcelo C. M. **A construção de recursos linguístico-computacionais para o português do Brasil**: o projeto de Unitex-PB. 2004. 72f. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional)- Instituto de Ciências Matemáticas, Universidade de São Paulo, São Carlos, 2004.

NADIN, Odair L. Dicionários escolares bilíngues de língua espanhola: reflexões sobre obras direcionadas ao aprendiz brasileiro. **Revista de Letras**, v.11, n.11, p.125-144, 2009.

NASCENTES, Antenor. O dialeto brasileiro. In: PINTO, Edith Pimentel. (Sel.). **O português do Brasil**: textos críticos e teóricos: 1920/1945 : fontes para a teoria e a história. Rio de Janeiro: Livros Técnicos e Científicos; São Paulo: EDUSP, 1981. v.2.

_____. **O linguajar carioca**. Rio de Janeiro: Organização Simões, 1953.

NETO, Nelson et al. Free tools and resources for brazilian portuguese speech recognition. **Journal of the Brazilian Computer Society**, v.17, n. 1, p.53-68, 2011.

OLIVEIRA, Catarina Alexandra Monteiro de. **Do grafema ao gesto**: contributos linguísticos para um sistema de síntese de base articulatória. 2009. 296f. Tese (Doutorado em Linguística)-Departamento de Línguas e Cultura, Universidade de Aveiro, Aveiro, 2009.

_____. *On european portuguese automatic syllabification*. In: CONGRESO INTERNACIONAL DE FONÉTICA EXPERIMENTAL, 3., 2007, Santiago de Compostela. **Congresso Internacional de Fonética Experimental**. Santiago de Compostela: Xunta de Galicia, 2007, p.461-473.

OLIVEIRA, Catarina Alexandra Monteiro de. On automatic european portuguese syllabification. **Journal of European Conference on Speech Communication and Technology**. Lisboa. p. 2933-2936, 2005.

_____.; MOUTINHO, Lurdes Castro; TEIXEIRA, Antonio. *Um novo sistema de conversão grafema-fone para PE baseado em transdutores*. In: CONGRESSO INTERNACIONAL DE FONÉTICA E FONOLOGIA, 1, 2004, São Luís. **Congresso Internacional de Fonética e Fonologia**, São Luís: Universidade Federal do Maranhão, 2004.

OLIVEIRA, Luís C. et al. **Progress report of project Dixi+**: A portuguese text-to-speech synthesizer for alternative and augmentative communication: relatório técnico-científico. Lisboa: INESC, 2001.

_____. **Síntese de fala a partir de texto**. Lisboa: Universidade Técnica de Lisboa, Instituto Superior Técnico, 1996.

OLIVEIRA, Luís C. et al. *A rule-based text-to-speech system for portuguese*. In: INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. 2, 1992, São Francisco. **International Conference on Acoustics, Speech, and Signal Processing**. San Francisco: Institute of Electrical and Electronics Engineers, 1992. p.73-76.

_____.; VIANA, Céu; TRANCOSO, Isabel. *DIXI - portuguese text-to-speech system*. In: EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY. 2., 1991, Genova. **European Conference on Speech Communication and Technology**. Genova, 1991. p.1239-1242.

OSTERMANN FILHO, Paulo Eduardo. **Desenvolvimento de regras de pronúncia para a síntese de fala em língua portuguesa**. 2002. 130f. Dissertação (Mestrado em Ciência da Computação)-Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

PARDAL, Ernesto A. **Aspects de la phonologie (générative) du portugais**. Lisboa: Instituto Nacional de Investigação Científica, 1977. (Série publicações do Centro de Linguística da Universidade de Lisboa, v.20)

PEARSON, Steve et al. *Automatic methods for lexical stress assignment and syllabification*. In: INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 6, 2000, Beijing. **International Conference on Spoken Language Processing**. Beijing: Institute of Electrical and Electronics Engineers, 2000. p. 423–426.

PEIXOTO, Jaqueline dos Santos. O ditongo em português: história, variação e gramática. **Revista Linguística**: Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro, Rio de Janeiro, v.7, n.1, p.48-72, 2011.

PIOTROWSKI, Tadeusz. Indication of english pronunciation in bilingual dictionaries. **Applied Linguistics**, Oxford, v.8, n.1, p.39-47, 1987.

PONTES, Antonio Luciano. Aspectos fonéticos em dicionários escolares. **Cadernos do XI Congresso Nacional de Linguística e Filologia**, Rio de Janeiro, v.11, n.4, p.39-47, 2007.

PROCTER, Paul. **Cambridge international dictionary of english**. Cambridge: Cambridge University Press, 1995.

QUICOLI, A. Carlos. Harmony, lowering and nasalization in brazilian portuguese. **Língua**, North-Holland, n.80, p. 295-331, 1990.

QUILIS, Antonio. Diccionarios de pronunciación. **Lingüística Española Actual**, v.4, n.2, p.325-331, 1982.

REDENBARGER, Wayne Jacob. **Articulator features and portuguese vowel height**. 1981. 191f. Tese (Doctoral dissertation on Linguistics)-Department of Romance Languages and Literatures of Harvard University, Cambridge, Massachusetts, 1981.

RIBEIRO, José. **Música caipira**: da roça ao rodeio. São Paulo: Editora 34, 1999.

RIBEIRO, Ricardo; OLIVEIRA, Luís; TRANCOSO, Isabel. *Morphosyntactic disambiguation for TTS Systems*. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 5, 2002, Las Palmas. **International Conference on Language Resources and Evaluation**. Las Palmas: ELRA, 2002. p. 1427-1431.

ROCHA LIMA, Carlos Henrique da. **Gramática normativa da língua portuguesa**. 22.ed. Rio de Janeiro: José Olympio, 1972.

_____. **Gramática normativa da língua portuguesa**. 49 ed.— Rio de Janeiro: José Olympio, 2011.

ROCHA, Willian; NETO, Nelson. *Implementação de um separador silábico gratuito baseado em regras linguísticas para o português brasileiro*. In: BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY, 9, 2013, Fortaleza. **Brazilian Symposium in Information and Human Language Technology**, Fortaleza: SBC. 2013. p. 108-115.

ROTHER-NEVES, Rui; REIS, Camila M. Uma bibliografia da nasalidade vocálica no português. **Letras de Hoje**, Porto Alegre, v.47, n.3, p.299-305, 2012.

SAMPAIO NETO, Nelson Cruz. **Ferramentas e recursos livres para reconhecimento e síntese de voz em português brasileiro**. 2011. 111f. Tese (Doutorado em Engenharia Elétrica)-Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2011.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. Editora Cultrix, 2008.

SCHWINDT, Luiz Carlos. A regra variável de harmonização vocálica no RS. In: BISOL, Leda; BRESCANCINI, Cláudia (Org.). **Fonologia e variação: recortes do português brasileiro**. Porto Alegre: EDIPUCRS, 2002. p.161-182.

SCLIAR-CABRAL, Leonor. **Princípios do sistema alfabético do português do Brasil**. São Paulo: Contexto, 2003.

SEARA, Izabel Christine. 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, p.79- 85, 2002.

_____. 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: BRAZILIAN TELECOMMUNICATION SYMPOSIUM, 19., 2001, Fortaleza. **Brazilian Telecommunication Symposium**, Fortaleza: SBrT, 2001. p. 1-6.

SEARA, Izabel Christine. **Estudo acústico-perceptual da nasalidade das vogais do português brasileiro**. 2000. 288f. Tese (Doutorado em Linguística). Universidade Federal de Santa Catarina, Florianópolis, 2000.

SELKIRK, Elisabeth. **Syntax of words**. Cambridge: MIT Press, 1982.

SHRIBERG, Lawrence; LOF, Gregory. Reliability studies in broad and narrow phonetic transcription. **Clinical Linguistics and Phonetics**, v.5, n.3, p.225-279, 1991.

SHULBY, Christopher; MENDONÇA, Gustavo; MARQUIAFÁVEL, Vanessa. *Automatic disambiguation of homographic heterophone pairs containing open and closed mid vowels*. In: BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY, n.9, 2013, Fortaleza. **Brazilian Symposium in Information and Human Language Technology**, Fortaleza, 2013. p 126-137. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/stil/2013/0014.pdf>>. Acesso em: 06 dez. De 2014.

SILVA, Denilson. **Algoritmos de processamento da linguagem e síntese de voz com emoções aplicados a um conversor texto-fala baseado em HMM**. 2011. 163f. Tese (Doutorado em Engenharia Elétrica)-COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

_____. *Separação das sílabas e determinação da tonicidade no português brasileiro*. In: SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES, n.1, 2008, Rio de Janeiro. **XXVI Simpósio Brasileiro de Telecomunicações**, Rio de Janeiro: SBrT. 2008. p. 1-5.

_____.; BRAGA, Daniela; RESENDE JÚNIOR, Fernando Gil Vicente. A rule-based method for homograph disambiguation in brazilian portuguese text-to-speech systems. **Journal of Communications and Information Systems**, v.1, n.27, p.01-09. 2012.

_____. et al. *A rule-based grapheme-phone converter and stress determination for brazilian Portuguese natural language processing*. In: INTERNATIONAL TELECOMMUNICATIONS SYMPOSIUM, n.6, 2006, Fortaleza. **International Telecommunications Symposium**, Fortaleza: SBrT, 2006. p. 992-996.

SILVA, Elda Lizandra Fernandes da; OLIVEIRA, Hélio Magalhães de. *Implementação de um algoritmo de divisão silábica automática para arquivos de fala na língua portuguesa*. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, n.1, 2012, Campina Grande. **Congresso Brasileiro de Automática**, Campina Grande, 2012. p. 4161-4166.

SILVA, Myrian Barbosa da. **As pretônicas no falar baiano**: a variedade culta de Salvador. 1989. 377f. Tese (Doutorado em Língua Portuguesa)-Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1989.

SILVEIRA, Ana Amélia Menegasso. **As vogais pretônicas na fala culta do noroeste paulista**. 2008. 143f. Dissertação (Mestrado em Estudos Linguísticos)-Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2008.

SILVEIRA, Regina Célia Pagliuchi da. **Uma pronúncia do português brasileiro**. São Paulo: Cortez, 2008.

SIMÕES, Flávio Olmo. **Implementação de um sistema de conversão texto-fala para o português do Brasil**. 1999. 186f. Dissertação (Mestrado em Engenharia Elétrica)-Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 1999.

SIRAVENHA, Ana Carolina. **Uso de regras fonológicas com determinação de vogal tônica para conversão grafema fone em português brasileiro**. 2009. 83f. Monografia (Trabalho de conclusão de Curso)-Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2009.

SOUSA, Ana Cláudia de. A sílaba fonética e a sílaba linguística: algumas considerações. **Working Papers em Linguística**. Florianópolis, n.2, p.117-133, 1998.

SOUZA, Paulo Chagas; SANTOS, Raquel Santana. Fonética. In: FIORIN, José Luís (Org.). **Introdução à linguística: princípios de análise**. 4.ed. São Paulo: Contexto, 2004. v.2. p.9-31.

STEFAN-ADRIAN, Toma; DORU-PETRU, Munteanu. *Rule-based automatic phonetic transcription for the Romanian language*. In: COMPUTATIONWORLD'09 - FUTURE COMPUTING, SERVICE COMPUTATION, COGNITIVE, ADAPTIVE, CONTENT, PATTERNS, 1, 2009, Atenas. **Computation World**. Atenas: Institute of Electrical and Electronics Engineers, 2009. p. 682-686.

STETSON, Raymond Herbert. **Motor phonetics**: a study of speech movements in action. Amsterdam: North-Holland, 1951.

TAYLOR, Paul. *Hidden Markov models for grapheme to phoneme conversion*. In: EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY, 9., Lisboa, 2005. **9th European Conference on Speech Communication and Technology**, Lisboa: Institute of Electrical and Electronics Engineers, 2005. p.1973-1976.

TEIXEIRA, António; OLIVEIRA, Catarina; MOUTINHO, Lurdes Castro. Machine learning of european portuguese grapheme-to-phone conversion using a Richer feature set. **Revista do DETUA**, Aveiro, v.4, n.6, 2006a. p.1-6.

_____. *On the use of machine learning and syllable information in european portuguese grapheme-phone conversion*. In: COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, 7., 2006, Itatiaia. **International Workshop of Computational Processing of the Portuguese Language**. Itatiaia: Springer Berlin Heidelberg, 2006b. p. 212-215.

TEIXEIRA, Graciete; DA COSTA, Margarida Faria; DA SILVA, Sofia Pereira (Ed.). **Grande dicionário da língua portuguesa**. Porto: Porto Editora, 2004.

TEIXEIRA, João Paulo Ramos. **A prosody model to TTS systems**. 2004. 207f. Tese de doutorado (Doutorado em Engenharia Electrotécnica e de Computadores)-Faculdade de Engenharia da Universidade do Porto, Porto, 2004.

_____. et al. *Multivox: conversor texto fala para português*. In: COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE, 3., 1998, Porto Alegre. **Encontro para o Processamento Computacional de Português Escrito e Falado (PROPOR 98)**, Porto Alegre: Springer Berlin Heidelberg, 1998. p. 88-98.

TESPRASIT, Virongrong; CHAROENPORNSAWAT, Paisarn; SORNLERLTLAMVANICH, Virach. *A context-sensitive homograph disambiguation in thai text-to-speech synthesis*. In: CONFERENCE OF THE NORTH AMERICAN CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS ON HUMAN LANGUAGE TECHNOLOGY, 2., 2003, Stroudsburg. **Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology**. Stroudsburg: Association for Computational Linguistics. 2003. p. 103-105.

TEYSSIER, Paul. **História da língua portuguesa**. São Paulo: Martins Fontes, 1997.

TRANCOSO, Izabel et al. *Rule based vs. neural network based approaches to letter-to-phone conversion for portuguese common and proper names*. In: INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 3., 1994, Yokohama. **3rd International Conference on Spoken Language Processing**, Yokohama: ISCA, 1994. p. 1767-1770.

TUFANO, Douglas. **Guia prático da nova ortografia: saiba o que mudou na ortografia brasileira**. São Paulo: Melhoramentos, 2008.

VAN BAEL, Cristophe Patrick Jan. **Validation, automatic generation and use of broad phonetic transcriptions**. 2007. 124f. Tese (Doutor em Germanic Languages)- Universidade de Radboud, Nijmegen, 2007.

VASILÉVSKI, Vera; ARAÚJO, Márcio José. *Um algoritmo prosódico para português do Brasil*. In: JORNADA DE DESCRIÇÃO DO PORTUGUÊS, 3., 2013, Fortaleza. **Jornada de Descrição do Português**. Fortaleza: Sociedade Brasileira de Computação, 2013. p.44-51.

_____. **Construção de um programa computacional para suporte à pesquisa em fonologia do português do Brasil**. 2008. 166f. Tese (Doutorado em Linguística), Universidade Federal de Santa Catarina, Florianópolis, 2008.

VEIGA, Arlindo. **Treino não supervisionado de modelos acústicos para reconhecimento de fala**. 2013. 134f. Tese (Doutorado em Engenharia Eletrônica e de Computadores)- Faculdade de Ciências e Tecnologia, Universidade de Coimbra, Coimbra, 2013.

_____; CANDEIAS, Sara; PERDIGÃO, Fernando. *Generating a pronunciation dictionary for european portuguese using joint-sequence model with embedded stress assignment*. **Journal of the Brazilian Computer Society**, Springer-Verlag. v. 19, n.2, p.127-134, 2013.

_____. *Conversão de grafemas para fonemas em português europeu: abordagem híbrida com modelos probabilísticos e regras fonológicas*. **Linguamática**, Braga, v.3, n.2, p.39-51, 2011.

VIANA, Maria do Ceu; D'ANDRADE, Ernesto. **CORSO I: um conversor de texto ortográfico em código fonético para o português**. Lisboa: CLUL, 1985. (Série relatórios técnicos do Grupo de Fonética e Fonologia, v. 6)

VIEGAS, Maria do Carmo. **O alçamento de vogais médias pretônicas e os itens lexicais**. 2001. 281f. Tese (Doutorado em Estudos Linguísticos)-Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2001.

VIEGAS, Maria do Carmo. **Alçamento de vogais médias pretônicas**: uma abordagem sociolinguística. 1987. 231f. Dissertação (Mestrado em Estudos Linguísticos)-Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 1987.

VIEIRA, Maria Nazaré. **Alteamento dos segmentos /e,o/ pretônicos e do segmento /o/ tônico no falar do médio-amazonas paraense**. 1990. 151f. Tese (Doutorado), Universidade Federal do Pará, Belém, 1990.

VILLAR, Mauro de Salles; BORBA, Francisco da Silva. O trabalho do dicionarista. In: XATARA, Cláudia; BEVILACQUA, Cleci Regina; HUMBLÉ, Philippe René Marie (Org.). **Dicionários na teoria e na prática**: como e para que são feitos. São Paulo: Parábola, 2011. p.17-25.

VIOLARO, Fábio et al. *Um conversor texto-fala para o português brasileiro com processamento linguístico de alta qualidade*. In: SIMPÓSIO BRASILEIRO DE MICROONDAS E OPTOELETRÔNICA. 1., 1996, Curitiba. **VII Simpósio Brasileiro de Microondas e Optoeletrônica**, Curitiba, 1996. p. 361-366.

WEISS, Helga Elisabeth. **Fonética articulatória**: guia e exercícios. Brasília: Linguistics, 1988.

WEISZFLOG, Walter. (Org.). **Michaelis**: moderno dicionário inglês-português, português-inglês. São Paulo: Melhoramentos, 2000.

WELKER, Herbert Andreas. **Dicionários**: uma pequena introdução à lexicografia. Brasília: Thesaurus, 2004.

WELLS, John C. Why phonetic transcription is important. **Journal of the Phonetic Society of Korea**, Seoul, v.31, n.32, p.239-242, 1996.

WERTZNER, Haydée Fiszbein et al. Características fonológicas de crianças com transtorno fonológico com e sem histórico de otite média. **Revista da Sociedade Brasileira de Fonoaudiologia**, São Paulo, v.12, n.1, p.41-47, 2007.

WETZELS, Leo W. Mid-vowel alternations in the brazilian portuguese verb. **Phonology**, Cambridge, v. 12, n. 02, p. 281-304, 1995.

_____. Mid vowel neutralization in brazilian portuguese. **Cadernos de Estudos Linguísticos**, Campinas, n.23, p.19-55, 1992.

_____. Harmonização vocálica, truncamento, abaixamento e neutralização no sistema verbal do português: uma análise autosssegmental. **Cadernos de Estudos Linguísticos**, Campinas, n.21, p.25-58, 1991.

YAROWSKY, David. Homograph disambiguation in text-to-speech synthesis. In: VAN SANTEN, Jan et al. (Ed.). **Progress in speech synthesis**. New York: Springer Verlag, 1996. p.157-172.

ZAVAGLIA, Claudia. *A lexicografia para o público infantil: uma análise macroestrutural de dicionários brasileiros*. In: XIII SIMPÓSIO NACIONAL DE LETRAS E LINGUÍSTICA, III SIMPÓSIO INTERNACIONAL DE LETRAS E LINGUÍSTICA, 2., 2011, Uberlândia.

SILEL, Uberlândia: EDUFU, 2011. p 1-19.

_____. **Os dicionários infantis**: uma análise de suas microestruturas: relatório de estágio de pós-doutorado. São Paulo: Universidade Presbiteriana Mackenzie, 2010. 106f.

_____. A prática lexicográfica multilíngue: questões concernentes ao campo das cores. In: ISQUERDO, Aparecida Negri; ALVES, Ieda Maria (Org.). **As ciências do léxico**: lexicologia, lexicografia, terminologia. São Paulo: Humanitas, 2007. v.3, p.209-222.

_____.; SILVA, Rosa Maria da. Dicionário multilíngue de regência verbal: uma proposta. **Cadernos de Tradução**, Florianópolis, v.2, n.18, p.67-92, 2006.

ANEXO B – Exemplo de saída produzida com o sistema de flexão e transcrição verbal

No primeiro quadro, temos a forma ortográfica flexionada do verbo <caber>. No segundo, a transcrição fonética dessa flexão. Como podemos notar, após a apresentação de uma tabela com as formas verbais conjugadas/pronunciadas, é mostrado um pequeno quadro com informações estatísticas quanto à quantidade de alterações ortográficas/fonéticas, alterações nos sufixos e radicais, bem como uma sequência dos radicais de um dado verbo. Todas essas alterações quantificadas são uma síntese do processo ocorrido para se obter as formas flexionadas do verbo irregular <caber>. As irregularidades do tipo inserção ou substituição estão destacadas em vermelho, enquanto que os apagamentos estão representados pelos pontos também vermelhos, o que torna sua identificação mais rápida e fácil ao usuário do sistema de flexão e transcrição de verbos em PB.

Verbo CABER flexionado - Grafia

Indicativo			
	Presente	Pretérito Imperfeito	Pretérito Perfeito
eu	caibo	cabia	coube
tu	cabes	cabias	coubeste
ele/ela/você	cabe	cabia	coube-
nós	cabemos	cabíamos	coubemos
vós	cabeis	cabíeis	coubestes
eles/elas/vocês	cabem	cabiam	couberam
	Pretérito Mais Que Perfeito	Futuro (Fut. do Presente)	Condicional (Fut. Pretérito)
eu	coubera	caberei	caberia
tu	couberas	caberás	caberias
ele/ela/você	coubera	cabirá	caberia
nós	coubéramos	caberemos	caberíamos
vós	coubéreis	cabereis	caberíeis
eles/elas/vocês	couberam	cabirão	caberiam
Subjuntivo			
	Presente	Pretérito Imperfeito	Futuro
eu	caiba	coubesse	couber
tu	caibas	coubesses	couberes
ele/ela/você	caiba	coubesse	couber
nós	caibamos	coubéssemos	coubermos
vós	caibais	coubésseis	couberdes
eles/elas/vocês	caibam	coubessem	couberem
Infinitivo Pessoal		Imperativo	Gerúndio
eu	caber	caiba	cabendo
tu	caberes	cabe	Particípio Passado
ele/ela/você	caber	caiba	cabido
nós	cabermos	caibamos	Infinitivo
vós	caberdes	cabei	caber
eles/elas/vocês	caberem	caibam	

Nº de alterações ortográficas: 31

Nº de alterações nos sufixos: 6

Nº de radicais diferentes: 5

Radicais: cab, caib, coube, coub, coubé

Quadro 2. Verbo CABER – Transcrição Fonética da flexão

Indicativo			
	Presente	Pretérito Imperfeito	Pretérito Perfeito
eu	'karbo	ka'bia	'koobi·
tu	'kabis	ka'bias	kou'bestʃi
ele/ela/você	'kabi	ka'bia	'koobi·
nós	ka'bemos	ka'biẽmos	kou'bemos
vós	ka'beis	ka'bieis	kou'bestʃis
eles/elas/vocês	'kabẽi	ka'biẽõ	kou'berẽw
	Pretérito Mais Que Perfeito	Futuro (Fut. do Presente)	Condicional (Fut. Pretérito)
eu	kou'bera	kabe'rei	kabe'ria
tu	kou'beras	kabe'ras	kabe'rias
ele/ela/você	kou'bera	kabe'ra	kabe'ria
nós	kou'berẽmos	kabe'remos	kabe'riẽmos
vós	kou'bereis	kabe'reis	kabe'rieis
eles/elas/vocês	kou'berẽõ	kabe'rẽõ	kabe'riẽõ
Subjuntivo			
	Presente	Pretérito Imperfeito	Futuro
eu	'karba	kou'besɪ	kou'beX
tu	'karbas	kou'besɪs	kou'berɪs
ele/ela/você	'karba	kou'besɪ	kou'beX
nós	ka'i'bẽmos	kou'besẽmos	kou'beXẽmos
vós	ka'i'bais	kou'besẽis	kou'beydʒɪs
eles/elas/vocês	'karbẽõ	kou'besẽi	kou'berẽi
	Infinitivo Pessoal	Imperativo	Gerúndio
eu	ka'beX	'karba	ka'bẽ·do
tu	ka'berɪs	'kabi	Particípio Passado
ele/ela/você	ka'beX	'karba	ka'bidu
nós	ka'beĩmos	ka'i'bẽmos	Infinitivo
vós	ka'beydʒɪs	ka'bej	ka'beX
eles/elas/vocês	ka'berẽi	'kaibẽõ	

Nº de alterações na pronúncia: 32

Nº de alterações nos sufixos: 25

Nº de radicais diferentes: 5

Radicais: [kab, kajb, kobi, kobe, kabẽ]

ANEXO C – ERROS OBTIDOS NO TESTE

acontece -> akõõ'tesi | akõõ'tesi | 0.909090909091
 adota -> a'dota | a'dota | 0.833333333333
 alegre -> a'legri | a'legri | 0.857142857143
 alguém -> av'gẽĩ | av'gõem | 0.428571428571
 alicerce -> ali'sersi | ali'sersi | 0.888888888889
 amsterdã -> amister'dẽ | amster'dẽ | 0.909090909091
 andréia -> ẽ'dreia | ẽ'dreia | 0.875
 aplaudir -> aplau'dzir | aplau'dzir | 0.888888888889
 apoiará -> apoia'ra | apoia'ra | 0.875
 aposta -> a'põsta | a'põsta | 0.857142857143
 arquiteto -> arki'tetu | arki'tetu | 0.888888888889
 assume -> a'sumi | a'sũmi | 0.857142857143
 bahia -> ba'ia | 'bahia | 0.666666666667
 baiana -> bai'ẽna | bai'ẽna | 0.875
 bairro -> 'baixu | 'baixu | 0.833333333333
 banheiro -> bẽ'neiru | ba'neiru | 0.777777777778
 basquete -> bas'kefi | bas'kefi | 0.875
 bebe -> 'bebi | 'bebi | 0.8
 biológico -> bio'loziku | bio'loziku | 0.8
 boa -> 'boa | 'boa | 0.75
 boas -> 'boas | 'boas | 0.8
 bola -> 'bola | 'bola | 0.8
 botafogo -> bõta'fogu | bõta'fogu | 0.888888888889
 breve -> 'brevi | 'brevi | 0.833333333333
 cabelo -> ka'belu | ka'belu | 0.857142857143
 caderno -> ka'dernu | ka'dernu | 0.875
 caiapó -> kaia'põ | kaia'põ | 0.857142857143
 certa -> 'serta | 'serta | 0.833333333333
 certas -> 'sertas | 'sertas | 0.857142857143
 certo -> 'sertu | 'sertu | 0.833333333333
 chefe -> 'fefi | 'fefi | 0.8
 cheia -> 'feia | 'feia | 0.8
 chorai -> fo'rai | fo'rai | 0.833333333333
 chuchu -> fu'fu | 'fufu | 0.6
 cineastas -> sine'astas | sini'astas | 0.9
 claudia -> 'klaudzia | klau'dzia | 0.625
 clóvis -> 'klõvis | 'klõvis | 0.857142857143
 coloque -> ko'lõki | ko'lõki | 0.857142857143
 companhia -> kõõpẽ'nia | kõõpa'nia | 0.833333333333
 complexidade -> kõõpleksi'dadzı | kõõplefi'dadzı | 0.866666666667

concessionária -> kōšsesio'naria | kōšsesio'naria | 0.9375
 concorda -> kōō'kōrda | kōō'kōrda | 0.909090909091
 congresso -> kōō'gresu | kōō'gresu | 0.909090909091
 conhecem -> ko'nesēĩ | ko'nesēĩ | 0.9
 conseguiui -> kōōse'giu | kōōsegu'iu | 0.727272727273
 considera -> kōōsi'dera | kōōsi'dera | 0.916666666667
 conversa -> kōō'versa | kōō'versa | 0.909090909091
 cooperativa -> kopera'fiva | 'koperafiva | 0.818181818182
 cooperativas -> kopera'fivas | 'koperafivas | 0.833333333333
 copa -> 'kōpa | 'kōpa | 0.8
 copo -> 'kōpu | 'kōpu | 0.8
 coragem -> ko'razēĩ | kō'razēĩ | 0.9
 coração -> kōra'sēō | kōra'sēō | 0.9
 coreanos -> kōre'ēnos | kōri'anos | 0.666666666667
 corpus -> 'kōrpus | kor'pus | 0.428571428571
 costa -> 'kōsta | 'kōsta | 0.833333333333
 covas -> 'kōvas | 'kōvas | 0.833333333333
 crepe -> 'krēpi | 'krēpi | 0.833333333333
 cresce -> 'krēsɪ | 'krēsɪ | 0.833333333333
 cuiabá -> kuia'ba | kuia'ba | 0.857142857143
 cuidado -> kui'dadu | kui'dadu | 0.875
 cuidar -> kui'dar | kui'dar | 0.857142857143
 cuidarei -> kuida'rei | kuida'rei | 0.888888888889
 decore -> de'kōri | de'kōri | 0.857142857143
 der -> 'dēr | 'dēr | 0.75
 deserto -> de'zertu | de'zertu | 0.875
 desindexação -> dezīdexa'sēō | dezīdeja'sēō | 0.866666666667
 desorganização -> dezorgēniza'sēō | dezorganiza'sēō | 0.888888888889
 desprezível -> despre'ziveu | despre'ziveu | 0.916666666667
 dessa -> 'dēsa | 'dēsa | 0.8
 desta -> 'dēsta | 'dēsta | 0.833333333333
 desvie -> dez'viu | 'dezviu | 0.714285714286
 deteriorando -> deterio'rēdu | deterio'rēdu | 0.923076923077
 deve -> 'dēvi | 'dēvi | 0.8
 devem -> 'dēvēĩ | 'dēvēĩ | 0.875
 dez -> 'dēz | 'dēz | 0.5
 dezessete -> deze'seĩ | deze'seĩ | 0.888888888889
 drogas -> 'drōgas | 'drōgas | 0.857142857143
 e -> 'e | 'e | 0.5
 eleitoral -> eleito'rau | eleito'rau | 0.9
 encoraja -> ēĩko'raza | ēĩko'raza | 0.909090909091
 era -> 'ēra | 'ēra | 0.75
 eram -> 'erēō | 'erēō | 0.857142857143

erra -> 'ɛxa | 'ɛxa | 0.75
 escreve -> ɪs'krɛvɪ | ɪs'krɛvɪ | 0.875
 espera -> ɪs'pɛra | ɪs'pɛra | 0.857142857143
 esportes -> ɪs'pɔrtɛs | ɪs'pɔrtʃɪs | 0.666666666667
 essa -> 'ɛsa | 'ɛsa | 0.75
 este -> 'ɛʃtɪ | 'ɪʃtɪ | 0.8
 europa -> ɛʊ'rɔpa | ɛʊ'rɔpa | 0.857142857143
 excelência -> ɛsɛ'lɛ̃nsia | 'ɛsɛlɛ̃nsia | 0.833333333333
 exhibe -> ɛ'zibɪ | ɛ'ʃibɪ | 0.833333333333
 exibição -> ɛzibɪ'sɛ̃õ | ɛʃibɪ'sɛ̃õ | 0.909090909091
 exige -> ɛ'zizɪ | ɛ'ʃizɪ | 0.833333333333
 exigência -> ɛzi'ʒɛ̃nsia | ɛʃi'ʒɛ̃nsia | 0.916666666667
 existe -> ɛ'zistʃɪ | ɛ'ʃistʃɪ | 0.857142857143
 existiam -> ɛzis'ʃiɛ̃õ | ɛʃis'ʃiɛ̃õ | 0.909090909091
 existência -> ɛzis'tɛ̃nsia | ɛʃis'tɛ̃nsia | 0.923076923077
 explode -> ɪs'plɔdʒɪ | ɪs'plɔdʒɪ | 0.875
 explorar -> ɪsplɔ'rɔr | ɪsplɔ'rɔr | 0.888888888889
 externo -> ɪs'tɛrnɔ | ɪs'tɛrnɔ | 0.875
 febre -> 'fɛbrɪ | 'fɛbrɪ | 0.833333333333
 fecha -> 'fɛʃɪ | 'fɛʃɪ | 0.8
 feminina -> femi'nina | femi'nina | 0.9
 foco -> 'fɔkɔ | 'fɔkɔ | 0.8
 foram -> 'forɛ̃õ | 'fɔrɛ̃õ | 0.875
 forte -> 'fɔrtʃɪ | 'fɔrtʃɪ | 0.833333333333
 foto -> 'fɔtɔ | 'fɔtɔ | 0.8
 franquadas -> frɛ̃kɛ'adas | frɛ̃kɛ'adas | 0.909090909091
 freia -> 'frɛia | 'frɛia | 0.833333333333
 freios -> 'frɛjus | frɛ'jus | 0.714285714286
 frota -> 'frɔta | 'frota | 0.833333333333
 gaiivotas -> gar'vɔtas | gar'vɔtas | 0.888888888889
 galope -> ga'lɔpi | ga'lɔpi | 0.857142857143
 ganhou -> gɛ̃'ɲɔʊ | ga'ɲɔʊ | 0.714285714286
 gilberto -> ʒiɔ'bertɔ | ʒiɔ'bertɔ | 0.888888888889
 gol -> 'gɔʊ | 'gɔʊ | 0.75
 gole -> 'gɔli | 'gɔli | 0.8
 goleada -> gole'ada | gɔli'ada | 0.875
 gols -> 'gɔʊs | 'gɔls | 0.8
 gostam -> 'gɔstɛ̃õ | 'gɔstɛ̃õ | 0.888888888889
 gravadoras -> grava'dɔras | grava'dɔras | 0.909090909091
 greve -> 'grɛvɪ | 'grɛvɪ | 0.833333333333
 guarani -> gɔarɛ̃'ni | gɔara'ni | 0.777777777778
 guerra -> 'gɛxa | 'gɛxa | 0.8
 habeas -> 'abeas | a'beas | 0.666666666667
 ibope -> i'bɔpi | i'bɔpi | 0.833333333333

ideal -> ide'ao | idr'ao | 0.833333333333
 idéia -> i'deia | i'deia | 0.833333333333
 idéias -> i'deias | i'deias | 0.857142857143
 ignora -> igi'nora | i'ginora | 0.75
 ignota -> igi'nota | i'ginota | 0.625
 importa -> i'porta | i'porta | 0.875
 impressionável -> i'presio'naveo | i'presio'naveo | 0.928571428571
 inclui -> i'klui | i'klui | 0.714285714286
 incorreto -> i'ko'xeto | i'ko'xeto | 0.888888888889
 indiscreta -> i'dzis'kreta | i'dzis'kreta | 0.909090909091
 infernos -> i'fernus | i'fernus | 0.888888888889
 influi -> i'flui | i'flui | 0.714285714286
 informa -> i'forma | i'forma | 0.875
 Inglaterra -> i'gla'texa | i'gla'texa | 0.9
 instante -> i's'teifi | i'nsteifi | 0.666666666667
 institucionalização -> i'stj'itucionaliza'seõ | i'stj'itucionaliza'seõ | 0.909090909091
 instituição -> i'stj'itui'seõ | i'stj'itui'seõ | 0.928571428571
 instrumento -> i'stru'meĩtu | i'stru'meĩtu | 0.928571428571
 interessa -> i'te'resa | i'te'resa | 0.888888888889
 joga -> 'zoga | 'zoga | 0.8
 jorge -> 'zorzi | 'zorzi | 0.833333333333
 jovens -> 'zovẽis | 'zovẽis | 0.888888888889
 judi -> 'judzi | 'zu'dzi | 0.6
 labaki -> la'baki | la'baki | 0.857142857143
 led -> 'ledzi | 'ledzi | 0.6
 leia -> 'leia | 'leia | 0.8
 levam -> 'levẽõ | 'levẽõ | 0.875
 lisboa -> liz'boa | liz'boa | 0.857142857143
 litoral -> lito'rao | lito'rao | 0.875
 livre -> 'livri | 'livri | 0.833333333333
 livro -> 'livro | 'livro | 0.833333333333
 livros -> 'livros | 'livros | 0.857142857143
 loja -> 'loza | 'loza | 0.8
 lojas -> 'lozas | 'lozas | 0.833333333333
 lopes -> 'lopi | 'lopi | 0.833333333333
 luisinho -> lur'ziho | lur'ziho | 0.444444444444
 maior -> mai'or | mai'or | 0.833333333333
 maioria -> mai'ria | mai'ria | 0.875
 marcus -> 'markos | 'markos | 0.571428571429
 mayrink -> mai'rĩki | mai'rĩki | 0.666666666667
 meia -> 'meia | 'meia | 0.8
 melhor -> me'lor | me'lor | 0.833333333333
 melhoramentos -> me'loram'eĩtus | me'loram'eĩtus | 0.933333333333

menor -> me'nɔr | me'nɔr | 0.833333333333
 menores -> me'nɔrɪs | me'nɔrɪs | 0.875
 mestre -> 'mɛstrɪ | 'mɛstrɪ | 0.857142857143
 miki -> 'mɪkɪ | 'mɪkɪ | 0.8
 milani -> mi'lɛni | mila'ni | 0.5
 modo -> 'mɔdɔ | 'mɔdɔ | 0.8
 moqueca -> mo'kɛka | mo'kɛka | 0.857142857143
 moradores -> mɔra'dɔrɪs | mɔra'dɔrɪs | 0.9
 morava -> mo'rava | mɔ'rava | 0.857142857143
 morre -> 'mɔxi | 'mɔxi | 0.8
 mortas -> 'mɔrtas | 'mɔrtas | 0.857142857143
 morte -> 'mɔrtʃɪ | 'mɔrtʃɪ | 0.833333333333
 mostra -> 'mɔstra | 'mɔstra | 0.857142857143
 mota -> 'mɔta | 'mɔta | 0.8
 mulher -> mu'ʎɛr | mu'ʎɛr | 0.833333333333
 mulheres -> mu'ʎɛrɪs | mu'ʎɛrɪs | 0.875
 nessas -> 'nɛsas | 'nɛsas | 0.833333333333
 nesta -> 'nɛsta | 'nɛsta | 0.833333333333
 neto -> 'nɛtɔ | 'nɛtɔ | 0.8
 neves -> 'nɛvɪs | 'nɛvɪs | 0.833333333333
 ninguém -> nĩ'gɛĩ | nĩ'gɔɛm | 0.5
 noal -> no'aʊ | no'aʊ | 0.8
 nobre -> 'nɔbrɪ | 'nɔbrɪ | 0.833333333333
 nordeste -> nor'dɛstʃɪ | nor'dɛstʃɪ | 0.888888888889
 norte -> 'nɔrtʃɪ | 'nɔrtʃɪ | 0.833333333333
 nossa -> 'nɔsa | 'nɔsa | 0.8
 nosso -> 'nɔsɔ | 'nɔsɔ | 0.8
 nossos -> 'nɔsɔs | 'nɔsɔs | 0.833333333333
 nova -> 'nɔva | 'nɔva | 0.8
 novamente -> nɔva'mɛĩʃɪ | nɔva'mɛĩʃɪ | 0.833333333333
 nove -> 'nɔvɪ | 'nɔvɪ | 0.8
 novecentos -> nɔvi'sɛĩtus | nove'sɛĩtus | 0.846153846154
 obra -> 'ɔbra | 'ɔbra | 0.8
 olhos -> 'ɔʎɔs | 'ɔʎɔs | 0.8
 onu -> 'ɔnu | o'nu | 0.4
 ordens -> 'ɔrdɛĩs | 'ɔrdɛĩs | 0.888888888889
 orestes -> o'rɛstʃɪs | o'rɛstʃɪs | 0.875
 organiza -> orgɛ'niza | orga'niza | 0.8
 pacote -> pa'kɔʃɪ | pa'kɔʃɪ | 0.857142857143
 palavra -> pa'lavra | pa'lavra | 0.875
 palestra -> pa'lestra | pa'lestra | 0.888888888889
 passear -> pase'ar | pasi'ar | 0.857142857143
 pede -> 'pɛdʒɪ | 'pɛdʒɪ | 0.8
 pelas -> 'pɛlas | 'pɛlas | 0.833333333333

pelos -> 'pelus | 'pelus | 0.833333333333
 perdem -> 'pɛrdẽĩ | 'pɛrdẽĩ | 0.888888888889
 perdoar -> pɛrdo'ar | pɛrdu'ar | 0.875
 perna -> 'pɛrna | 'pɛrna | 0.833333333333
 perplexo -> pɛr'plɛksu | pɛr'plɛʃu | 0.7
 pessoal -> pɛso'au | pɛsu'au | 0.857142857143
 pessoas -> pɛ'soas | pɛ'suas | 0.857142857143
 petróleo -> pɛ'trɔliu | pɛ'trɔleu | 0.888888888889
 peça -> 'pɛsa | 'pɛsa | 0.8
 pior -> pi'ɔr | pi'or | 0.8
 pneus -> pi'neus | 'pineus | 0.714285714286
 pobre -> 'pɔbrɪ | 'pɔbrɪ | 0.833333333333
 pode -> 'pɔdʒɪ | 'pɔdʒɪ | 0.8
 podem -> 'pɔdẽĩ | 'pɔdẽĩ | 0.875
 popoca -> po'pɔka | po'pɔka | 0.857142857143
 porque -> por'ke | por'ki | 0.833333333333
 porta -> 'pɔrta | 'pɔrta | 0.833333333333
 português -> portu'ges | portu'gues | 0.9
 possa -> 'pɔsa | 'pɔsa | 0.8
 posse -> 'pɔsɪ | 'pɔsɪ | 0.8
 posso -> 'pɔsu | 'pɔsu | 0.8
 praia -> 'praia | 'praia | 0.833333333333
 processo -> pro'sɛsu | pro'sɛsu | 0.875
 profissionais -> profisio'nais | profisio'nais | 0.923076923077
 projeto -> pro'ʒɛtu | pro'ʒɛtu | 0.875
 projetos -> pro'ʒɛtus | pro'ʒɛtus | 0.888888888889
 promete -> pro'mɛʃi | pro'mɛʃi | 0.875
 proposta -> pro'pɔsta | pro'pɔsta | 0.888888888889
 prossegue -> pro'sɛgi | pro'sɛgi | 0.875
 provoca -> pro'vɔka | pro'vɔka | 0.875
 próxima -> 'prɔsɪma | 'prɔʃɪma | 0.75
 psiquiatra -> pɪsiki'atra | pɪsiqui'atra | 0.833333333333
 puder -> pu'dɛr | pu'dɛr | 0.833333333333
 quadrilha -> kɔa'driɫa | kɔadʒɪ'riʃa | 0.444444444444
 quadro -> 'kɔadɾu | 'kɔadʒɪɾu | 0.75
 qual -> 'kɔau | qu'au | 0.4
 qualquer -> kɔau'kɛr | kɔau'kɛr | 0.875
 quatro -> 'kɔatɾu | 'kɔaʃɪɾu | 0.75
 quatrocentos -> kɔatro'sɛĩtus | kɔaʃɪ'rosɛĩtus | 0.75
 que -> 'ke | 'qui | 0.25
 queda -> 'kɛda | 'kɛda | 0.8
 queimadura -> keima'dura | kei'madura | 0.8
 quem -> 'kẽĩ | qu'ẽĩ | 0.5

quente -> 'kẽĩfĩ | 'koẽĩfĩ | 0.888888888889
 quer -> 'kɛr | qu'er | 0.0
 quero -> 'kɛrɔ | 'kerɔ | 0.8
 quisier -> ki'zɛr | ki'zer | 0.833333333333
 raia -> 'xara | 'xaia | 0.8
 reais -> xe'ais | xɪ'ais | 0.666666666667
 reajuste -> xea'zʊstʃĩ | xɪa'zʊstʃĩ | 0.888888888889
 real -> xe'au | xɪ'au | 0.8
 realizadas -> xeali'zadas | xɪali'zadas | 0.909090909091
 realizar -> xeali'zar | xɪali'zar | 0.888888888889
 reação -> xea'sẽõ | xɪa'sẽõ | 0.888888888889
 redor -> xe'dɔr | xe'dor | 0.833333333333
 refletem -> xe'fletẽĩ | xe'fletẽĩ | 0.909090909091
 reforma -> xe'fɔrma | xe'forma | 0.875
 repete -> xe'pɛʃĩ | xe'peʃĩ | 0.857142857143
 requiãõ -> xeki'ẽõ | xequi'ẽõ | 0.8
 reserva -> xe'zɛrva | xe'zerva | 0.875
 resta -> 'xɛsta | 'xesta | 0.833333333333
 resto -> 'xɛstɔ | 'xestɔ | 0.833333333333
 reta -> 'xeta | 'xeta | 0.8
 ricota -> xi'kɔta | xi'kota | 0.857142857143
 rodas -> 'xɔdas | 'xodas | 0.833333333333
 roraima -> xo'raima | xɔ'raima | 0.875
 ruins -> 'xuĩs | 'xuĩs | 0.833333333333
 rótula -> 'xɔtula | 'xɔtula | 0.857142857143
 sanitário -> sɛni'tario | sani'tario | 0.818181818182
 sete -> 'sɛʃĩ | 'seʃĩ | 0.8
 setecentos -> sete'sẽĩtus | sete'sẽĩtus | 0.923076923077
 sexo -> 'sɛksɔ | 'seʃɔ | 0.5
 sexual -> seksu'au | seʃu'au | 0.714285714286
 simultânea -> simuɔ'tẽnia | simuɔ'tẽnea | 0.916666666667
 soares -> so'aris | 'sɔaris | 0.571428571429
 soberania -> soberẽ'nia | sobera'nia | 0.818181818182
 solo -> 'sɔlɔ | 'solɔ | 0.8
 sozinho -> sɔ'zĩɲɔ | so'zĩɲɔ | 0.875
 sucesso -> su'sesɔ | su'sesɔ | 0.857142857143
 sugere -> su'zɛrɪ | su'zerɪ | 0.857142857143
 suor -> su'ɔr | su'or | 0.8
 teatro -> te'atrɔ | tɪ'atrɔ | 0.857142857143
 tecnologia -> tekɪnolo'zia | tekɪnolo'zia | 0.916666666667
 temporada -> tẽĩpo'rada | tẽĩpɔ'rada | 0.916666666667
 terra -> 'tɛxa | 'tɛxa | 0.8
 terras -> 'tɛxas | 'tɛxas | 0.833333333333
 tese -> 'tɛzi | 'tezi | 0.8

testa -> 'tɛstɐ | 'testɐ | 0.833333333333
 teste -> 'tɛstʃi | 'testʃi | 0.833333333333
 toca -> 'tɔkɐ | 'tokɐ | 0.8
 torce -> 'tɔrsɪ | 'torsɪ | 0.833333333333
 torna -> 'tɔrnɐ | 'tornɐ | 0.833333333333
 torta -> 'tɔrtɐ | 'tortɐ | 0.833333333333
 trimestre -> tri'mɛstrɪ | tri'mɛstrɪ | 0.9
 trouxe -> 'trɔusɪ | 'trɔʊʃi | 0.857142857143
 tupãzinho -> tupɐ̃'ziɲɔ | tu'pɛ̃ziɲɔ | 0.8
 táxi -> 'taksɪ | 'taʃi | 0.6
 uma -> 'umɐ | 'ũmɐ | 0.8
 une -> 'unɐ | 'ũnɪ | 0.6
 urbanismo -> urbɐ̃'nizmɔ | urba'nizmɔ | 0.818181818182
 urbanista -> urbɐ̃'nistɐ | urba'nistɐ | 0.818181818182
 velha -> 'vɛlɐ | 'vɛlɐ | 0.8
 velho -> 'vɛlɔ | 'vɛlɔ | 0.8
 verbas -> 'vɛrbɐs | 'vɛrbɐs | 0.857142857143
 verso -> 'vɛrsɔ | 'vɛrsɔ | 0.833333333333
 veta -> 'vɛtɐ | 'vɛtɐ | 0.8
 vieram -> vi'erɐ̃ | vi'erɐ̃ | 0.888888888889
 volta -> 'vɔltɐ | 'vɔltɐ | 0.833333333333
 vossa -> 'vɔsɐ | 'vɔsɐ | 0.8
 voto -> 'vɔtɔ | 'vɔtɔ | 0.8
 vídeo -> 'vidɛɔ | 'vidɛɔ | 0.833333333333
 zero -> 'zɛrɔ | 'zɛrɔ | 0.8
 zimbábue -> zĩ'babue | 'zĩbabue | 0.777777777778