



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE HUMANIDADES
DEPARTAMENTO DE LETRAS VERNÁCULAS
PROGRAMA DE PÓS – GRADUAÇÃO EM LINGUÍSTICA

CID IVAN DA COSTA CARVALHO

**TRANSDUTOR DE ESTADOS FINITOS PARA CONVERSÃO DE GRAFEMA PARA
A PRONÚNCIA DA VARIEDADE LINGUÍSTICA POTIGUAR**

FORTALEZA

2016

CID IVAN DA COSTA CARVALHO

TRANSDUTOR DE ESTADOS FINITOS PARA CONVERSÃO DE GRAFEMA PARA
A PRONÚNCIA DA VARIEDADE LINGUÍSTICA POTIGUAR

Tese apresentada ao Programa de Pós-graduação em Linguística, do Departamento de Letras Vernáculas, da Universidade Federal do Ceará, como requisitos para obtenção do título de Doutor em Linguística.

Linha de pesquisa: Linguística Aplicada

Orientador: Prof. Dr. Leonel Figueiredo de Alencar Araripe.

FORTALEZA

2016

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

C322t Carvalho, Cid Ivan da Costa.

Transdutor de estados finitos para conversão de grafema para a pronúncia da variedade linguística potiguar / Cid Ivan da Costa Carvalho. – 2016.

160 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística, Fortaleza, 2016.

Orientação: Prof. Dr. Leonel Figueiredo de Alencar Araripe.

Coorientação: Prof. Dr. Moisés Batista da Silva.

1. Fonética. . 2. Fonologia.. 3. Transcrição fonética.. 4. SAMPA. I. Título.

CDD 410

CID IVAN DA COSTA CARVALHO

TRANSDUTOR DE ESTADOS FINITOS PARA CONVERSÃO DE GRAFEMA PARA
A PRONÚNCIA DA VARIEDADE LINGUÍSTICA POTIGUAR

Tese apresentada ao Programa de Pós-graduação em Linguística, da Universidade Federal do Ceará, como requisitos para obtenção do título de Doutor em Linguística. Linha de pesquisa: Linguística Aplicada.

Aprovada em 06 de maio de 2016.

BANCA EXAMINADORA

Prof. Dr. Leonel Figueiredo de Alencar Araripe (orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Moisés Batista da Silva (coorientador)
Universidade do Estado do Rio Grande do Norte (UERN)

Prof. Dra. Maria Elias Soares (Membro interno)
Universidade Federal do Ceará (UFC)

Prof. Dra. Rosemeire Selma Monteiro-Plantin (Membro interno)
Universidade Federal do Ceará (UFC)

Prof. Dra. Vlândia Celia Pinheiro (Membro externo)
Universidade de Fortaleza –UNIFOR

Prof. Dra. Elisangela Nogueira Teixeira (Membro externo)
Universidade Federal do Ceará (UFC)

In memoriam

Helânio Gomes de Paiva

AGRADECIMENTOS

Agradeço

ao meu Deus e Pai: Jesus Cristo, pela graça e misericórdia sempre presente em minha vida;

à minha esposa: Kaliane,
à minha filha: Ana Beatriz,
ao meu filho: João Vítor,
aos meus pais: Manoel e Zilmar,
aos meus irmãos: Meire, Francisco, Magna, José, Sidney, Mônica e Fátima, pelo amor, conforto, apoio, ajuda e confiança a mim depositada;

Aos meus amigos: Alexandre, Kaline, Nagib e Neidson, pelas boas conversas que sempre tivemos;

Aos meus colegas do PPGL: Hélio Leonam Barroso Silva, Davis Macedo Vasconcelos, Ednardo Luiz da Costa, Mardônio França, Jéssica Oliveira de Souza, Gleiberson Nogueira, pelas leituras e contribuições à escrita da tese e ao desenvolvimento do Potigrafone.

Ao meu orientador: Leonel Figueiredo de Alencar Araripe, pelo trabalho dedicado a minha pesquisa;

Ao meu coorientador: Moisés Batista da Silva, pelas contribuições ao desenvolvimento deste trabalho;

Aos membros da banca: Maria Elias Soares, Rosimeire Selma Monteiro-Plantin, Vlândia Celia Pinheiro, Elisângela Nogueira Teixeira, pela leitura e sugestões a este trabalho.

RESUMO

O Potigrafone é um transdutor de estados finitos que faz a transcrição automática das formas gráficas da língua portuguesa para as formas fonéticas da variedade linguística potiguar. Com o desenvolvimento desse sistema, conseguimos identificar e separar os padrões fonéticos da variedade linguística potiguar, através do Atlas linguístico do Centro-Oeste potiguar e no Atlas Geolinguístico do Litoral Potiguar; apresentar uma ferramenta de transcrição fonética utilizando o alfabeto fonético *Speech Assessment Methods Phonetic Alphabet* (SAMPA) e, além disso, conseguimos avaliar o desempenho do sistema em *corpus* de textos escritos de língua portuguesa por meio de duas medidas multiníveis: *Exact Match Ratio* e *Labelling Fscore*. Para isso, foram atendidos passos fundamentais: primeiro, foram necessárias a identificação e a separação dos padrões fonológicos presentes nas cartas fonéticas dos Atlas supracitados; segundo, a definição de alguns critérios pertinentes às entradas (*inputs*) do sistema; terceiro, a implementação das regras que relacionam a forma subjacente com a forma superficial e; por último, a escolha das duas medidas adequadas à avaliação de um sistema que apresentava mais de uma forma de transcrição. Os resultados mostram que o desempenho do Potigrafone foi de 81% para a medida *Exact Match Ratio* e 89% para a medida *Labelling FScore*, apresentando, assim, boa performance na exatidão e nos acertos totais e parciais, no que se refere à transcrição sem a marcação da sílaba tônica e sem os fenômenos fonéticos da nasalidade, da palatalização, da ditongação, da monotongação, do apagamento, da epêntese; outro resultado importante é que os módulos de consoantes e de núcleo silábico também interferem no desempenho, uma vez que os valores com todos os módulos e sem esses dois módulos se assemelham. O Potigrafone apresenta bom desempenho na transcrição fonética, mas existem alguns erros pontuais que torna necessário a constituição de novas regras para corrigi-los e, conseqüentemente, melhorar o desempenho do sistema, tanto na exatidão quanto na correspondência total e parcial da transcrição.

Palavras-chave: Potigrafone. Variedade Linguística Potiguar. Atlas linguístico do Centro-Oeste potiguar. Atlas Geolinguístico do Litoral Potiguar. SAMPA.

ABSTRACT

The Potigrafone is a finite-state transducer that automatically transcribes the graphic forms from the Portuguese language to the phonetic forms belonging to the “potiguar” linguistic variety. It was through the development of this system, by using of the Atlas Linguístico do Centro-Oeste Potiguar and to the Atlas Geolinguístico do Litoral Potiguar, that it has been possible to identify and separate the phonetic patterns in the linguistic variety belonging to the studied region, presenting a phonetic transcription tool while using the Speech Assessment Methods Phonetic Alphabet - SAMPA; and, in addition, a performance assessment in the corpus of written texts in Portuguese has been made through the use of two multi-level measures: Exact Match Ratio and Labelling Fscore. To do so, the identification and separation of phonological patterns present in the phonetic letters of Atlas, the definition of some relevant criteria to inputs in the system, the implementation of rules relating to underlying form with the surface form and, also, two system evaluation measures have been needed. The results have showed that the Potigrafone performance was 81% for the Exact Match Ratio measurement and 89% for the Labelling FScore, presenting thus, increased accuracy in total or partial correct attempts, when the transcript is made without stress markup and without the phonetic phenomena of nasality, the palatalization, the diphthongization, the monophthongization, the deletion, the epenthesis; another important result is that the consonant modules and the syllabic core also interfere with this performance, since the values with all modules and without these two modules are similar. The Potigrafone performs well during the phonetic transcription, but there are some occasional errors that have made necessary the creation of new rules to correct them and, consequently, help to improve the system performance, both in accuracy and in full and partial transcription match.

Keywords: Potigrafone. Linguistic variety Potiguar. Atlas Linguístico do Centro-Oeste potiguar. Atlas Geolinguístico do Litoral Potiguar. SAMPA.

LISTA DE FIGURAS

Figura 1 - Regra de nasalidade	57
Figura 2 - Regra de vozeamento	57
Figura 3 - Regra de harmonia vocálica	58
Figura 4 - Regra de palatalização	58
Figura 5 - Regra de lenição ou enfraquecimento	59
Figura 6 - Regra de apagamento	60
Figura 7 - Regra de monotongação	61
Figura 8 - Regra de ditongação	61
Figura 9 - Regra de epêntese	62
Figura 10 - Grafo de transição de L_1	74
Figura 11 - Língua formal, expressão regular e rede de estados finitos	75
Figura 12 - Mapa da Região Centro-Oeste Potiguar	83
Figura 13 - Carta fonética do Atlas Linguístico do Centro-Oeste Potiguar	85
Figura 14 - Mapa da Região do Litoral Potiguar	86
Figura 15 - Carta fonética do Atlas Linguístico do Litoral Potiguar	87
Figura 16- Alinhamento de grafema com fone	92
Figura 17 - Arquitetura do sistema Potigrafone	94
Figura 18 - Interface gráfica do Potigrafone	96
Figura 19 - Alinhamento padrão dos dicionários e das transcrições fonéticas ...	117
Figura 20 - Obtenção dos dados pelo sistema estatístico	119

LISTA DE GRÁFICOS

Gráfico 1 – Modelo de estrutura silábica baseado em onset-rima	65
Gráfico 2 – A sílaba no português com glide em posição pós-vocálica	66
Gráfico 3 – A sílaba no português com glide em posição pré-vocálica	66
Gráfico 4 – Níveis da representação lexical	67
Gráfico 5 – Hierarquia de sonoridade	68
Gráfico 6 – Grafo de transição de palavra não concatenada	104
Gráfico 7 – Grafo de transição de palavra concatenada	104
Gráfico 8 – O uso do operador <i>Kleene plus</i> numa língua formal	106
Gráfico 9 – O uso do operador <i>Kleene star</i> numa língua formal	107
Gráfico 10 - Concatenação do conjunto A	107
Gráfico 11 - Concatenação do conjunto B	107
Gráfico 12 - Concatenação dos transdutores A e B	108
Gráfico 13 - União de transdutores $A \cup B$	108
Gráfico 14 - Intersecção de transdutores $A \cap B$	109
Gráfico 15 - Autômato do ataque complexo do português	111
Gráfico 16 - Autômato da coda silábica	113
Gráfico 17 - Autômato da rima	113
Gráfico 18 - Medida <i>Exact Match Ratio</i>	121
Gráfico 19- Medida <i>Labelling Fscore</i>	122
Gráfico 20 - Dados da avaliação do sistema	123

LISTA DE TABELAS E QUADROS

Tabela 1 -	Resultados do modelo sem as regras fonológicas	23
Tabela 2 -	Resultados do modelo com as regras fonológicas	23
Tabela 3 -	Erros ocorridos no conversor de Barbosa <i>et alia</i>	27
Tabela 4 -	Apresentação da transcrição fonética pelo SAMPA e as ocorrências nas cartas fonéticas	88
Tabela 5 -	Classificação das medidas multi-tópico para avaliação do sistema.	117
Quadro 1 -	Exemplo de transcrição fonética pelo sistema Petrus	28
Quadro 2 -	Relação dos símbolos fonéticos com as letras e os dígrafos consonantais	43
Quadro 3 -	Segmentos vocálicos para a variedade linguística potiguar	46
Quadro 4 -	Relação dos símbolos fonéticos com as letras e dígrafos vocálicos	47
Quadro 5 -	Representação dos ditongos decrescentes orais e nasais	48
Quadro 6 -	Representação dos ditongos crescentes orais e nasais	49
Quadro 7 -	Representação dos tritongos orais e nasais	49
Quadro 8 -	Símbolos utilizados nas regras fonológicas	54
Quadro 9 -	Máximo onset na língua <i>Ulwa</i>	65
Quadro 10 -	Constituição da sílaba por meio da linguagem Foma	69
Quadro 11 -	Visualização da composição de língua regular	77
Quadro 12 -	Principais símbolos utilizados no Foma	106
Quadro 13 -	Ataque complexo	110
Quadro 14 -	Dígrafos consonantais	112
Quadro 15 -	Conjunto de <i>corpus</i> ouro para a avaliação do sistema	118
Quadro 16 -	Execução do sistema na amostra <i>dicio_corpusbr.txt</i>	120

LISTA DE ABREVIATURAS E SIGLAS

AFI	Alfabeto Fonético Internacional
ALiB	Atlas Linguístico do Brasil
ALiPTG	Atlas Geolinguístico do Litoral Potiguar
CARTs	<i>Classification and Regression Trees</i>
CetenFolha	Corpus de Extractos de Textos Electrónicos NILC/Folha de S. Paulo
FSM	<i>Finite-State Machine</i>
FST	<i>Finite State Transducer</i>
G ₂ P	<i>Grapheme to phoneme</i>
HMM	Hide Markov Model
LTS	<i>Letter to Sound</i>
NILC	Núcleo Interinstitucional de Linguística Computacional
PB	Português Brasileiro
PE	Português Europeu
PER	<i>Phoneme error rate</i>
Regex	<i>Regular expression</i>
QFF	Questionário Fonético-Fonológico
QSL	Questionário Semântico-Lexical
QSM	Questionário Morfossintático
SAMPA	<i>Speech Assessment Methods Phonetic Alphabet</i>
STT	<i>Speech-To-Text</i>
SF	Sem a transcrição dos Fenômenos fonéticos
ST	Sem a marcação de Tonicidade
STF	Sem a marcação de Tonicidade e sem Fenômenos fonéticos
TTS	<i>Text-To-Speech</i>
WFSTs	<i>Finite state transducer</i>
WER	<i>Word error rate</i>
XFST	<i>Xerox Finite State Transducer</i>

LISTA DE SÍMBOLOS

*	Zero ou mais vezes
+	Uma ou mais vezes
	União
()	Opcionalidade
~	Complemento de um conjunto
" "	Espaço em branco é a concatenação entre elementos
:	Produto cruzado
.o.	Composição
->	Substituição direta da esquerda para direita
<	Precedente
>	Seguintes
/	Ignora o conjunto de elementos
...	Marcador de substituição
@	Contexto de restrição
_	Substituição ou restrição em contexto específico
.#.	Marcador de fronteira de palavras
#	Inserção de comentário
	Substituição direcional específica
0	<i>String</i> vazia
\$	Fator de um conjunto
^	Concatenação n-ária
Σ	Conjunto de elementos

SUMÁRIO

1	INTRODUÇÃO	11
2	SISTEMAS GRÁFICOS FÔNICOS PARA O PORTUGUÊS	16
2.1	Gráficos fônicos para o Português Europeu	16
2.1.1	<i>Sistema por regras linguísticas</i>	17
2.1.2	<i>Sistema por máquinas de estados finitos</i>	18
2.1.3	<i>Sistema por Máxima Entropia</i>	20
2.1.4	<i>Sistema híbrido Grafone</i>	22
2.2	Gráficos fônicos para o Português Brasileiro	24
2.2.1	<i>SiWeb</i>	24
2.2.2	<i>Tradutor grafema fonema baseado em autômatos adaptativos</i> ...	26
2.2.3	<i>Sistema de conversão baseado em regras</i>	27
2.2.4	<i>Petrus: Um conversor baseado em regras para PB</i>	27
2.2.5	<i>Nhenhém 1.0 - Ferramenta de suporte à pesquisa fonológica</i>	29
2.3	Considerações sobre os sistemas gráficos fônicos	30
3	FUNDAMENTOS LINGUÍSTICOS e COMPUTACIONAIS	34
3.1	Formas gráficas e formas fonéticas	34
3.1.1	<i>Formas gráficas da língua</i>	35
3.1.2	<i>Formas fonéticas</i>	37
3.1.2.1	<i>Alfabeto fonético</i>	38
3.1.2.2	<i>Símbolos fonéticos para variedade linguística potiguar</i>	39
3.1.2.2.1	Consoantes	40
3.1.2.2.2	Vogais	45
3.1.2.2.3	Encontros vocálicos: ditongos, tritongos e hiatos	48
3.2	Conceitos fonológicos	50
3.2.1	<i>Conceitos-chave</i>	50
3.2.2	<i>Regras e processos fonológicos</i>	53
3.2.2.1	<i>Regras fonológicas</i>	53
3.2.2.2	<i>Processos fonológicos</i>	55
3.2.2.2.1	Processo de assimilação	56
3.2.2.2.2	Processo de reestruturação silábica	59
3.2.2.2.3	Processo de neutralização	62
3.3	Estrutura silábica e tonicidade das palavras	63
3.3.1	<i>Estrutura silábica do português</i>	63
3.3.2	<i>Tonicidade das palavras</i>	70

3.4	Conceitos computacionais	72
3.4.1	<i>Língua formal, expressão regular e transdutores</i>	72
3.4.2	<i>Formas subjacentes e formas de superfície</i>	75
4	METODOLOGIA, IMPLEMENTAÇÃO E ANÁLISE DO SISTEMA	80
4.1	Base empírica	81
4.1.1	<i>Atlas Linguísticos do Centro-Oeste Potiguar</i>	82
4.1.2	<i>Atlas Linguísticos do Litoral Potiguar</i>	85
4.2	Potigrafone - conversor gráfico e fônico para a variedade linguística potiguar	88
4.2.1	<i>Transcrição fonética automática</i>	89
4.2.2	<i>Arquitetura e algoritmo do sistema</i>	93
4.2.3	<i>Implementação do sistema do sistema na linguagem Foma</i>	101
4.2.3.1	<i>Foma: comandos e operadores</i>	101
4.2.3.2	<i>Módulos do sistema</i>	109
4.3	Avaliação e análise do sistema	114
4.3.1	Corpus de referência	115
4.3.2	Medidas de avaliação do sistema	116
4.3.3	Avaliação e resultados do sistema	118
5	CONCLUSÃO	125
	REFERÊNCIAS	130
	APÊNDICE A - TRANSCRIÇÃO DOS CORPUS PELO ALFABETO FONÉTICO SAMPA	134
	APÊNDICE B – CÓDIGO FONTE DO POTIGRAFONE	141
	APÊNDICE C – EXECUTAR O CONVERSOR POTIGRAFONE	146
	ANEXO A - CORRESPONDÊNCIA ENTRE OS SÍMBOLOS DO SAMPA COM OS SÍMBOLOS DO ALFABETO FONÉTICO INTERNACIONAL	149
	ANEXO B - CLASSIFICAÇÃO E REPRESENTAÇÃO DOS SÍMBOLOS FONÉTICOS DO SAMPA	151
	ANEXO C - SÍMBOLOS FONÉTICOS USADOS PELO SISTEMA DE ESCRITA FONÉTICA SAMPA PARA O PORTUGUÊS	152
	ANEXO D - MATRIZ DE TRAÇOS DAS VOGAIS E CONSOANTES DO PB	153
	ANEXO E - ALGORITMO POTIGRAFONE	154

1 INTRODUÇÃO

Um transdutor de estados finitos é uma máquina construída por duas fitas: uma de entrada e uma de saída. Dessa forma, um transdutor traduz o conteúdo da entrada para o conteúdo da saída, reconhecendo a cadeia de caracteres de entrada e gerando outra cadeia na saída. Esse sistema é frequentemente utilizado nas pesquisas e aplicações de processamento de linguagem natural, em especial, na fonologia e na análise morfológica. Uma maneira comum de utilizá-lo é fazendo uso de operações em cascata, em que o sistema está combinado em um único transdutor com aplicação repetida.

Há sistemas que fazem a conversão de grafema para fonema e são comumente chamados de *Grapheme to Phoneme* (G₂P) ou de *Letter to Sound* (LTS). Esses conversores são aplicados no pré-processamento dos sistemas da tecnologia de fala e como ferramenta de pesquisa em Linguística. Também fazem parte do pré-processamento para que o sistema de *síntese* de fala e reconhecimento de voz consigam “adivinhar” a pronúncia correta das palavras (TEIXEIRA *et alia*, 2006). A síntese de fala (*Text-To-Speech - TTS*) e o reconhecimento de voz (*Speech-To-Text - STT*) estão presentes em serviços de telecomunicações, na Educação, em controle de equipamentos industriais, em alarmes para situações de risco, em ajuda às pessoas com algum tipo de necessidade especial (especialmente em pessoas com deficiência visual ou auditiva), em controles de lista de espera em hospitais, sinalização de bagagens em transportes públicos, sistemas de navegação GPS, *smartphones*, etc.

Além do uso no pré-processamento, os conversores são muito úteis como instrumento de auxílio aos linguistas, mais especificamente, aos foneticistas que trabalham com transcrição fonética de *corpora* de língua portuguesa e, também, aos lexicógrafos na construção da transcrição fonética dos verbetes de dicionários. A transcrição manual de *corpora*, tradicionalmente feita por peritos em fonética, está associada aos inconvenientes de um processo muito moroso, carente de procedimentos *standard* amplamente aceitos e, muitas vezes, mais sujeito a erros humanos e à falta de coerência na aplicação dos critérios entre anotadores.

A conscientização desses problemas sublinha a necessidade de se recorrer a sistemas que executem de forma automática o processo de transcrição, tornando possível a anotação de grandes quantidades de material de fala através da

aplicação de um conjunto fixo de critérios objetivos, sem intervenção humana. Claro que os resultados de um conversor não têm a mesma precisão dos da transcrição manual; no entanto, uma vez que os erros são de natureza sistemática e pontual, a correção é feita com maior facilidade. Acreditamos que esses motivos já são suficientes para se compreender os benefícios que o desenvolvimento de novos conversores pode trazer para o campo da pesquisa linguística e do desenvolvimento tecnológico.

Este trabalho, portanto, tem como objetivo apresentar o desenvolvimento do conversor *Potigrafone* - um transdutor de estados finitos - que faz a transcrição automática dos grafemas para os sons (fones) da variedade linguística potiguar. Destacamos dois fatores motivadores para o seu desenvolvimento: a inexistência de sistema para a variedade potiguar e a presença de dados empíricos que dê consistência à transcrição do sistema. Outros sistemas computacionais foram e estão sendo desenvolvidos com outras abordagens e para outras variedades linguísticas como, por exemplo, o sistema Petrus que foi desenvolvido numa abordagem de regras linguísticas para a variedade paulista.

Até agora, porém, existem poucos sistemas e algoritmos implementados para o português brasileiro e nenhum deles foi desenvolvido para a variedade linguística falada no Rio Grande do Norte – RN. Sendo assim, este trabalho é pioneiro na transcrição automática do registro para a fala potiguar, mas muitos de seus aspectos servem, também, para outras variedades linguísticas do Brasil. É o caso do fenômeno da nasalidade que ocorre na maioria das variedades linguísticas faladas no nordeste. Além disso, esse transdutor preenche uma lacuna deixada por outros conversores para a transcrição fonética e fonológica automática do português, pois, como destacamos no capítulo 2, esses conversores não fazem a transcrição fonética das palavras escritas, considerando, na transcrição, os fenômenos fonéticos que, geralmente, estão presentes nas variedades linguísticas, como a nasalidade, monotongação, ditongação, etc.

Nesse trabalho, então, respondemos à seguinte pergunta: Os conversores de grafema para fone existentes para o português fazem a transcrição fonética apresentando marcas dos fenômenos fonéticos? De acordo com o que apresentamos no próximo capítulo, acreditamos que os conversores atuais desenvolvidos para essa língua não fazem a transcrição fonética considerando muitos dos fenômenos fonéticos presentes nas variedades linguísticas. Quando

fazem a transcrição fonética, transcrevem as formas gráficas para o sistema fonológico ou abordam uma transcrição restrita à forma mais próxima da variedade padrão da língua.

Outro aspecto importante a destacar é que a transcrição automática desse transdutor tem como ponto de partida os processos fonológicos presentes no *Atlas Linguístico do Centro-Oeste Potiguar* e no *Atlas Geolinguístico do Litoral Potiguar* (ALiPTG). Esses atlas nortearam a compreensão dos processos fonológicos presentes no dialeto potiguar e contribuem para que a descrição fonética do sistema não seja baseada apenas no conhecimento linguístico sobre a variedade. Além disso, os atlas sistematizam as variantes linguísticas e delimitam as áreas dialetais desse estado, oferecendo dados consistentes para o desenvolvimento do transdutor.

Por isso, definimos como principais objetivos específicos:

- identificar e separar os padrões fonéticos da variedade linguística potiguar presentes nos *corpus* do *Atlas Linguístico do Centro-Oeste Potiguar* e do *Atlas Geolinguístico do Litoral Potiguar*;
- apresentar os padrões fonéticos no alfabeto fonético *Speech Assessment Methods Phonetic Alphabet* (SAMPA);
- avaliar o desempenho do sistema em *corpus* de textos escritos de língua portuguesa, para compreender quais fenômenos o sistema executa com maior propriedade e quais os de menor confiabilidade.

Destacamos que esta tese se fundamenta nos pressupostos teóricos da fonologia computacional como uma aplicação das técnicas formais e computacionais para a representação e para o processo de informações fonológicas.

O trabalho está estruturado em cinco capítulos, sendo que o primeiro está reservado a esta introdução.

No capítulo dois, mostramos o estado da arte no que diz respeito à transcrição fonética e fonológica sob três pontos fundamentais: a abordagem, o alfabeto fonético e o sistema de entrada e saída dos conversores. Esse capítulo foi estruturado da seguinte forma: os conversores para o PE (Português Europeu) e os conversores para o PB (Português Brasileiro).

No capítulo três, apresentamos os conceitos linguísticos da teoria de base

e os conceitos computacionais que norteiam a construção do Potigrafone. Nesse capítulo, distinguimos as formas de entrada das formas de saída do transdutor, ou seja, as formas gráficas das palavras e as formas fonéticas. Como veremos, a expressão "formas gráficas" se refere à escrita ortográfica das palavras e não à escrita grafemática. Além disso, os conceitos estão relacionados com as aplicações da fala potiguar, através das tabelas que listam os sons consonantais, os vocálicos e os encontros vocálicos.

Apresentamos, também, os conceitos-chave da fonologia, a saber, fone, fonema, alofone, além dos conceitos de regras e processos fonológicos e de estrutura silábica e de tonicidade das palavras. A exposição desses conceitos deixa o texto mais didático e ajuda-nos a compreender melhor a fonologia computacional como uma aplicação das técnicas formais e computacionais para a representação e processo de informações fonológicas (BIRD, 2009). Aqui destacamos que os transdutores de estados finitos são um exemplo bem-sucedido quando a informação fonológica é tratada como uma cadeia de símbolos, ou seja, como as cadeias superior e inferior estão relacionadas de forma que as letras representam os sons da fala.

No capítulo quatro, descrevemos a metodologia empregada no desenvolvimento do transdutor Potigrafone, mostrando a forma de identificação e separação dos padrões fonológicos presentes nas cartas fonéticas dos Atlas, a definição de alguns critérios pertinentes às entradas (*inputs*) do sistema. Descrevemos também o desenvolvimento do transdutor Potigrafone que executa a transcrição fonética das formas gráficas para a variedade linguística potiguar, destacando o tipo de transcrição fonética e os diacríticos para a marcação da tonicidade, da nasalização e da palatalização das palavras; a relação da forma gráfica com a forma fonética e suas respectivas peculiaridades; a constituição do algoritmo e arquitetura do sistema e a implementação dos módulos utilizando os comandos e operadores da linguagem de programação Foma. Mostramos, além disso, a avaliação do sistema feita com base em duas métricas: a *Exact Match Ratio* e *Labelling FScore*. Essas métricas são mais adequadas à avaliação de sistemas com mais de uma saída e apresentam resultados mais precisos sobre os erros e acertos. Elas fazem a classificação multiníveis e avaliam através da correspondência de rótulos e de classes parcial ou total, considerando todas as classes e seus rótulos como equivalentes. A classificação dos erros norteará a implementação de novas

regras.

No último capítulo, apresentamos algumas conclusões sobre o sistema desenvolvido e algumas considerações referentes ao desempenho desse transdutor e as pesquisas futuras concernentes à transcrição automática.

2 SISTEMAS GRÁFICOS E FÔNICOS PARA O PORTUGUÊS

Neste capítulo, mostramos o estado da arte dos estudos sobre conversores desenvolvidos para o português sob três pontos fundamentais: a abordagem, o alfabeto fonético e o sistema de entrada e saída. Esses conversores são rotulados pela sigla G_2P (*Grapheme to Phoneme*) e/ou pela sigla LTS (*Letter-to-Sound*). Os primeiros enfatizam a conversão de um conjunto de grafemas para uma sequência de símbolos fonológicos de uma determinada língua, enquanto os segundos executam a transcrição da sequência de grafema para os símbolos fonéticos que representam uma variedade linguística, tentando transcrever a palavra mais próxima possível da pronúncia.

A escolha do primeiro ou do segundo tipo de conversor interfere no resultado final, na forma de transcrição. Por exemplo, por um lado, o Grafone faz a transcrição fonológica da palavra, isto é, a transcrição feita por esse sistema não apresenta símbolos que representam a fala de uma variedade específica. Por outro lado, o Petrus é um sistema que transcreve a palavra considerando a sua pronúncia para a variedade linguística paulista. Então, dizemos que o primeiro conversor se enquadra como um G_2P e o segundo como um LTS.

No entanto, por haver muita similaridade entre esses tipos de conversores gráficos e fônicos, ao ponto de, às vezes, haver autores que confundem um com o outro, não faremos distinção quanto à classificação do sistema, se o sistema pertence ao primeiro ou ao segundo tipo. Destacamos, também, que não é relevante a classificação dos conversores para este trabalho.

Pensando assim, dividimos este capítulo em três partes: a primeira e a segunda mostram alguns sistemas desenvolvidos para o Português Europeu (doravante PE) e para Português Brasileiro (doravante PB), respectivamente; a terceira apresenta considerações sobre o estado da arte relativo à conversão de grafema para fonema.

2.1 Sistemas gráficos e fônicos para o Português Europeu

Apresentamos a seguir quatro conversores para o PE e as abordagens que fundamentam a implementação básica de cada sistema. O primeiro é um sistema regido por regras linguísticas, desenvolvido por Braga, Coelho e Resende

(2006); o segundo é um conversor construído por máquinas de estados finitos, implementado por Caseiro *et alia* (2002); o terceiro é um sistema probabilístico pelo método de máxima entropia, por Barros e Weiss (2006) e, por último, o sistema híbrido *grafone*, por Veigas, Candeias e Perdigão (2011). Além desses, poderíamos expor outras propostas como o sistema baseado em redes neuronais, desenvolvido por Trancoso *et alia* (1994) e o sistema por *Classification and Regression Trees* (CARTs), por Oliveira *et alia* (2001). Todavia, restringimo-nos à apresentação dos quatro primeiros.

2.1.1 Sistema por regras linguísticas

Braga, Coelho e Resende (2006) construíram um algoritmo que faz as transcrições fonológica e fonética baseadas em regras linguísticas e é adequado à variedade linguística do PE. No entanto, os autores alegam que o algoritmo pode ser muito útil para outras variedades de Português, inclusive para o PB. Essas regras foram implementadas por meio de padrões fonológicos e utilizando, na transcrição dos grafemas, os símbolos fonéticos do alfabeto fonético SAMPA¹ para o PE.

O conjunto completo de regras de transcrição para cada grafema usado em PE foi construído com base nas ocorrências de contexto. O algoritmo verifica se a sequência do grafema <am>, por exemplo, num dado contexto, representa o fonema vocálico nasal /a~/, isto é, se a sequência está dentro dos padrões para essa representação fonológica; em caso afirmativo, o sistema executa a transcrição. O sistema considera todos os padrões gráficos, incluindo os grafemas estrangeiros e os ditongos crescentes e decrescentes que incorporam as regras dos grafemas vocálicos “i” e “u”, como nas palavras “cárie” e “ouvir”.

Ao conceber essas regras, eles tentaram projetá-las, tanto quanto possível, de acordo com os padrões gráficos, de forma a reduzir as dependências de outros módulos de processamento de texto. Em nota de rodapé, os autores destacaram que houve apenas duas exceções relativas aos padrões de regras: o marcador de sílaba tônica e o separador silábico.

Partindo daí, eles implementaram as regras de transcrição agrupando-as de acordo com os caracteres comuns. Realizaram vários testes que foram

¹ O anexo C apresenta os símbolos do alfabeto SAMPA para o português.

ordenados por frequência dos grafemas na língua e testaram das mais frequentes até a menos frequente. Esse teste foi feito a partir da análise de cerca de 1 milhão de caracteres contidos no *corpus* FEUP IPB, que é um banco de dados de sinal sonoros rotulados para o português padrão europeu. Com este procedimento, os desenvolvedores conseguiram reduzir o número de 146 regras para apenas 100.

Feitos esses procedimentos, o algoritmo foi testado usando textos de jornais *online*, contendo um número total de 10.884 caracteres, distribuídos entre 1802 palavras que foram usadas como a entrada do sistema. Os fonemas transcritos pelo sistema foram verificados manualmente e foram convertidos corretamente 98,80% dos caracteres. Os principais erros cometidos pelo sistema foram decorrentes das palavras homógrafas, com a transcrição das vogais <e> e <o> e com os nomes próprios.

Esse conversor baseado em regras, se comparado com outros, apresenta uma boa performance, mas, como apontaremos no tópico 2.3, essa abordagem possui alguns inconvenientes que os autores não deixaram explícitos no seu trabalho. Talvez isso tenha ocorrido devido à limitação no espaço de publicação do artigo e a forma sucinta e tabelar com que eles mostram as regras do sistema.

2.1.2 Sistema por máquinas de estados finitos

Caseiro *et alia* (2002) apresentam um conversor de grafema para fonema baseado em máquina de estados finitos, construído a partir das regras do Dixi - sistema de regras já existente para o PE, contendo cerca de 200 regras e, no caso das regras de atribuição de tonicidade e de outras regras restantes, usaram o *Speech Compiler for Your Language* (Scyla) – compilador de regras desenvolvido pelo *Centro Studi e Laboratori Telecomunicazioni* (CSELT).

Os autores afirmam que a estrutura multi-nível desse compilador permitiu que cada procedimento servisse para ter acesso, simultaneamente, a dados resultantes dos procedimentos anteriores, de modo que as regras poderiam se referir, ao mesmo tempo, a vários níveis, tais como o nível de grafema, o nível de fonema, o nível de *sândi*, etc.

O primeiro desafio mencionado pelos desenvolvedores, dentro dessa estrutura, foi o de converter as regras do Dixi e do Scyla em um conjunto de regras de estados finitos, no padrão das formas gerativas do tipo $A \rightarrow B / C _ D$ em que A,

B, C e D podem ser convertidas em expressões regulares referentes a um ou mais níveis. Eles apontaram também que o uso dessas regras leva à dependência de outras regras desnecessárias devido à substituição dos grafemas pelos fonemas.

No caso do Dixi, no entanto, alguns desses problemas foram evitados quando implementaram regras do sistema no nível de grafema – como marcação de tonicidade. No nível do fonema, foram implementadas camadas diferentes da estrutura de dados multiníveis. Assim, não existem estágios intermediários de representação e nenhuma regra cria ou destrói o contexto necessário para a aplicação de uma outra regra.

Além disso, eles utilizaram as regras de reescrita com as seguintes condições: (1) a sequência grafema g_1, g_2, \dots, g_n é traduzida em $g_{1_}, g_{2_} \dots g_{n_}$, onde o *underline* _ é um símbolo vazio usado como suporte para os fonemas; (2) cada regrasubstituirá o _ por um fonema correspondente ao grafema anterior. Desse modo, o contexto das regras pode agora referir-se livremente aos grafemas e (3) a última regra remove todos os grafemas, deixando apenas uma sequência de fonemas. A língua de entrada e a de saída dos transdutores de regra é formada, portanto, de um subconjunto de grafemas e de fonemas, respectivamente, no qual ambos se sobrepõem.

Tendo em vista a construção multinível, Caseiro *et alia* (2002) dividiram o sistema em três partes: o conjunto de regras de especificações, as fases do grafema para fonema e as regras para o transdutor.

O conjunto de regras de especificação se relaciona com os símbolos não terminais. Essas regras são organizadas em várias fases, cada uma representada por transdutores que podem ser compostos para construir o sistema completo. Cada fase exhibe as seguintes funções: (1) a fase da tonicidade consistindo de 27 regras que marcam a vogal tônica da palavra; (2) a da introdução dos fonemas nos espaços reservados ao fonema vazio após cada grafema e (3) a do prefixo no léxico. Esta fase consiste de 92 regras de pronúncia para palavras compostas.

A compilação dessas regras resulta em um grande número de transdutores de estados finitos (*Finite State Transducer* - WFSTs) que precisam ser compostos para construir um único transdutor grafema para fone. Após a composição, eles obtiveram um pequeno conjunto de 8 transdutores agrupados com um tamanho médio de 410k. O conjunto de regras de introdução dos fonemas e de remoção dos grafemas compostas com outros WFSTs formam o conjunto final de 10

transdutores.

A avaliação do sistema foi feita em duas etapas. A primeira consistiu na comparação da abordagem de regra baseada na WFST com o sistema baseado em regras DIXI anterior. Nessa etapa, os autores mostram que o transdutor obteve desempenho menor do que o Dixi. A segunda etapa foi executada após o alinhamento dos grafemas aos fonemas e a construção de um modelo n -grama. Na maioria dos casos, o alinhamento é feito com base na correspondência de um grafema para um fone, mas também foram modelados alguns casos de dois grafemas para um fone e, em outros casos, de um ou dois grafemas para um mais fones.

Segundo os autores, esse alinhamento serviu para a modelagem das sequências usando um modelo n -grama. Este modelo baseia-se na probabilidade de uma correspondência entre os grafemas e os fones particularmente. O modelo foi primeiro convertido em um “aceitador” de estados finitos (WFSA) ao longo de pares de símbolos e, em seguida, em um transdutor de estado finito que transforma cada par de símbolos numa entrada e numa saída. Nessa fase, o desempenho do sistema, após a abordagem n -grama, apresentou bons resultados, cerca de 96% para as palavras sem marcação de sílaba tônica e de 91% para as palavras com marcação de sílabas tônicas.

Da mesma forma que ocorre com o conversor de Braga, Coelho e Resende (2006), esse transdutor apresenta bom desempenho para a transcrição de palavras. No entanto, as regras de reescrita, quando mal utilizadas, deixam o sistema com pouca eficácia e com muita sobrecarga de regras inoperantes. Por um lado, a nossa proposta se baseia nesse tipo de abordagem e sabemos que é necessário muita cautela na aplicação das regras. Todavia, por outro lado, escolhemos o desenvolvimento do sistema com base no reconhecimento da estrutura silábica que contribui na redução e no uso de regras que auxiliam na marcação da sílaba tônica e nas regras de transformação.

2.1.3 Sistema por Máxima Entropia

Barros e Weiss (2006) apresentam um sistema probabilístico baseado em modelos de máxima entropia. Esse sistema executa a transcrição gráfica e fônica, insere a acentuação e faz a separação silábica das palavras para o PE.

Esses autores apresentam duas justificativas para o uso desse modelo num sistema gráfico e fônico. A primeira se fundamenta no fato de que os sistemas probabilísticos são mais flexíveis quanto ao som sintético natural do discurso contínuo, uma vez que os modelos estatísticos podem ser treinados de tal maneira que são proferidos no ritmo da fala, destinando-se ao uso em determinada aplicação em síntese de fala. A segunda motivação é que a abordagem de máxima entropia já é bem conhecida na resolução de ambiguidade no processamento de linguagem natural, em que muitos problemas podem ser reformulados como um problema de classificação, que é justamente o que ocorre no caso dos sistemas gráficos e fônicos. Há símbolos gráficos que apresentam uma ou mais formas de transcrição fonológica.

A tarefa básica apontada pelos autores de um sistema probabilístico foi a de incluir um contexto para as transcrições e prever uma classe correta para o fenômeno transcrito, com o objetivo de estimar uma função de X em Y , na qual se prevê um objeto x que pertence a X e um y que pertence a Y . No sistema, cada fonema do inventário representa uma classe; assim, há 38 classes que correspondem ao número de fonemas para o PE. Nesse caso, o valor de Y representa as classes predefinidas para cada tarefa de prever a relação de grafema para fonema.

Eles acrescentam que o valor de X consiste de traços linguísticos aos quais foram incluídos os contextos, e a entrada, resultante da classificação, é um vetor de traços, contendo o próprio objeto que tem que ser classificado, bem como o contexto. Para a atribuição do valor de X , eles fizeram uma classificação binária em que a classe é verdadeira para as sílabas tônicas e falsas para as não tônicas. Essa tarefa também serviu para a separação silábica.

Baseando-se nos valores de X e Y , o conversor foi dividido em três modelos distintos e, para cada um desses modelos, foi utilizado um *corpus* de treino: a tarefa de transcrição foi treinada em um *corpus* que consistia de 7.352 palavras ortográficas com suas respectivas transcrições fonéticas, gerando o modelo de transcrição. A tarefa de silabação foi treinada em um *corpus* de 4.283, gerando o módulo de silabação e a tarefa de acentuação foi treinada em um *corpus* que continha 4.219 palavras fonéticas, que gerou o módulo de acentuação.

Barros e Weiss (2006) apresentam a avaliação dos modelos separadamente. O cálculo do desempenho foi mostrado com base no número de

elementos classificados corretamente, dividido pelo número de elementos gerais e multiplicado por cem, obtendo os seguintes resultados: O modelo de transcrição obteve o resultado de 88,94 %, o de silabação obteve 97,64 % e o de tonicidade 85,57%.

Em relação a essa abordagem, diríamos que esse sistema obteve bons resultados no modelo que faz a silabação das palavras, mas os resultados de transcrição e tonicidade ficam aquém dos outros dois modelos já mencionados. Outro fato é que essas três tarefas obtiveram uma média abaixo da de outros sistemas: 91% de acerto. Isso talvez se deva ao fato de que a extensão do *corpus* de treino, utilizado para criar o modelo probabilístico, seja muito pequeno. Uma das dificuldades que podemos apontar sobre os sistemas de probabilidades, em língua portuguesa, é a falta de grande quantidade de dados linguísticos prontos para o treino do sistema. Além disso, esses sistemas não captam um contexto suficientemente amplo de forma a impedir que a estrutura fonológica da língua seja violada.

2.1.4 Sistema híbrido Grafone²

O Conversor de grafema para fonema do Português Europeu – Grafone – foi desenvolvido pela equipe do Instituto de Telecomunicação, no Laboratório de Processamento de Sinais em Coimbra, Portugal. Esse laboratório vem pesquisando e desenvolvendo sistemas de *reconhecimento de fala* e *síntese de fala*, os quais exigem o desenvolvimento de um conversor de grafema para fonema de alto nível.

Segundo Veiga, Candeias e Perdigão (2011), o Grafone faz a conversão gráfica e fônica a partir de uma abordagem híbrida, utilizando um modelo probabilístico de língua e um conjunto de regras fonológicas.

Como aponta os autores, a construção do modelo de língua requereu a existência de "um dicionário fonológico, necessário para estimar as probabilidades dos padrões encontrados" (VEIGA, CANDEIAS E PERDIGÃO, 2011, p. 40). O treinamento desse modelo foi feito em um dicionário de pronúnciação, intitulado *dic_CETEMP_40k*, contendo 41.586 vocábulos derivados do *corpus* CETEMPúblico. Esse *corpus* apresenta um conjunto de 180 milhões de palavras, advindo de uma

² O sistema Grafone está disponível em: <http://www.co.it.pt/~labfala/g2p/index.html>.

coleção de extratos do jornal Público. Com base na criação desse dicionário, foram construídos mais quatro dicionários intitulados dic_CETEMP_40k_alinhado; dic_CETEMP_40k_acentuado; dic_CETEMP_40k_alinhado-digrafos e dic_CETEMP_40k_acentuado_alinhado_digraf.

“Para testar o modelo estatístico, cada um destes dicionários foi particionado em 5 dicionários de treino e 5 dicionários de teste, de forma rotativa. O dicionário inicial foi dividido em 5 partes, cada uma com 20% dos vocábulos (8317), escolhidos de forma aleatória.” (VEIGAS, CANDEIAS e PERDIGÃO, 2011, p. 48).

Além disso, os autores adicionaram ao modelo as regras fonológicas de acentuação vocálica, ou seja, as regras para o reconhecimento do núcleo de sílaba tônica de cada vocábulo e as regras para a identificação da correspondência exata entre um grafema e o respectivo fonema, de acordo com o contexto.

Feito isso, eles verificaram o desempenho do sistema por meio de duas taxas médias de erros de conversão verificados nos dicionários de teste: taxa média de erro de fonemas (PER – "*phoneme error rate*") e taxa média de erro de vocábulos (WER – "*word error rate*"). As tabelas abaixo sumarizam os resultados obtidos através dos modelos com as regras fonológicas e sem as regras fonológicas, respectivamente.

Tabela 1 - Resultados do modelo sem as regras fonológicas

n-grama	2	3	4	5	6	7	8
WER (%)	35.1	15.5	7.90	5.96	6.08	6.51	7.01
PER(%)	4.69	1.86	0.95	0.72	0.74	0.79	0.86

Fonte: Veiga, Candeias e Perdigão (2011, p. 48).

Tabela 2 - Resultados do modelo com as regras fonológicas

n-grama	2	3	4	5	6	7	8
WER (%)	9.73	4.70	2.60	2.31	2.42	2.58	2.82
PER(%)	1.27	0.60	0.33	0.30	0.31	0.33	0.36

Fonte: Veiga, Candeias e Perdigão (2011, p. 48).

Os autores acrescentam que mais de 80% dos erros são provocados por confusão gerada pelo grau de abertura dos fonemas que devem corresponder aos grafemas <o> , <e> e <a>, os quais podem ser pronunciados [o] ou [O], [e] ou [E], [ɔ] ou [a], dependendo do contexto. Por exemplo, a palavra escrita como <peso> tem o som fechado /"pezu/ quando é substantivo e tem o som aberto /"pEsu/ quando é

verbo.

Ressaltamos que, dentre os sistemas mencionados acima para o PE, o Grafone é o único sistema disponível para consulta dos usuários. Os outros estão restritos aos pesquisadores ou são utilizados como ferramentas de pré-processamento da fala ou de reconhecimento de voz. Enfatizamos, também, que esse conversor faz a transcrição fonológica nos alfabetos SAMPA e AFI, mas não executa a transcrição fonética.

2.2 Sistemas gráficos fônicos para o Português Brasileiro

Nos subtópicos a seguir, apresentamos alguns conversores gráficos fônicos para o PB como o Silweb, desenvolvido por Matte, Meireles e Fráguas (2006); o tradutor baseado em autômatos adaptativos, implementado por Shibata (2008); o conversor baseado em regras, construído por Barbosa *et alia* (2003); o sistema Petrus, desenvolvido por Marquiefável, Bokan e Zavaglia (2014) e o sistema Nhenhem 1.0, por Vasilévski (2008).

2.2.1 *SilWeb*³

Matte, Meireles e Fráguas (2006) apresentam um analisador fonológico chamado *SilWeb* para o regresso de composição prosódica e constitutiva como o acento, a informação fonêmica e o número de sílabas de uma transcrição fonológica das palavras escritas em Português do Brasil. Essa ferramenta computacional foi desenvolvida pelo Laboratório de Fonética e Psicolinguística (LAFAPE) da Universidade Estadual de Campinas (Unicamp) e nasceu da necessidade de analisar textos escritos do Corpus de Extratos de Textos Eletrônicos NILC/Folha de São Paulo (CENTEN/Folha), no âmbito do projeto Integrando Parâmetros Contínuos e Discretos em Modelos do Conhecimento Fônico Lexical. O uso eficaz desse sistema depende de uma prévia transcrição fonológica que é feita pelo Ortofon – sistema de transcrição fonológica.

Segundo esses autores, o Ortofon executa algumas tarefas de transcrição das palavras escritas e as fornece ao Silweb. Essas tarefas são a marca de

³ Os autores afirmam que o sistema está disponível no sítio: <http://silweb.sourceforge.net/>. Todavia, não conseguimos acessá-lo e nem pudemos fazer o descarregamento do sistema.

tonicidade nas palavras e o fechamento do <e> e do <o> em final de palavras e nos monossílabos átonos. No entanto, esse sistema não faz a separação dos encontros consonantais; não executa a mudança do <l> pós-vocálico para o glide /ʌ/ e dos dígrafos <lh> e <nh> nas consoantes palatais /d/ e /ɲ/, respectivamente. Também não converte a redução da vogal átona <a> nos monossílabos.

Como aponta esses autores, o Silweb acessa o ortofon para verificar a transcrição fonológica das palavras e faz uma análise mais profunda do sistema fonológico, pois esse sistema verifica o tamanho, os finais das sílabas e a quantidade de caracteres de cada palavra e recupera cada caractere e cada sílaba analisados. Além disso, o sistema retorna a tonicidade da sílaba na palavra, o número de sílabas das palavras, a classe acentual da palavra, conforme seja oxítona, paroxítona e proparoxítona e, também, retorna as máscaras de distinção entre uma análise que considera apenas as vogais e uma análise que considera as semivogais.

Esse sistema se estrutura com base em três entidades principais: a palavra, a sílaba e o caractere. A primeira entidade realiza a tarefa de obter a classe acentual da palavra de entrada para classificá-la em oxítona, paroxítona ou proparoxítona. Para a realização dessa tarefa, o sistema depende de outras duas: uma que ocorre no âmbito da sílaba e a outra no âmbito do caractere. Para obter a classe acentual, o sistema acessa o número de sílabas e a posição delas na palavra, para depois acentuá-las. Essas duas tarefas são tributárias às análises realizadas nos caracteres. Essas três entidades realizam tarefas em cadeias que vão dos caracteres para as palavras, porém a tarefa de base, ou seja, a tarefa nos caracteres é feita no Ortofon e é passada para o Silweb.

Esse conversor apresenta muitas opções de verificação da análise fonológica como a quantidade de sílabas e de caracteres e a classe acentual da sílaba tônica, etc. No entanto, acreditamos que uma de suas desvantagens é o fato de ele ser dependente de outro sistema, o ortofon, para fazer a transcrição fonológica.

Outro fato importante é que os autores apresentam link: <https://sourceforge.net/projects/silweb/>, para que os usuários possam utilizá-lo, mas, por meio desse endereço, não é possível em que não é possível acessá-lo e/ou encontrá-lo para o descarregamento no computador.

2.2.2 Tradutor grafema/fonema baseado em autômatos adaptativos

Shibata (2008) apresenta a criação de um método baseado em autômatos adaptativos para a tradução de grafema-fonema⁴ na variedade linguística de São Paulo e sobre uma possível aplicação deste método num *software* de tradução texto-voz.

Para esse autor, o autômato adaptativo mimetiza o comportamento humano no tratamento de regras de tonicidade, separação de sílabas e as influências que as sílabas exercem sobre suas vizinhas. Esses aspectos tornam o método facilmente utilizável para outras variações da língua portuguesa, considerando que essas características são invariantes em relação à localidade e à época da variedade linguística.

Partindo dessas características, o autor nos apresenta uma arquitetura utilizada para representar a estrutura dos autômatos adaptativos desenvolvido por ele e que se baseia em três classes: *Automaton*, *Configuration* e a *Interface Action*. Ele acrescenta que a primeira classe possui o método chamado *recognize* que "dispara o processo de reconhecimento de uma sequência de símbolos e retorna o resultado na forma de um valor *boolean*." (SHIBATA, 2008, p. 69). A segunda representa uma configuração do autômato durante o processo de execução e a terceira representa ações adaptativas utilizadas para alterar o conjunto de transições de uma configuração do autômato. Nesses três elementos, encontram-se as regras de funcionamento do simulador.

Além desse modelo, o autor afirma que foi necessário criar um método para a escolha da representação fonética correta, em caso de ambiguidade, e também o uso de dois *softwares*, um para a simulação de autômatos adaptativos e outro para a tradução grafema-fonema de palavras.

Por fim, o autor testou e avaliou o tradutor tomando um conjunto de palavras que serviu de base para a transcrição da língua contemporânea falada na cidade de São Paulo. Para essa variedade, o modelo apresentou resultados satisfatórios, superando 95% de acerto na tradução grafema-fonema de palavras e chegando a 90% de acerto, levando em consideração a resolução de dúvidas geradas por palavras que podem possuir duas representações sonoras.

O autor deixa claro que esses autômatos adaptativos podem ser

⁴ O autor utilizou o termo "fonema" para se referir ao "fone" da variedade linguística paulista.

aplicados a outras variedades linguísticas, mas não fica claro como fazer isso ou que caminhos para obter os resultados.

2.2.3 Sistema de conversão baseado em regras

Barbosa *et alia* (2003) descrevem as regras de um algoritmo que faz a transcrição de grafema para fone da variedade linguística falada no Rio de Janeiro. Eles afirmam que os estudos se concentram na ideia de propor regras para lidar com palavras estrangeiras, nomes, ditongos e para aperfeiçoar a transcrição da fricativa velar surda [s] e da alternância vocálica das vogais abertas [O] e [E] e fechadas [o] ou [e] e, também, propor as regras para diferenciar os níveis de nasalização e de acentuação para algumas unidades acústicas.

Esse algoritmo foi testado no banco de dados do CETEN-Folha. Os fones transcritos foram verificados e 98,43% foram transcritos corretamente em um conjunto de 7.805 fones de 1.364 palavras.

Os autores apresentam a tabela 3, replicada abaixo, que mostra o desempenho desse sistema de regras para a transcrição de um determinado grafema para o fone.

Tabela 3 - Erros ocorridos no conversor de Barbosa *et alia* (2003)

Tipos de erros	Ocorrência	Ocorrência %
[e] or [E]	22	0.28%
[o] or [O]	18	0.23%
Fones incorretos para as palavras estrangeiras	31	0.40%
Ditongos	35	0.44%
Fones acrônimos incorretos	17	0.22%

Fonte: Barbosa *et alia* (2003, p. 30).

2.2.4 Petrus: um conversor baseado em regras para PB

Marquiefável, Bokan e Zavaglia (2014) apresentam o sistema Petrus⁵ que faz a transcrição fonética automática das sequências de grafemas do português para o dialeto paulista, utilizando os símbolos do Alfabeto Fonético Internacional. Os autores explicam que “Por dialeto paulista, deve-se entender a fala de pessoas cultas oriundas do Estado de São Paulo.” (MARQUIAFÁVEL e ZAVAGLIA, 2011, p.

⁵ Disponível em: <http://www.nilc.icmc.usp.br/petrus>.

156) e acrescentam que não há qualquer sistema desse tipo para dialeto padrão de São Paulo.

O Petrus é um ambiente computacional ligado ao Núcleo Interinstitucional de Linguística Computacional (Nilc/ICMC) e que dispõe de uma interface gráfica de fácil utilização, servindo de apoio à transcrição fonética automática de palavras isoladas ou de conjunto de palavras.

Segundo esses autores, esse sistema é composto por seis módulos. O primeiro é responsável pela transcrição fonética das palavras homógrafas heterofônicas, como nas palavras 'g[o]sto (substantivo) e 'g[O]sto (verbo). Esse recurso linguístico foi construído usando um dicionário com pares de palavras heterofônicas com quase 1.000 homográficos transcritos.

O segundo módulo executa a identificação da presença de prefixo em uma palavra antes de transcrevê-la, uma vez que os prefixos em PB têm um comportamento específico no sentido de alternância vocálica, e o recurso linguístico utilizado foi um algoritmo baseado com 145 regras.

O terceiro módulo identifica o diacrítico [ˈ], a vogal acentuada na palavra que está sendo analisada, usando um algoritmo baseado em regras.

O quarto módulo identifica e marca os limites de sílabas e usa um algoritmo com 25 regras baseados em *Hidden Markov Model* (HMM).

O quinto módulo executa a etiquetagem das palavras - *Part-de-Speech tagger* - usando UNITEX.

No último módulo, toda essa informação linguística é utilizada para transcrição fonética das palavras escritas, por meio de um conjunto de 160 regras, sendo que 62 delas fazem a transcrição das consoantes e 98 fazem das vogais.

Com a utilização desses módulos, o Petrus faz a transcrição das palavras, separa as sílabas e classifica a palavra transcrita quanto à classe gramatical. Podemos ver um exemplo de transcrição com a palavra <peso> na quadro 1, a seguir:

Quadro 1 - Exemplo de transcrição fonética pelo sistema *Petrus*

Grafema	Transcrição fonética
Peso	['pezʊ] Syllable boundaries: pe.so Part of speech: Substantivo

['pezʊ]
Syllable boundaries: pe.so
Part of speech: Verbo

Fonte: Dados verificado on line pelo Petrus, acesso em:10/11/2015.

Consideramos que o Petrus é o mais completo sistema de transcrição fonética para a variedade linguística paulista, pois nele podemos encontrar a separação silábica, a classe gramatical das palavras e a transcrição fonética. Além disso, está disponível para o usuário. No entanto, o conversor ainda comete erros básicos de silabação, como na palavra <ca.rrro>, onde as letras dobradas ficam na mesma sílaba e também a transcrição fonética focaliza apenas dois fenômenos fonéticos: a alternância vocálica, como foi o caso da palavra <peso>, e da epêntese, como é o caso da palavra <pneu>, transcrita por ['pineʊ], acrescida de [ɪ] epentético. Deixa de lado, portanto, outros fenômenos muito relevantes e que estão presentes na transcrição do Potigrafone. Como veremos no capítulo 4, o transdutor que desenvolvemos para a variedade linguística potiguar faz a transcrição fonética dessa última palavra, <pneu>, considerando outras variantes possíveis, como [pinew] e [penew].

2.2.5 Nhenhém 1.0 - Ferramenta de suporte à pesquisa fonológica

Vasilévski, Scliar-Cabral e Araújo (2013) descrevem a estrutura do *software* apelidado de Nhenhém. Esse programa é usado para o tratamento de textos escritos e são codificados em símbolos fonológicos do AFI, formando um *corpus* para análise estatística. Esse sistema é baseado em regras gramaticais de conversão do grafema para o fonema.

Os desenvolvedores afirmam que o aplicativo lê massas de dados relativamente grandes e fornece relatórios de conversão fonológica e relatório estatístico dos fonemas dos textos convertidos. Esse sistema foi produzido para uma aplicação na Linguística de *Corpus*, na qual os dados estatísticos são muito relevantes para uma pesquisa quantitativa dos fonemas.

Segundo Vasilevski (2008, p. 128), o sistema executa três etapas principais: a conversão automática do texto para o alfabeto fonético, a edição de texto transcrito para ajustes e correções e a pesquisa do texto traduzido em dado numérico referente aos fenômenos fonológicos. Após a execução de cada uma dessas etapas, pode ser impresso um relatório da situação em um arquivo.

Outra informação importante sobre o Nhenhém 1.0, apontada por Vasilevski (2008), diz respeito à entrada e à saída de dados. A entrada do sistema pode receber palavras digitadas ou transferidas de outro arquivo, permitindo ao usuário trabalhar com conversão de texto ou de palavras individuais, como ocorre no *Petrus*. A conversão se dá de forma simultânea à escrita. A saída permite extrair dados estatísticos de um ou mais textos para a transcrição fonológica.

Essa autora afirma também que foram feitos testes preliminares, a partir de um *corpus* montado nos pressupostos da teoria pertinente. Depois disso, a execução do aplicativo mostrou que o sistema converte a ortografia da língua portuguesa em uma descodificação de símbolos fonológicos, com aproximadamente 98% de acerto. Além disso, o programa fornece a tonicidade das palavras da língua portuguesa com acerto de quase 99%.

Enfatizamos que esse conversor faz a transcrição fonológica das palavras escritas e permite verificarmos dados estatísticos sobre o texto transcrito. O enfoque do Nhenhém 1.0 não é a transcrição fonética para uma variedade linguística e a autora destaca que é uma ferramenta de suporte à pesquisa fonológica.

2.3 Considerações sobre os sistemas gráficos e fônicos

Nos tópicos acima, apresentamos alguns conversores para o PE e para o PB e tivemos a oportunidade de destacar os seus desempenhos na transcrição de grafema para fonema. Todavia, não faremos uma comparação da *performance* entre eles, pois cada sistema foi desenvolvido com finalidades e com critérios diferentes. Por exemplo, o *Petrus* serve como uma ferramenta de auxílio aos lexicógrafos na transcrição de palavras para os dicionários eletrônicos; o Nhenhém 1.0 é ferramenta de auxílio ao ensino da fonologia do português e à pesquisa na Linguística de *Corpus*; o Grafone auxilia na transcrição fonológica para o português. No entanto, comentaremos sobre três pontos que consideramos relevantes para a compreensão do estado da arte na transcrição gráfica e fônica: a abordagem utilizada nos sistemas, o alfabeto fonético e a relação da entrada com a saída nos conversores.

No que se refere à abordagem dos conversores, os sistemas baseados em regras são muito úteis para gerar aplicação padrão nas transcrições de síntese ou para construir novos *corpora* para sistemas de aprendizado de máquina, como o sistema desenvolvido por Braga, Coelho e Resende (2006) para o PE e, também,

como o de Barbosa *et alia* (2003), o Petrus de Marquiefável, Bokan e Zavaglia (2014), o Nhenhém 1.0 de Vasilévski (2008), para o PB.

Porém, como bem apontam Barros e Weiss (2006), esses sistemas apresentam alguns inconvenientes, uma vez que a aplicação de uma regra pode ter consequências desastrosas em outras. As regras são construídas em forma de cascatas, ou seja, após a execução da primeira regra, o sistema lança como *output* para a segunda e assim sucessivamente. Essa característica serve tanto para os sistemas de regras quanto para o sistema de estados finitos, baseado em regra do padrão gerativo. Outro problema é o fato de serem mais caros do que os sistemas probabilísticos, pois é necessário um linguista perito para configurar todas as regras e exceções com o fim de produzir os bons resultados.

Os sistemas que têm uma abordagem de estados finitos - como o de Caseiro *et alia* (2002) para o PE e o de Shibata (2008) para o PB, são mais flexíveis na integração entre as múltiplas fontes de informação linguística - como a integração de um gráfico e fônico com um etiquetador morfossintático - além de integração de conhecimento em métodos baseados em dados dirigidos. No entanto, deve-se tomar cuidado com o uso das regras de reescritas, pois a má utilização pode tornar o programa ineficaz.

Quanto aos alfabetos fonéticos, foram utilizados para transcrição fonológica e fonética automáticas o SAMPA e o AFI. Este primeiro alfabeto é mais conhecido pelos linguistas e o outro é mais conveniente na aplicação computacional, pois não é necessária a instalação da fonte fonética para o reconhecimento dos caracteres. Braga, Coelho e Resende (2006), Barbosa *et alia* (2003) e o Grafone optaram pelos símbolos fonéticos do primeiro, enquanto os outros usam os símbolos do segundo alfabeto fonético.

Quanto à relação da entrada com a saída nos conversores, devemos considerar que a entrada (*input*) e a saída (*output*) dos dados pelos conversores apresentam quase a mesma estrutura. Com exceção do Petrus, os sistemas apresentam apenas uma transcrição fonológica ou fonética possível para cada palavra, ou seja, existe apenas um *output* para um conjunto de grafemas do *input*. Porém, nem sempre essa relação satisfaz o uso da pronúncia numa variedade linguística da língua portuguesa. Além disso, o Petrus apresenta duas transcrições fonéticas apenas para o fenômeno da alternância vocálica, como vimos no quadro 1 (p. 28).

Uma palavra escrita pode ter duas ou mais formas de pronunciá-la. Como exemplo, citamos a palavra <cebola> que pode apresentar duas pronúncias, [si'bole] e [se'bole], numa mesma variedade linguística. Como dizem Black, Lenzo e Pagel (1998, p.77), "o número de letras em uma palavra e o número de sons em uma pronúncia, em geral, não são de uma para uma. Como temos pesquisado, as letras podem ser mapeadas por zero, um, dois, três ou vários sons."⁶ Se uma letra pode ser mapeada por zero, um ou mais sons, então os sistemas que executam a transcrição de grafema para sons deveriam ter a possibilidade de ter uma ou mais saídas.

Isso constitui o fato de que os sistemas que possuem uma ou mais saídas para a mesma palavra podem ser considerados conversores do tipo *LTS*, como afirmam Black, Lenzo e Pagel (1998). Esses sistemas se enquadram numa transcrição grafema-fone, como é feita pelo *Petrus*. Desse modo, apresentamos aqui um sistema que está adequado à proposta *LTS* para a pronúncia da variedade linguística potiguar. Não desenvolvemos um transdutor que faz a transcrição de uma palavra escrita para o sistema fonológico - como é feita por meio do Grafone e do Nhenhém1.0 - nem para uma variedade padrão - como se pode ver no tradutor baseado em autômatos adaptativos de Shibata (2008) e o *Petrus*.

O Potigrafone - sistema de transcrição fonética para a variedade linguística potiguar, apresentado por este trabalho, faz a transcrição fonética das palavras escritas da língua portuguesa, utilizando os símbolos do alfabeto fonético SAMPA, para representar os sons da variedade linguística potiguar, considerando, para cada símbolo de entrada, a possibilidade de ser mapeada por zero, um ou mais símbolos que apresentam um mesmo som.

Com relação a isso, salientamos que Braga, Coelho e Resende (2006) e Caseiro *et alia* (2002) utilizaram o termo fone⁷ para referir-se ao fonema da língua portuguesa, de modo que os percentuais verificados acima correspondem à relação do grafema com o fonema. Não se trata de um sistema de conversão de grafema para a pronúncia da língua falada em Portugal. O tradutor de grafema para fonema de Shibata (2008) e o sistema *Petrus* de Marquiasfável, Bokan e Zavaglia (2014) exibem a ideia mais próxima da que apresentamos nesse trabalho. No entanto,

⁶ "The number of letters in a word and the number of phones in its pronunciation in general are not a one to one match. For the languages we have investigated, letters can map to zero, one, two or very exceptionally three phones."

⁷ Por enquanto não será apresentada a diferença que há entre os termos fone e fonema. Para essa distinção veja o tópico 4.1.

esses sistemas fazem a transcrição da escrita para a fala padrão de São Paulo e não apresentam a transcrição dos fenômenos fonéticos.

Neste capítulo, destacamos os principais alfabetos fonéticos utilizados pelos conversores. Por um lado, os sistemas desenvolvidos para a integração com os reconhecedores e com os sintetizadores de voz utilizaram o alfabeto fonético SAMPA; por outro lado, os sistemas implementados para utilização na transcrição fonética ou fonológica por linguistas utilizaram o alfabeto fonético IFA. Além disso, pontuamos os tipos de abordagem dos sistemas para compreendermos qual a abordagem mais utilizada e a mais eficaz na transcrição fonética e enfatizamos o desempenho de alguns conversores.

No capítulo 4, apresentaremos um transdutor de estados finitos que faz a transcrição fonética observando outros fenômenos fonéticos que não estão presentes nesses sistemas, como a monotongação, ditongação, redução vocálica, etc. que serão descritos no subtópico 3.2.2. As sequências de grafemas para fones foram modeladas através de um algoritmo que executa o alinhamento entre grafemas e fones para os dígrafos consonantais e vocálicos, para as informações adjacentes ao contexto silábico da língua portuguesa e para a acentuação tônica das palavras.

3 FUNDAMENTOS LINGUÍSTICOS E COMPUTACIONAIS

No capítulo anterior, apresentamos alguns sistemas computacionais que fazem a conversão de grafema para fonema ou de grafema para fone, utilizando, como formas de saída⁸, os símbolos do Alfabeto Fonético Internacional (AFI) ou do SAMPA. A escolha de um ou de outro alfabeto não implica no aumento ou na diminuição do desempenho do sistema, nem na qualidade da transcrição. No entanto, destacamos que, se o sistema estiver disponível como um programa executável no computador, o primeiro alfabeto exigirá que o usuário faça a instalação da fonte fonética adequada ao programa; caso esteja disponível na *web*, como é o caso do Grafone e do Petrus, não é necessária a instalação. É bem verdade que o AFI é mais conhecido pelos linguistas do que o SAMPA, todavia, este foi baseado naquele e, além disso, apresenta um conjunto de símbolos reconhecível pelos computadores, como veremos neste capítulo.

Feitas essas considerações, dividimos este capítulo em quatro partes fundamentais que norteiam o conhecimento linguístico presente no desenvolvimento do sistema e na aplicação computacional.

A primeira parte descreve as formas de entrada, destacando apenas o conjunto de símbolos gráficos do português, os sinais diacríticos e os dígrafos e detalha as formas de saída, mostrando o alfabeto fonético utilizado como forma de saída e a representação dos símbolos utilizados na variedade linguística potiguar.

A segunda parte apresenta os conceitos linguísticos da teoria de base que fundamentam o processo de transcrição fonética.

O terceiro mostra a estrutura silábica como unidade fundamental e organizacional e a marcação da tonicidade das palavras.

A última parte relaciona os conceitos computacionais necessários para a implementação do sistema e distingue as representações subjacentes e as superficiais que estão envolvidas na construção de um conversor.

3.1 Formas gráficas e formas fonéticas

Neste tópico, descrevemos as formas de entrada e as formas de saída do

⁸ Os sistemas Petrus e tradutor grafema para fonema utilizam os símbolos fonéticos do AFI para a variedade linguística paulista.

transdutor, ou seja, as formas gráficas das palavras e as formas fonéticas. Destacamos que a expressão 'formas gráficas' se refere à escrita ortográfica das palavras e não à escrita grafemática. Segundo Gomes (2007), esse tipo de escrita é uma tentativa de representar, o mais próximo possível, a fala; por exemplo, a forma ortográfica para a palavra menino é menino e a forma grafemática é mininu.

As formas de saída refletem as características sonoras das palavras para as variantes linguísticas potiguaras. As tabelas de 5 a 10 listam os sons consonantais e vocálicos e os encontros vocálicos típicos da variedade linguística potiguar, por meio dos símbolos do alfabeto fonético SAMPA. As formas fonéticas representam a fala potiguar para aquela palavra.

3.1.1 Formas gráficas

É fato inconteste que a forma escrita da língua portuguesa é diferente da forma como os falantes a pronunciam. Para Bechara (2005, p. 53), três fatores contribuem para que o sistema ortográfico (ou gráfico) do português não possua uma ortografia ideal, ou seja, um sistema em que cada letra corresponde a um som da língua: (1) a adoção do alfabeto latino que nem sempre é capaz de representar os fonemas da língua; (2) as mudanças que ocorreram com os fonemas das línguas neolatinas depois de terem adotado esse alfabeto e (3) a indecisão permanente das convenções ortográficas entre a opção do sistema fônico ou etimológico.

Considerando a adoção do alfabeto latino, a forma escrita dessa língua possui um conjunto de 26 letras⁹: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z. Desse conjunto de caracteres, *cinco* representam os fonemas vocálicos, *dezessete* representam os fonemas consonantais e *um* sem correspondência fonológica - símbolo mudo, o *h* – usado por herança etimológica e tem uso restrito, e *três* com aplicação nas palavras estrangeiras.

Esse alfabeto não contempla a representação para todos os 33 fonemas da língua. Por isso, segundo Bechara (2005), o alfabeto absorve diversas variações e extensões, fundindo letras - como ocorre nos dígrafos - e modificando-as com o uso dos *diacríticos*, atribuindo funções especiais a duplas de letras ou absorvendo letras completamente novas ao alfabeto.

⁹ Para conhecimento sobre o novo acordo ortográfico e a inserção das letras K, W e Y, acesse o site: <http://www.portaldalinguaportuguesa.org/acordo.php?action=acordo&version=1911>. Destacamos que essas letras não fazem parte do alfabeto do Potigrafone, como justificaremos na metodologia.

Esse autor acrescenta que os **sinais diacríticos** são certos sinais gráficos que se juntam às vogais, geralmente, para lhes dar um valor fonético especial e permitir a pronúncia correta das palavras. Um motivo plausível para o uso dos sinais diacríticos é a eliminação de ambiguidades na escrita. Quando um mesmo grafema representa vários fonemas, existem casos em que a palavra pode ficar ambígua. Nesses casos, o diacrítico pode ser útil para eliminar a ambiguidade. Outro motivo também é o fato de que determinadas palavras na oralidade guardam semelhanças entre si, que podem ser evidenciadas na escrita; é o caso das parônimas. Uma forma de diferenciá-las é adotar o mesmo sinal para essas palavras.

Os diacríticos apresentam as seguintes funções básicas:

- a) o *acento agudo* (´) serve para indicar o som aberto das vogais;
- b) o *til* (ã) marca a nasalização das vogais ou do ditongo, exemplo: irmã, irmão, não, etc.;
- c) o *acento circunflexo* (â) serve para indicar o fechamento do som das vogais;
- d) o *acento grave* (à) serve para indicar o fenômeno da crase;
- e) a *cedilha* (ç) que se coloca sob o c, unicamente antes das vogais "a", "o" e "u".

Nos casos em que não se empregam diacríticos, usam-se os cinco grafemas romanos básicos para representação das vogais (<a>, <e>, <i>, <o>, <u>) e fica, a cargo do leitor, a identificação da vogal representada, o que é facilmente conseguido pela observação do contexto. Além do mais, os acentos agudos e circunflexos são empregados em nossa ortografia com outra função: a de indicar a posição da sílaba tônica na palavra.

O **dígrafo** é o emprego de duas letras para representar um fonema. Esse fenômeno gráfico ocorre devido à inexistência de letra para representar um som na língua portuguesa. Para a formação desse grupo de letras, usa-se uma *letra diacrítica*, aquela "que se junta a outra para lhe dar valor fonético especial e constituir um dígrafo" (BECHARA, 2005, p. 73), contribuindo para uma escrita menos ambígua.

Na forma gráfica da língua portuguesa, ocorrem dois tipos de dígrafos: os dígrafos consonantais e os vocálicos. O primeiro tipo ocorre através das letras <h>, <r>, <s>, <c>, <ç>, <u> que são acrescentadas às consoantes, formando os pares de letras: <ch>, <lh> e <nh>, as geminadas, <rr> e <ss>, <xc>, <sc>, <sç> e o caso

especial de <gu> e <qu>. O segundo tipo se apresenta por meio das letras <m> e <n> que se juntam às vogais para representar as vogais nasais:<am>, <an>, , <en>, <im>, <in>, <om>, <on>, <um> e <um>. Estes últimos pares de letras podem ser acrescentados de um diacrítico quando estão em sílaba tônica, como na palavra <lâmpada>.

Essas letras diacríticas merecem destaque, pois elas apresentam funções bem definidas no transdutor; são alinhadas com as formas fonéticas especiais. As letras diacríticas das consoantes têm o valor fonético zero [0] e as letras diacríticas vocálicas possuem o valor fonético de nasalização da vogal. Estas últimas são representadas pelo símbolo til [~], do alfabeto fonético SAMPA. Esse símbolo é usado também para indicar o fenômeno da nasalidade dos segmentos vocálicos, seguido de uma consoante nasal, como se pode ver na alínea a, do tópico 4.2.1.

Para melhor compreensão do uso dessas letras pelo transdutor, observe os exemplos no quadro 2, onde se encontram as representações para consoantes e dígrafos e, no quadro 3, para as representações das vogais orais e nasais.

3.1.2 Formas fonéticas

No tópico anterior, mencionamos que a língua portuguesa possui 33 fonemas e que os dígrafos são pares de letras que representam um fonema. O termo fonema foi empregado tendo em vista o conhecimento implícito do funcionamento dos sons da língua, ou seja, "como o sistema de sons está representado em nossa mente." (SEARA, NUNES E LAZZAROTTO-VOLCÃO, 2015, p. 98). Como bem aponta Matzanauer (2015), para não haver cruzamento de noções da Fonética e da Fonologia, utilizamos os termos *fonemes* ou *sons* para nos referirmos aos segmentos emitidos por um falante do PB, quando ele produz a fala. As definições dos termos fone, fonema e alofone são apresentadas no tópico 3.2.1.

A distinção desses conceitos é relevante para a compreensão dos argumentos apresentados neste trabalho e para sua plena adequação no sistema computacional, porque a determinação do que é uma forma fonética e do que é uma forma fonológica têm sentidos diferentes para a Linguística e, conseqüentemente, para a aplicação computacional. Como visto no capítulo anterior, a escolha do *output* direciona o conversor para um tipo de transcrição fonética ou fonológica. Nesse caso, podemos mencionar como exemplos o conversor fonológico Grafone e o

fonético Petrus.

Diferenciamos, portanto, a forma escrita da forma fonológica e da fonética por meio dos recursos gráficos. As formas escritas aparecem entre parênteses angulares < >, as formas fonológicas são representadas entre barras / / e as formas fonéticas são representadas entre chaves []. Usamos, por exemplo, a palavra <carro>, para ilustrar a diferença entre a transcrição fonética e a fonológica. O Grafone faz a transcrição fonológica /k'aru/ e o Petrus faz a transcrição fonética ['kaxɔ]. O dígrafo, no primeiro, está representado por um arquifonema e, no segundo, por meio de um retroflexo que caracteriza a variedade linguística paulista.

Esclarecemos, todavia, que o propósito deste trabalho é a apresentação de um transdutor que faça a transcrição da pronúncia para variedade linguística potiguar; a conversão que busca fazer é de grafema para fone (som) e não de grafema para fonema.

As formas fonéticas do transdutor são expressas por símbolos do alfabeto fonético SAMPA para a variedade linguística potiguar. Os subtópicos a seguir apresentam a importância do uso desse alfabeto e o conjunto de fones representados para essa variedade.

3.1.2.1 Alfabeto fonético

O *Speech Assessment Methods Phonetic Alphabet* - SAMPA¹⁰ - apresenta algumas características especiais que justificam seu uso nos sistemas G₂P e LTS:

- a) é um sistema de escrita fonética legível por computadores, pois usa um conjunto de caracteres do *Código Padrão Americano para o Intercâmbio de Informação* (ASCII) de 7 bits, que é a codificação padrão dos teclados dos computadores;
- b) foi desenvolvido a partir do mapeamento dos símbolos do AFI e adota o máximo de caracteres desse alfabeto;
- c) pode ser uma alternativa para a transcrição quando não se tem em mãos "as fontes que são formadas, entre outras, por muitas letras gregas e símbolos próprios." (SEARA, NUNES e VOLCÃO, 2015, p.24);
- d) não necessita da instalação de uma fonte no computador para o reconhecimento dos símbolos, como é o caso do AFI.

¹⁰ Para informações sobre esse alfabeto consulte o site: <<https://pt.wikipedia.org/wiki/SAMPA>> e para ver a diferença entre os símbolos do alfabeto fonético internacional (IPA) e o SAMPA, acesso em: <http://www.lfsag.unito.it/ipaindex_en.html> ou consulte os anexos A e B deste trabalho.

O conjunto de símbolos foi desenvolvido em consulta com falantes nativos de cada língua para a qual foram aplicadas, mas a padronização do sistema é de nível internacional. Além disso, ele foi originalmente desenvolvido por um grupo internacional de foneticistas que o aplicou primeiramente nas Comunidades Europeias de línguas dinamarquesa, holandesa, Inglesa, francesa, onde *Unicode* (ISO 10646) não estava disponível. Esse alfabeto constitui a base de colaboração internacional mais robusta para uma codificação de leitura óptica padrão de notação fonética. Todos os símbolos do AFI, que coincidem com as letras minúsculas do alfabeto latino, permanecem os mesmos. Quando isso não é possível, utilizam-se outros sinais disponíveis, como ocorre com a vogal baixa em posição final que, no AFI, é representada pelo símbolo [e] e no SAMPA pelo símbolo [6].

Todavia, destacamos que "os símbolos do SAMPA são adaptados diferentemente para cada uma das línguas que transcreve e, por isso, um dado conjunto desses símbolos só é válido para o idioma ao qual ele foi adaptado". (SEARA, NUNES e VOLCÃO, 2015, p.24). No anexo C, apresentamos os símbolos fonéticos usados para o sistema de escrita fonética do português e do qual fundamentamos nossa adaptação para a variedade linguística potiguar.

3.1.2.2 Símbolos fonéticos para a variedade linguística potiguar

Os segmentos consonantais e vocálicos que fazem parte da transcrição fonética para a variedade linguística potiguar foram constituídos a partir dos símbolos fonéticos do SAMPA para o português e estão disponíveis no *anexo C* deste trabalho, como bem salientamos no tópico anterior.

A distribuição dos segmentos partiu da concepção de que a produção da fala se caracteriza por um conjunto complexo de fatores que contribuem para a distinção dos sons. O primeiro fator é a necessidade de vibração das moléculas de ar e estas serem expelidas pela compressão do volume dos pulmões, tornando o fluxo de ar audível pela vibração das pregas vocais, ou seja, a produção das vogais e das consoantes. O segundo fator se caracteriza pelo fluxo de ar que sofre modificações em função das alterações promovidas dentro do aparelho fonador, a partir dos movimentos e posicionamentos dos órgãos articuladores que estão acima da laringe. Assim,

os órgãos articuladores, como língua, lábios, ao se moverem (articularem) para produzir os variados sons do português, transformam o trato vocal em um tubo acústico ou uma caixa de ressonância que irá amplificar e modificar os sons produzidos. SEARA, NUNES e VOLCÃO (2015, p. 38).

Como afirmam as autoras, a divisão entre vogais e consoantes em nível de articulação deve ser entendida a partir da liberação do fluxo de ar dos pulmões. Para a emissão das primeiras, não há qualquer impedimento à passagem de ar, porém, para as segundas, ocorre obstrução no trato vocal, seja parcial ou total. Desse modo, neste tópico, mostraremos a classificação da produção do sons consonantais e vocálicos, incluindo ditongos e tritongos, da variedade linguística potiguar e seus respectivos símbolos do SAMPA.

3.1.2.2.1 Consoantes

Para Seara, Nunes e Volcão (2015), os segmentos consonantais apresentam algum tipo de resistência, total ou parcial, ao fluxo de ar no trato oral. O *modo de articulação* se caracteriza pela forma como o ar passa pelas cavidades supraglotais; o *lugar* ou *ponto de articulação* define a posição dos articuladores passivos e ativos, quando se produz uma consoante e, por fim, o vozeamento ou sonoridade¹¹ distingue os sons surdos (não vibram as pregas vocais) e sonoros (vibram as cordas vocais). O quadro 2, abaixo, apresenta a relação dos grafemas consonantais e dos dígrafos com os sons consonantais da variedade potiguar, os quais fazem parte da transcrição fonética do Potigrafone.

O **modo de articulação** está relacionado ao tipo de obstrução no trato vocal. Nesse caso, a produção de uma consoante ocorre por meio de alguns obstáculos à passagem da corrente de ar como uma oclusão, uma constricção, uma oclusão parcial que desvia a direção da corrente de ar, uma tremulação da língua que permite a vibração da corrente. De acordo com Câmara Jr. (2008), o primeiro obstáculo põe a consoante na classe das *oclusivas*; o segundo, na das *constritivas* ou *fricativas*; o terceiro, na classe das *nasais*, com oclusão ou constricção na boca, mas com ressonância pelas fossas nasais; o quarto, na das *laterais*, com oclusão num ponto central da língua e desvio laterais da corrente de ar e o quinto, na classe das *vibrantes*.

¹¹ Silva (2014) prefere utilizar os termos vozeada e desvozeadas, Seara, Nunes e Volcão (2015) empregam os termos surdas ou não vozeada e sonoras ou vozeadas. Utiliza-se aqui esse termos como sinônimos.

Apresentamos, a seguir, o modo de articulação das consoantes da fala potiguar, as quais fazem parte do português brasileiro, mas não são sons exclusivos do dialeto potiguar, pois tais consoantes estão presentes também em outras variantes do PB. Acreditamos, como diz Matzenauer (2015), que esses segmentos representam as formas fonéticas das variantes da fala potiguar e não são fonemas do PB.

As *consoantes oclusivas* são produzidas quando o fluxo de ar encontra uma interrupção total, seja pelo fechamento dos lábios, seja pela pressão da língua sob a arcada dentária ou sob o palato duro. Na variedade linguística potiguar, essas consoantes são [p, b, t, d, k, g]. Em <capa> [ˈkapɔ] e <cabo> [ˈkabu], os sons [p] e [b] são emitidos com obstrução total dos lábios; em <pata> [ˈbetɔ] e <seda> [ˈsedɔ], os sons [t] e [d] são produzidos com uma obstrução total na região que vai dos dentes até os alvéolos e, nas palavras <cada> [ˈkadɔ] e <gata> [ˈgatɔ], os sons são produzidos a partir da obstrução total do véu palatino. A maioria desses sons tem representação direta entre a pronúncia e a letra para a variedade.

As *consoantes fricativas* ocorrem quando há um estreitamento da passagem do ar, que resulta em um ruído semelhante ao de uma fricção. Elas são representadas pelos símbolos fonéticos [f, v, s, z, ʃ, ʒ, h]. Alguns desses sons apresentam muitas divergências com a grafia. O fonema [s] é representado na escrita por diversas letras, tornando a relação indireta, de modo que apenas o contexto silábico ou lexical¹² define a representação sonora da letra. Ressaltamos que o som da fricativa glotal surda [h] é representada na escrita pela letra <r> no início da palavra ou em posição pós-vocálica e, ainda, pelo dígrafo <rr>, como nas palavras "rato", "mar" e "carro", respectivamente (SILVA, 2014).

Como bem destaca Silva (2014), em todos os dialetos¹³ do português falado no Brasil, há distinção fonêmica entre a vibrante simples /r/ e a vibrante múltipla /r/ em posição intervocálica. Os segmentos fonéticos das vibrantes múltiplas podem ocorrer em posição pós-vocálica como uma consoante fricativa [h], [ç], [x] [ɦ] ou retroflexa [ɻ]. Por isso, utiliza-se o arquifonema /R/ para denotar a neutralização existente no contraste fonético desses segmentos. Assim, nessa posição, o /R/ pode

¹² O termo lexical está sendo usado como sinônimo de palavra, vocábulo.

¹³ Nesse trabalho, os termos *dialeto* e *falar* são sinônimos de *variedade linguística*, ou seja, referem-se ao falar característico de determinado grupo social e/ou regional. (Ver Coelho *et al*, 2015, p. 15).

ser pronunciado por um desses segmentos, dependendo do dialeto, como na palavra “mar” que pode ser emitido [ˈmar], [ˈmah], [ˈmaɦ], [ˈmaɣ], [ˈmax] e [ˈmaɾ].

Na variedade linguística potiguar, ocorre o contraste fonêmico em posição intervocálica entre o “r fraco” [r] e o “R forte” [r̄]. O primeiro é pronunciado como uma vibrante simples entre vogais e seguindo consoante na mesma sílaba e o segundo é pronunciado como fricativa glotal entre vogais, em início de palavra, seguindo consoante em outra sílaba e em coda medial e final.

As *consoantes nasais* são aqueles sons que, na sua realização, parte da corrente de ar que sai pelo trato vocal e parte pelas fossas nasais devido ao abaixamento simultâneo do véu do palato. Quando iniciam as sílabas, as letras <m> e <n> e o dígrafo <nh> são representadas pelos símbolos fonéticos [m], [n] e [ɲ], respectivamente, como nas palavras <mar> [ˈma0], <nuvem> [ˈnuve~j] e <banho> [ˈba~J0u]. Quando as duas primeiras letras estiverem em posições pós-vocálicas, são representadas pelos símbolos de nasalização [~], pois nessa posição elas constituem a segunda letra do dígrafo vocálico, como foi mencionado no final do *tópico 3.1*.

As *consoantes laterais* são emitidas quando a língua, ao tocar os alvéolos, obstrui a passagem do ar nas vias superiores, mas permite que o ar passe através das paredes laterais da boca. Quando iniciam uma sílaba ou quando é o segundo elemento do ataque complexo, a letra <l> representa a lateral alveolar [l] e o dígrafo <lh> representa a lateral palatal [L], respectivamente. Segundo Silva (2014), quando a primeira estiver em posição pós-vocálica ocorre o fenômeno da vocalização e é representada pelo glide [w].

A *vibrante simples* [r] se caracteriza pelo movimento vibratório e rápido da língua, provocando, assim, breves interrupções na corrente de ar. Esse som é "uniforme em posição intervocálica", como na palavra <caro>, e "seguindo consoante em todos os dialetos do português brasileiro" (SILVA, 2015, p.39), como na palavra <prarato>.

O quadro 2 apresenta a relação dos símbolos fonéticos do SAMPA com as letras e os dígrafos que representam as consoantes para o português brasileiro.

Quadro 2 – Relação dos símbolos fonéticos com as letras e os dígrafos consonantais

Modo	Classificação		Fone Grafema Dígrafo		Exp.	Transcrição fonética*
	Ponto	Vozeamento				
Oclusivas	Bilabial	Desvozeado	[p]	<p>	soprar	[so"pra0]
		Vozeada	[b]		Bem	[be~j]
	Alveolar	Desvozeada	[t]	<t>	con t ar	[ko~"ta0]
		Vozeada	[d]	<d>	na d ar	[na"dah]
	Velar	Desvozeada	[k]	<c>	Co m o	["ko~mu]
Vozeada		[g]		<gu>	gu er ra	["g0Eh0]
Fricativas	Labio dental	Desvozeada	[f]	<f>	Fogo	["fogu]
		Vozeada	[v]	<v>	árvore	["ahvori]
	Alveolar	Desvozeada	[s]	<c>	prec i sar	[presi"za0]
				<ç>	cabeça	[ka"bes6]
				<s>	ser	[se0]
				<x>	máximo	["masimu]
				<z>	feliz	[fe"lis]
				<ss>	pássaro	["pas0aru]
				<sc>	nascer	[na"s0e0]
				<sç>	de s ça	["des06]
				<xc>	exceto	[e"s0Etu]
				<xs>	exsudar	[e"s0da0]
		Vozeada	[z]	<s>	pe s ado,	[pe"zadu]
			<z>	dizer,	[di"ze0]	
			<x>	exílio	[e"zilju]	
	Desvozeada	[S]	<x>	Xá	["Sa]	
	Glotal	Desvozeada		<ch>	Chá	["S0a]
			[h]	<r>	rato,	["hatu]
				<rr>	porta	["poht6]
					Carro	["kah0u]
Nasais	Bilabial	Vozeada	[m]	<m>	mar	[mah]
	Alveolar	Vozeada	[n]	<n>	nuvem	["nuve~j]
	Palatal	Vozeada	[J]	<nh>	banho	["ba~J0u]
Laterais	Alveolar	Vozeada	[l]	<l>	Eles	["eles],
	Palatal	Vozeada	[L]	<lh>	mulher	[mu"LEh]
Vibrantes	Alveolar	Vozeada	[r]	<r>	Caro	["karu]
					Prato	["pratu]

Fonte: Adaptado de Silva (2014), Masip (2014) e Seara, Nunes e Volcão (2015).

* A transcrição dessas palavras foram testadas pelo sistema e todas apresentam essas formas fonéticas, com exceção das palavras <ser>, <máximo> e <exsutar>.

O **ponto de articulação** se relaciona aos articuladores envolvidos na constrictão de consoantes e se define a partir do articulador ativo em relação ao passivo. As consoantes são classificadas como:

- *bilabiais* - ocorrem quando o lábio inferior toca no lábio superior, a exemplo das consoantes [p, t, m] nas palavras, <pa~~t~~o> ["pa~~t~~u], <ba~~t~~o> ["ba~~t~~u] e <ma~~t~~o> ["ma~~t~~u];
- *labiodentais* - são produzidas quando o lábio inferior vai em direção aos dentes incisivos superiores. São exemplos as consoantes [f, v] na

palavra <fava>["fava] em que as letras e os sons testados são labiodentais;

- *alveolares* - ocorrem quando a ponta ou a lâmina da língua toca ou vai na direção dos alvéolos. As consoantes [t, d, s, z, n, r] nas palavras <tudo> ["tudu], <assado> [a"s0adu], <casa> ["kaza], <nora> ["nOra] são exemplos desse tipo.
- *alveopalatais* - ocorrem quando a parte anterior da língua toca ou dirige para a região medial do palato duro, como os sons consonantais [S, Z] nas palavras <chá> [S0a], <jeito> ["Zeitu] e <gente> [Ze~t@];
- *palatais* - ocorrem quando a parte média da língua toca ou se encaminha em direção ao palato duro, como as consoantes [L, J] nas palavras <palha> ["paL0a] e <ganho> ["gaJ0u];
- *velares* - ocorrem quando o dorso da língua toca ou vai na direção do palato mole ou véu palatino, como as consoantes [k, g] nas palavras <casa> ["kaza] e <galo> ["galu];
- *glotal* - ocorre quando os músculos da glote são os articuladores desse segmento, a exemplo da pronúncia do "erre" nas palavras <carro> ["kah0u], <rato> ["hatu] e <marte> ["maht@].

Vale ressaltar as consoantes palatais quanto ao seu ponto de articulação. Silva (2014, p.64, 65) destaca que a consoante lateral palatal correspondente ao dígrafo <lh> apresenta três alternativas de articulação. A primeira diz respeito à articulação como consoante lateral palatal que apresenta obstrução da passagem da corrente de ar na região palatal e é representada pelo símbolo [L]. A segunda diz respeito ao fato de que o dígrafo <lh> representa os casos em que uma consoante lateral alveolar é articulada com a propriedade secundária de palatalização, sendo representado pelo uso de um diacrítico [l']. A terceira articulação ocorre quando os falantes pronunciam as palavras <teia> e <telha> de maneira idêntica. Nesse caso, tem-se uma vogal que ocupa a posição consonantal correspondente ao dígrafo <lh>. No Potigrafone, essa posição é ocupada pelo mesmo símbolo da consoante palatal, como podemos ver na descrição do fenômeno da palatalização na página 56 deste trabalho.

As consoantes são classificadas quanto ao vozeamento e podem ser

surdas ou desvozeada, quando não ocorre a vibração das cordas vocais na produção da consoante. Isso acontece com os sons oclusivos [p, t, k] e fricativos [f, s, S, h]. Os sonoros ou vozeados ocorrem quando há vibração das cordas vocais na produção da consoante, como nos sons oclusivos [b, d, g], fricativos [v, z, Z], laterais [l, L], nasais [m, n, J] e na vibrante [r].

Esse subtópico apresentou a classificação dos segmentos consonantais para a fala potiguar e a relação destes com as letras e os dígrafos da escrita do português, relacionando as formas de entrada com as formas de saída para as consoantes presentes nessa variedade.

3.1.2.2.2 Vogais

Os segmentos vocálicos, segundo Seara, Nunes e Volcão (2015), são sons produzidos com o ar saindo dos pulmões, de modo que o estreitamento gerado pelo movimento dos articuladores não produzem fricção. Para essas autoras, as vogais do PB e, conseqüentemente, da variedade potiguar são descritas segundo os aspectos articulatórios oralidade e nasalidade, movimento de altura da língua, movimento de avanço e recuo da língua e estiramentos dos lábios. Tais aspectos são apresentados no quadro 3.

As vogais podem ser classificadas como **orais** e **nasais**. Na produção das orais, o véu palatino fecha a passagem para a cavidade nasal, fazendo com que o ar saia apenas pelo trato oral. Nas nasais, ao contrário, o véu do palato encontra-se abaixado, permitindo que o ar saia também pelas cavidades nasais.

As vogais são definidas quanto ao *movimento de altura da língua*, podendo ser altas, médias ou baixas. As *altas* ocorrem quando o corpo da língua se eleva ao máximo, estreitando o trato vocal, mas sem produzir fricção, como no caso da produção dos sons [i] e [u]. As *médias* ocorrem quando a língua se encontra numa posição intermediária entre alta e baixa. Quando se encontra na posição mais baixa, produzem-se os sons [E] e [O] e, quando está numa posição mais alta, emitem-se os sons [e] e [o]. As vogais médias ocorrem quando o corpo da língua se encontra na posição mais baixa do trato vocal, como na produção do som [a]. Essa classificação também é chamada de classificação quanto ao eixo vertical.

As vogais podem ser classificadas quanto ao avanço ou recuo da língua como anteriores, posteriores e centrais. As *centrais* são realizadas na posição mais

central do trato vocal, como as vogais orais [a] e [ɒ] que são produzidas com a boca aberta e a língua na posição quase de repouso (relaxada) e exigindo menor esforço. As *vogais anteriores* se dirigem para a parte anterior do trato vocal, como as vogais orais [E], [e] e [i] que são emitidas quando a ponta da língua vai avançando até o palato duro, diminuindo a abertura bucal e aumentando a abertura da faringe. As *posteriores* se movimentam para a parte posterior do trato vocal na direção do palato mole, sem apresentar bloqueio à passagem de ar. As vogais orais [O], [o] e [u] ocorrem quando o dorso da língua se eleva, recuando em direção ao véu palatino e provocando uma diminuição da abertura bucal e um arredondamento progressivo dos lábios.

As vogais também são classificadas quanto ao timbre e podem ser abertas, fechadas e reduzidas. O timbre é o efeito acústico resultante da distância entre o dorso da língua e o véu do palatino, tendo a cavidade bucal como caixa de ressonância. As *abertas* [a], [E] e [O] são produzidas quando a língua se acha baixa e as *fechadas* [e], [o], [i] e [u] ocorrem quando a língua se eleva. As formas *reduzidas* são proferidas anulando a oposição entre aberta e fechada como é o caso das soantes orais [a], [i] e [u] e das soantes nasais do quadro 3, abaixo.

Na tabela a seguir apresentamos a classificação dos sons vocálicos para a fala potiguar, considerando a oralidade e a nasalidade - o abaixamento ou levantamento do véu palatino - o movimento de altura, recuo e avanço da língua e o timbre das vogais.

Quadro 3 – Segmentos vocálicos para a variedade linguística potiguar

		Anterior		Central		Posterior	
		Oral	Nasal	Oral	Nasal	Nasal	Oral
Tônica	Alta	i	i~			u~	u
	Média alta	e	e~			o~	o
	Média baixa	E					O
	Baixa			a	a~		
Pós-tônica	Alta	i				u	
	Média alta						
	Baixa			ɒ			

Fonte: Adaptação do quadro 7 de Seara, Nunes e Volcão (2015, p. 64).

No quadro 4, abaixo, apresentamos a relação entre as representações ortográficas para as vogais orais e nasais. Silva (2015, p. 60) observa que "*ortograficamente* a letra **m** ocorre em final de sílaba e em final de palavra (como em 'campo' e 'fim'). Nesse caso, a letra **m** marca a nasalização da vogal anterior e não a articulação de uma consoante nasal". A escrita marca a nasalização da vogal por uma consoante nasal na mesma sílaba. Por esse motivo, as letras <m> e <n> foram destacadas como letras diacríticas no tópico 3.1, formando dígrafos vocálicos.

A autora acrescenta que "a nasalidade de uma vogal ocorre quando uma vogal tipicamente oral é seguida por uma das consoantes nasais: [m, n, ɲ]." (SILVA (2015, p 93). Esse fenômeno se apresenta entre duas consoantes nasais, como nas palavras "mono" [mo~nu], mínimo" [mi~ni~mu] ou imediatamente antes de uma consoante nasal, como em "cama" [ka~m6], "cana" [ka~n6], ocorrendo assim a assimilação do traço nasal da consoante. Esse fenômeno marca a variação dialetal, pois as variedades nordestinas, como a potiguar, preferem o uso da nasalidade, enquanto que a variedade paulista expressa a falta desse fenômeno.

Quadro 4 - Relação dos símbolos fonéticos com as letras e dígrafos vocálicos.

Fone	Grafema	Exemplo	Transcrição*
[a]	<a> - <á>	ca <u>s</u> a, ca <u>j</u> á	["kaz6] - [ka"Za]
[6]	<a>	Salada <u>a</u>	[sa"lad6]
[a~]	<am> - <an> <âm> - <ân> <ã>	am <u>p</u> aro, am <u>o</u> res âmb <u>i</u> to, ânc <u>o</u> ra irmã <u>a</u>	[a~"paru] - [a~"tis] [a~"bitu], [a~"kor6] [ih"ma~]
[e]	<e> - <ê>	pe <u>s</u> o, pêssego	["pezu] - ["pes0egu]
[e~]	 - <en> <êm> - <ên>	te <u>m</u> po, rend <u>a</u> tê <u>m</u> , ênfase	["te~pu] - ["he~da] ["te~] - ["e~fasi]
[E]	<e> - <é>	celest <u>e</u> , caf <u>e</u>	[se"leSti] - [ka"fE]
[i]	<i> - <í> <e>	pic <u>o</u> , ítal <u>o</u> pass <u>e</u>	["piku], ["italu] ["pas0i]
[i~]	<im> - <in> <ím> - <ín>	in <u>f</u> lamável, simp <u>l</u> es, íntim <u>o</u> , ímp <u>e</u> to	[i~"fla~"mavew], [i~"si~plis] [i~"timu], [i~"pÉtu],
[O]	<o> - <ó>	ov <u>a</u> l, vov <u>o</u>	[o"vaw] - [vo"vO]
[o]	<o> - <ô>	ov <u>o</u> , mofad <u>o</u>	["ovu] - [mo"fadu]
[o~]	<om> - <on> <ôn> <õ>	somb <u>r</u> a, ond <u>e</u> de <u>o</u> ntico bot <u>õ</u> es	["so~br6] - [o~"di] [de"o~tiku] [bo"to~js]
[u]	<u> - <ú> <o>	uv <u>a</u> , sujeir <u>a</u> , últim <u>o</u> porc <u>o</u>	["uv6], [su"Zejr6] ["uwtimu] ["pohku]
[u~]	<um> - <un>	umbig <u>o</u> , junt <u>o</u> ,	[u~"bigu], [Zu~tu]

Fonte: Adaptado de Masip (2014) e Seara, Nunes e Volcão (2015).

* Todas as palavras foram transcritas no sistema Potigrafone.

3.1.2.2.3 Encontros vocálicos: ditongos, tritongos e hiatos

Segundo Masip (2014), as semivogais¹⁴ [j] e [w], tanto oral quanto nasal, não são sons prolongáveis e uniformes como o [i] e o [u], respectivamente, mas fruto de uma emissão momentânea, com tendência a fechar progressivamente sua abertura palatal. Assim, não se tornam núcleo silábico, mas acompanham uma vogal em uma sílaba, formando os ditongos ou os tritongos. Na escrita, elas são representadas por meio das letras “i” e “e” para o símbolo [j], e “u” e “o” para o símbolo [w].

Os ditongos, segundo Seara, Nunes e Volcão (2015, p. 66), são formados por dois segmentos vocálicos em que há possibilidade de duas sequências: uma vogal-semivogal e uma semivogal-vogal. Elas ressaltam que a primeira sequência é sempre inseparável e é chamada de ditongo decrescente, pois termina pela vogal com menor proeminência acentual; a segunda, chamada ditongo crescente, é finalizada pelo segmento de maior proeminência e, por isso, há a possibilidade de constituir sílabas separadas em alguns casos.

Nos quadros 5 e 6, a seguir, listam e relacionam os segmentos que constituem os ditongos crescentes e decrescentes, orais e nasais para o PB e que estão presentes na variedade linguística potiguar. Ressaltamos que a relação desses encontros vocálicos é apresentada com a ausência do fenômeno da monotongação e da ditongação que ocorre com alguns ditongos nessa mesma variedade. Para compreender as regras e o processo fonológico desses fenômenos veja o tópico 4.1.2.

Quadro 5 – Representação dos ditongos decrescentes orais e nasais

Ditongos decrescentes							
Orais				Nasais			
Fone	Grafema	Expl.	Transcrição	Fone	Grafema	Expl.	Transcrição
[aj]	ai	pa <u>i</u>	["paj]	[a~w]	ão	Mão	["ma~w]
[aw]	au	pa <u>u</u>	["paw]	[a~j]	ãe	Mãe	["ma~j]
[Ej]	éi	ré <u>i</u> s	["rEjs]	[e~j]	em	Tem	["te~j]
[ej]	ei	le <u>i</u> te	[lejti]	[o~j]	õe	Põe	["po~j]
[Ew]	éu	cé <u>u</u>	["sEw]	[u~j]	ui	Mui	["mu~j]
[ew]	eu	me <u>u</u>	["meu]				
[iw]	iu	vi <u>u</u>	["viu]				
[Oj]	ói	ró <u>i</u>	["hOj]				

¹⁴ No próximo capítulo, as semivogais serão consideradas glides. Alguns autores preferem usar a expressão "vogal assilábica", como Câmara Jr. (2008); outros utilizam o termo semivogal, como Masip (2014).

[oj]	oi	foi	["foj]
[ow]	ou	sou	["sow]
[uj]	ui	cuida	["kujd6]

Fonte: Adaptado de Silva (2015, p. 94-97) e Seara, Nunes e Volcão (2015, p.67).

* Todas as palavras foram transcritas no sistema Potigrafone.

Quadro 6 – Representação dos ditongos crescentes orais e nasais

Ditongos crescentes							
Orais				Nasais			
Fone	Grafema	Expl.	Transcrição*	Fone	Grafema	Expl.	Transcrição
[ja]	la	polícia	[po"lisja]	[ja~]	lan	Criança	["krja~sa]
	Ea	área	["arja]	[wa~]	Uan	Quando	["kwa~du]
[ju]	lo	sério	["sErju]	[we~]	uen	Cinquenta	["si~kwe~t@]
	Eo	aéreo	[a"Erju]		Uên	Sequência	[sE"kwe~sja]
[ji]	le	série	["sErji]	[wi~]	Uin	Arguindo	[ah"gwi~du]
[wi]	Ue	tenue	["te~nwi]				
[wu]	Uo	árduo	["ahdwu]				
[wa]	Ua	nódoa	["nOdwa]				
	Oa	mágoa	["magwa]				

Fonte: Adaptado de Silva (2015, p. 94-97) e Seara, Nunes e Volcão (2015, p.67).

* Todas as palavras foram transcritas no sistema Potigrafone. As palavras <mágoa> e <nódoa> apresentaram erros na vogal final e as palavras <cinquenta>, <sequência> e <arguindo> foram transcritas como se fossem dígrafos na sequência de <gu> e <qu>.

Os tritongos são formados por três segmentos vocálicos em que somente um deles ocupa o pico silábico. "Alguns estudiosos consideram os tritongos como a fusão de um ditongo crescente e um decrescente; outros consideram que tritongos, precedidos de [k] e [g], seriam na verdade consoantes complexas seguidas de ditongo." (SEARA, NUNES e VOLCÃO, 2015, p.68).

Quadro 7 – Representação dos tritongos orais e nasais

Tritongos							
Orais				Nasais			
Fone	Grafema	Expl.	Transcrição*	Fone	Grafema	Expl.	Transcrição
[waj]	uai	Uruguai	[uru"gwaj]	[wa~w]	Uão	Saguão	[sa"gwa~w]
[wej]	uei	averigui	[averi"gwej]				
[wiw]	uiu	redarguiu	[redar"giw]				
[wow]	uou	enxaguou	[e~Sa"gwow]				

Fonte: Masip (2014, p. 39) e Seara, Nunes e Volcão (2015).

* Todas as palavras foram transcritas no sistema Potigrafone. Apenas as palavras <averigui>, <redarguiu> e <arguindo> foram transcritas como se fossem dígrafos na sequência de <gu> e <qu>.

Quanto aos hiatos, são encontros vocálicos nos quais cada vogal constitui uma sílaba, a exemplo das palavras "saí" [sai] e "baú" [bau]. "Um hiato pode ser intravocabular, quando ocorre dentro de uma palavra, ou intervocabular, quando é consequência do encontro entre uma vogal final de uma palavra e a vogal inicial de outra." (SEARA, NUNES E VOLCÃO, 2015, p. 69).

As autoras também apresentam outras ocorrências de hiatos: com vogais iguais átonas ("caatinga" e "coordenação"); com vogais iguais, em que a primeira é tônica ("voo" e "veem"); com vogais iguais, em que a segunda é tônica ("alcoólico" e "xiita"); com vogais diferentes átonas ("doação" e "estereotipado"); com vogais diferentes, sendo a primeira tônica ("Maria" e "pavio") e com vogais diferentes, em que a segunda é tônica ("freada").

3.2 Conceitos fonológicos

Neste tópico, apresentamos os conceitos linguísticos da teoria de base que fundamenta o processo de transcrição fonética e deixamos para o último tópico deste capítulo os conceitos computacionais necessários para a implementação do sistema. Essa distinção deixa o texto mais didático e nos ajuda à compreensão de que a Fonologia Computacional se apresenta como uma aplicação das técnicas formais e computacionais para a representação e processo de informações fonológicas (BIRD, 2009).

Nesse sentido, dividimos este tópico em dois subtópicos: o primeiro apresenta a distinção dos conceitos-chave de fone, fonema e alofone; o segundo relaciona as regras e os processos fonológicos.

3.2.1 Conceitos-chave

A Fonética é a área da Linguística que se ocupa das propriedades físicas dos sons da fala, tanto do ponto de vista da percepção como do ponto de vista da produção. A Fonologia estuda a organização dos sistemas de sons das línguas através da representação e formalização de modelos e processos sonoros, para que os falantes tenham as unidades mínimas sonoras da sua própria língua.

Essa diferença na perspectiva sobre a produção sonora contribui para a distinção dos conceitos de fone, fonema e alofone. Consideramos que as unidades físicas da fala estão em função dos sons da língua, ou seja, ambas as unidades são fundamentais para a compreensão dos fenômenos fonéticos presentes em uma variedade linguística. Por isso, os conceitos-chave que estão expostos nos subitens abaixo dizem respeito às definições específicas dos termos técnicos que estão presentes na Fonologia e na Fonética.

Assim, para a produção de um som em qualquer língua, faz-se necessário o uso de uma parte específica do corpo humano que se denomina aparelho fonador¹⁵. Os sons da fala, a partir de seus movimentos articulatórios, são vistos como fones, uma vez que as suas características são independentes da função que eles desempenham numa língua determinada. (SEARA, NUNES E LAZZAROTTO-VOLCÃO, 2015).

Por meio do conceito de fone, podemos extrair as escalas físicas que descrevem os aspectos do evento da fala como alta, média e baixa, para as vogais e labial, dental, palato alveolar e velar para as consoantes. Além dessas, também descrevem as que aparecem em pares e representam a presença ou ausência de atributos ou do ponto de produção ou da recepção, por exemplo, nasal/oral, sonoro/surdo, arredondada/não arredondada, posterior/anterior, etc. Por isso, se diz que o som consonantal [b] é articulado com uma corrente de ar pulmonar egressiva, com vibração das cordas vocais e com obstrução de ar seguida de uma explosão; diz-se, também, que os três *as* na palavra <salada> podem ser considerados como realizações um pouco distintas, mas os falantes do português reconhecem esses sons como pertencendo à mesma categoria, ou seja, trata-se da mesma vogal [a]. (MASSINI-CAGLIARI e CAGLIARI, 2011).

Mostramos, no tópico anterior, o conjunto de fones (sons) para o dialeto potiguar, a partir dos movimentos articulatórios, ou seja, do ponto de vista de sua produção pelo aparelho fonador. É bem verdade que não detalhamos os fenômenos fonéticos que estão presentes nessa variedade linguística, mas a maioria dos segmentos consonantais e vocálicos foi constituída com base empírica da fala potiguar, mostrada na transcrição fonéticas dos Atlas Linguístico do Centro-Oeste Potiguar e do ALiPTG, segundo Silva (2012) e Pereira (2008), respectivamente.

O **fonema** é uma unidade sonora mínima que distingue duas palavras entre si, ou seja, o fonema tem um papel **contrastivo** e concreto na pesquisa linguística. "A determinação de fonema se dá a partir da identificação de pares mínimos para um grupo de dois segmentos." (SILVA, 2014, p. 127). E o par mínimo caracteriza os fonemas por contraste em ambiente idêntico. Pensando assim, pode se dizer que o par mínimo /p/ e /b/ é constituído por dois fonemas do PB, uma vez que eles distinguem uma palavra da outra, como <pata> de <bata> e a ausência

¹⁵ Silva (2014) e Seara, Nunes e Lazzarotto-Volcã (2015) apresentam uma descrição detalhada dos mecanismos de produção da fala e da articulação dos sons para o PB, os quais não foram dirimidas neste trabalho.

deles nessa sequência de fonema origina uma nova palavra, como <ata> por exemplo. Os fonemas são unidades sonoras capazes de marcar a diferença de significado, ou seja, exercem as funções relacionadas às diferenças de significado, de modo que, quando substituídos ou eliminados, mudam o sentido das palavras.

Os registros das cartas fonéticas, nesses atlas, consideraram as variantes de um fonema, de vários fonemas correspondentes a um único fonema mais antigo ou de determinadas séries de fonemas. Foram comprovadas nos pontos investigados pelas pesquisas e mostradas nas figuras 13 e 15 deste trabalho.

No entanto, quando dois sons não implicam em mudanças de significado da palavra para o determinado contexto do fonema, diz-se que ocorre um **alofone** ou **variante fonológica**¹⁶. Cada fonema pode ser emitido pelo falante de forma diferente, resultando em sons diferentes para a mesma unidade distintiva. Os alofones são sons cujos traços comuns permitem associá-los a um fonema da língua. Eles são identificados por meio da *distribuição complementar* e da *variação livre*.

Para Silva (2014), o primeiro ocorre quando dois segmentos realizam-se em ambientes exclusivos, ou seja, onde uma das variantes ocorre a outra variante não acontecerá. O segundo caso se apresenta quando dois segmentos ocorrem no mesmo ambiente sem prejuízo de significado.

Um exemplo de distribuição complementar ocorre com o fonema /t/ que pode ser produzido como um som africado [tʃ] ou como um som oclusivo velar [t], dependendo do contexto. Quando esse fonema vem precedido das vogais [i, i~], como na palavra <tia> é pronunciado como [ˈtʃia] ou [ˈti6]; não se realiza como africada quando vem precedida das outras vogais. Esses alofones são foneticamente distintivos, mas apenas variação do som dependendo do contexto e da comunidade de falantes, ou seja, os sons [t] e [tʃ] são alofones ou variantes previsíveis de um mesmo segmento abstrato, o fonema /t/.

As diferentes realizações da vibrante múltipla /R/, nas variedades do PB, são um exemplo de alofonia em distribuição livre, ou seja, a realização desse fonema não é fonologicamente motivada. As variantes desvozeadas como a consoante fricativa glotal [h] e a fricativa velar [x] e como também as consoantes vozeadas fricativa velar [ɣ], a fricativa glotal [ɦ] e a retroflexa alveolar [ɹ] podem se

¹⁶ Lembra-se que a representação fonêmica é feita entre barras simples e as variantes é expressa entre colchetes, como se pode ver no capítulo anterior.

alternar livremente, pois não há nada na estrutura do PB que motive essa variação. Silva (2014, p. 133) diz que a variação livre "é condicionada por fatores extralinguísticos como a localização geográfica, grau de escolaridade, classe social, sexo, idade, entre outros."

Esses três conceitos-chave da fonologia fundamentam linguisticamente parte da construção do Potigrafone, uma vez que os fones são os segmentos vocálicos e consonantais encontrados na transcrição fonética, ou seja, no *output* são símbolos fonéticos; os fonemas distinguem funcionalmente das outras unidades da língua de modo que uma transcrição vai distinguir uma palavra da outra e os alofones relacionam a manifestação fonética da variedade potiguar de um fonema da língua portuguesa.

3.2.2 Regras e processos fonológicos

Numa visão gerativista, Seara, Nunes e Volcão (2015, p. 140) afirmam que "o componente fonológico é formado por um conjunto de representações subjacente e por regras que definem como essas representações emergem na superfície." As regras fonológicas atuam na modificação, no acréscimo ou no apagamento dos segmentos das formas de entrada. Por isso, segundo Silva (2011), essas regras são aplicadas às representações *subjacentes*, transformando-as e gerando novas formas de representação até o término do processo derivacional, quando se tem a representação *superficial*.

Os subtópicos a seguir enfatizam a formalização das regras fonológicas e a apresentação dos processos fonológicos presentes na conversão de grafema para o som da variedade potiguar.

3.2.2.1 Regras fonológicas

Silva (2014, p. 198, 199) resume a formalização das três principais regras fonológicas: transformação, inserção e cancelamento dos segmentos. Essas regras podem ser encontradas com mais detalhes em Schane (1975). O quadro 8 abaixo apresenta os principais símbolos utilizados na descrição das regras fonológicas.

Quadro 8 - Símbolos utilizados nas regras fonológicas

Símbolo	Nome	Descrição
---------	------	-----------

_	traço	Serve para marcar a posição exata em que ocorre o Segmento, cujo contexto será caracterizado pelo que o precede e segue.
.	Ponto	Separador silábico
+	Kleene plus	Marca a fronteira entre as sílabas.
/	barra	Serve para indicar o que vem depois é a informação de um contexto, ou seja, separa o ambiente do restante da regra. à esquerda à direita
∅	Conjunto vazio	A eliminação é indicada pelo ∅, mas poderiam ser utilizados <i>colchetes vazios</i> ou 0 (zero)
\$	cifrão	Representa a sílaba.
→	seta	Substituição direta em contexto significa que o elemento à esquerda transforma-se no da direita.
#	Cerquinha	Limite de palavra

Fonte: adaptado de Cagliari (2002, p.29)

a) regra de transformação

$$// \rightarrow [w] / _.\$$$

Lê-se essa regra da seguinte forma: o segmento // transforma-se em glide [w], quando estiver em posição final de sílaba. Exemplo: /sal/ → ['saw] e /salta/ → [saw'ta].

b) regra de cancelamento

$$// \rightarrow \emptyset / 'V_+s$$

Lê-se essa regra da seguinte forma: o segmento // é cancelado, quando precedido de vogal acentuada e seguido do morfema de plural s. Essa regra não é válida para as palavras adjetivas terminadas em <vel>, como <amável>. Porém, na palavra <sal> pode ser aplicada. Exemplo: 'sal+S → ['sa+S].

c) regra de inserção

$$\emptyset \rightarrow [i] / 'V_+s$$

Lê-se essa regra da seguinte forma: insira o segmento [i] quando uma vogal acentuada é seguida do morfema de plural S. Exemplo: 'sa+S → ['saiS] “sais”.

O uso dessas regras tem o propósito de explicitar o mecanismo fonológico e fornecer as descrições estruturais determinantes das mudanças que ocorrem entre as representações gráficas e as representações fonéticas. Essas regras não são aplicadas ao acaso, mas segundo critérios que satisfaçam

determinadas exigências. Além disso, as "regras fonológicas expressam processos fonológicos e idealmente o fazem de maneira simples, econômica e em caráter generalizador." (SILVA, 2014, p. 198).

3.2.2.2 Processos fonológicos¹⁷

Os processos fonológicos "podem ser considerados como um sistema de regras que relaciona a estrutura profunda de um item lexical à sua estrutura fonética (de superfície)." (SEARA, NUNES, VOLCÃO 2015, p.142).

Conforme essas autoras¹⁸, os processos fonológicos são organizados em quatro categorias: assimilação, reestruturação silábica, enfraquecimento e reforço, e neutralização. Todavia, este trabalho destaca apenas três (assimilação, reestruturação silábica e neutralização), as quais constituem a base de transcrição do sistema computacional. Esses processos foram subcategorizados em fenômenos fonéticos¹⁹ referenciados no trabalho de Silva (2011, 2014).

Cada fenômeno está exemplificado por meio de uma palavra retirada do Atlas Linguísticos do Centro-Oeste Potiguar, por Silva (2012) e/ou do Atlas Geolinguístico do Litoral Potiguar, por Pereira (2008). Como definido no tópico 3.2, a forma escrita aparece entre parênteses angulares < > e a forma fonética entre colchetes [] e, entre essas formas, uma seta → que aponta para direita, <A> → [A]. Cada exemplo pode ser visto no apêndice A deste trabalho. Após a exemplificação, são formalizadas as regras que relacionam a forma escrita (forma de entrada do sistema) com a forma fonética (forma de saída do sistema). O tópico 4.3.2 relaciona a estrutura subjacente, que é constituída pela forma escrita, com a estrutura de superfície, que é formada pela forma fonética.

¹⁷ Há autores que, como Schane (1975), preferem a expressão processos fonológicos e outros, como Silva (2014), utilizam a expressão fenômeno fonético. Por uma opção prática de escrita do texto e para melhor compreensão dos fenômenos descritos pelo conversor, procurou-se subcategorizar os processos fonológicos em fenômenos fonéticos. O primeiro tem um sentido amplo e integra os segundos. Essa escolha será compreendida com a exposição do texto.

¹⁸ Optamos pela fundição dos processos de enfraquecimento e reforço no processo de reestruturação, uma vez que o apagamento e o reforço (p. 146) acarretam uma reestruturação silábica. Além disso, com base em Cagliari (2002) e Silva (2011), foram acrescentados e enquadrados, nos processos fonológicos, outros fenômenos não mencionados pelas autoras, como o de nasalidade, mas que fazem parte da descrição fornecida pelo Potigrafone.

¹⁹ As autoras não usam a expressão *fenômeno fonético*, mas a expressão *processo de palatalização*, *de nasalização*, etc. Preferimos a utilização da primeira expressão apenas para diferenciar quando se está tratando de um processo específico e de um conjunto de processos. Um processo específico será designado pela expressão *fenômeno de palatalização*, *de nasalidade*, etc. já o processo de assimilação, por exemplo, engloba os fenômenos da nasalidade, de vozeamento, etc.

3.2.2.2.1 Processo de assimilação

A **assimilação** ocorre quando segmentos diferentes se tornam mais semelhantes, ou seja, um segmento assume os traços distintivos de um segmento vizinho. Dependendo do segmento precedente ou subsequente, a assimilação, segundo Silva (2011), pode ser de dois tipos: regressiva e progressiva. A primeira ocorre quando um som compartilha alguma(s) de sua(s) propriedade(s) com o som precedente. A segunda quando compartilha alguma(s) de sua(s) propriedade(s) com o som subsequente.

Aqui, os fenômenos fonéticos de nasalidade, vozeamento, harmonia vocálica, palatalização e lenição expressam o processo de assimilação.

O *fenômeno da nasalidade*²⁰ ocorre "quando o véu do palato começa a se abrir para a produção da consoante nasal em momento em que a vogal anterior à consoante ainda está sendo produzida." (SEARA, NUNES, VOLCÃO 2015, p.143). Dito de outra forma por Silva (2014, p. 93), "A nasalidade de uma vogal ocorre quando uma vogal tipicamente oral é seguida por uma das consoantes nasais: [m, n, nh]". Esse fenômeno ocorre com maior frequência nas variedades linguísticas do Nordeste e é mais perceptível auditivamente com a vogal [a], todavia, ocorre também com as outras vogais. Logo, a nasalidade é um "fenômeno em que uma vogal oral tem a propriedade opcional de ressonância na cavidade nasal." (SILVA 2011, p. 157).

Esse fenômeno é muito frequente no dialeto potiguar. Para exemplificá-lo, tem-se a palavra <anel> → [a~"new] na carta fonética 51, do Atlas Linguístico do Centro-Oeste Potiguar, onde se pode ver que a vogal oral <a> assimila a nasalidade da consoante subsequente [n].

Esse fenômeno está expresso pela seguinte regra de inserção do traço de nasalidade na vogal que o assimila da consoante subsequente²¹.

²⁰ Foi utilizada a expressão fenômeno da nasalidade, com base em Silva (2011, p. 157), ao invés de processo de nasalização, para que não haja confusão entre os termos nasalização e nasalidade definidos no final do tópico 3.3.2, quando tratamos sobre as vogais nasais.

²¹ A matriz fonética dos traços utilizados nas regras fonológicas está disponível no anexo D deste trabalho.

Figura 1 - Regra de nasalidade

$$V \rightarrow [+nasalizada] / _ .(C) \\ +Nas$$

Fonte: Seara, Nunes, Volcão (2015, p.151).

O *fenômeno de vozeamento* "ocorre quando uma consoante se torna surda ou sonora, dependendo se a consoante adjacente a ela é surda ou sonora, respectivamente." Segundo Seara, Nunes, Volcão (2015, p.153). A carta fonética 55 do atlas mencionado acima e a carta 11 do ALiPTG apresentam a marca desse fenômeno na variedade potiguar nas palavras <mesmo> → [ˈmezmu] e <costas> → [ˈkoSt6s], respectivamente. Nota-se que, no primeiro caso, a consoante fricativa alveolar desvozeada <s> é produzida como uma consoante fricativa alveolar vozeada [z] quando está diante de uma consoante vozeada; no segundo caso, ela é emitida como uma fricativa palatal desvozeada [S], quando está diante de uma consoante desvozeada, ocorrendo, assim, a assimilação do traço vozeado para o primeiro caso e desvozeado para o segundo.

Observando essas palavras, podemos verificar em que condições ocorre esse fenômeno para a formalização de uma regra que estabeleça o fato do vozeamento dessa consoante, como se pode ver na regra a seguir.

Figura 2- Regra de vozeamento

$$\left(\begin{array}{l} +cons \\ -soant \\ +cor \\ -son \end{array} \right) \rightarrow \left[\begin{array}{l} +son \end{array} \right] _ . \# \left(\begin{array}{l} +cons \\ +son \end{array} \right)$$

Fonte: Seara, Nunes, Volcão (2015, p.153).

O *fenômeno da harmonia vocálica* é um processo de assimilação que ocorre com a vogal média alta em posição pretônica. A carta fonética 16 do ALiPTG apresenta um exemplo desse fenômeno na palavra <gordura> → [guhˈdura]. Nesse caso, ocorre mudança na vogal média-alta pretônica (<o>) da palavra <gordura> que foi produzida como uma vogal alta [u], assemelhando-se em altura à vogal tônica, o que resulta na pronúncia [guhˈdura] com uma harmonização da altura da vogal pretônica em relação à vogal tônica da palavra.

Esse fenômeno ocorre também com as vogais médias alta [e] que assimila o traço da vogal alta [i], em posição pré-tônica, emitida como se fosse uma vogal alta [i], com a vogal tônica da palavra. A regra de transformação abaixo mostra a formalização desse fenômeno.

Figura 3 - Regra de harmonia vocálica

$$\overset{V}{[-\text{acento}]} \rightarrow [+alto] \quad / \quad \left[\begin{array}{c} \overset{V}{} \\ +\text{acento} \\ +\text{alto} \end{array} \right]$$

Fonte: Seara, Nunes, Volcão (2015, p.155).

O *fenômeno da palatalização* é o fenômeno que "adquire uma articulação palatal ou próxima à região palatal. No português brasileiro, ocorre a palatalização de oclusivas alveolares antes da vogal alta ou glide palatal" (SILVA, 2011, p. 168). Isso ocorre devido ao posicionamento da língua para a emissão da vogal se sobrepor ao gesto consonantal da consoante adjacente.

Pereira (2008) diz que a consoante alveolar [t] pode ocorrer como variante palatalizada [tʃ] decorrente de contextos em que se tem presente o segmento <to> no final de sílaba, tendo como antecedentes os ditongos orais ou nasais [ej] ou [uj], apresentados nas cartas fonéticas 20 <muito> → [ˈmu~jtSu] e 23 <peito>. → [ˈpejtSu].

Esse fenômeno pode ser formalizado pela seguinte regra de transformação:

Figura 4 - Regra de palatalização

$$\left(\begin{array}{c} \overset{C}{} \\ -\text{cont} \\ +\text{cor} \\ +\text{ant} \\ -\text{voz} \end{array} \right) \rightarrow [-alto] / _ \overset{V}{} \left[\begin{array}{c} +\text{rec} \\ +\text{alt} \end{array} \right]$$

Fonte: Adaptada de Seara, Nunes, Volcão (2015, p.152).

O *fenômeno da lenição* é o "enfraquecimento de um segmento consonantal que se torna mais sonoro ou é produzido com menor grau de constrição

no trato vocal." (SILVA, 2011, p. 141). Esse fenômeno ocorre com as consoantes palatais [L] e [J]. O enfraquecimento da palatal [L] se manifesta por meio de uma despalatalização, como na palavra <mulher> → [mu"IE] e da iotização, como em <telha> → [teja], como afirma Silva (2012) no Atlas Linguístico do Centro-Oeste Potiguar. A regra abaixo apresenta a formalização desse fenômeno.

Figura 5 - Regra de lenição

$$\left(\begin{array}{c} \text{C} \\ -\text{cont} \\ +\text{cor} \\ +\text{ant} \\ -\text{voz} \end{array} \right) \rightarrow \begin{array}{c} [-\text{alto}] \\ [+solt \text{ ret}] \end{array} / _ \left(\begin{array}{c} \text{V} \\ +\text{rec} \end{array} \right)$$

Fonte: Figura elaborada pelo autor.

3.2.2.2.2 Processo de reestruturação silábica

Os processos de **reestruturação silábica** "ocorrem quando há alteração na distribuição das consoantes e das vogais, podendo esses segmentos serem inseridos ou eliminados". (SEARA, NUNES, VOLCÃO 2015, p.144). Os fenômenos fonéticos do apagamento, da monotongação, da epêntese e da ditongação manifestam o processo de reestruturação silábica.

O *apagamento* é um fenômeno fonético em que um segmento consonantal ou vocálico é cancelado. Segundo Silva (2011), esse fenômeno ocorre com as vogais em sílabas átonas e com as consoantes nas bordas das palavras ou em encontros consonantais.

Em palavras proparoxítonas, a penúltima vogal desaparece, transformando a palavra em paroxítona. É o que acontece com a pronúncia potiguar para a palavra <fósforo>. A carta fonética 14 do ALiPTG apresenta uma síncope da vogal postônica, como se pode-se perceber <fósforo> → [ˈfOsfru]. Uma vogal próxima à outra acentuada foi eliminada. O cancelamento das consoantes ocorre, geralmente, nas bordas das palavras ou em encontro consonantal. A carta fonética 10 desse mesmo atlas mostra o apagamento da consoante glotal na palavra <colher> → [koˈLe0].

No primeiro, a estrutura silábica é formada por três sílabas: CVC.CV.CV e passa por uma reestruturação silábica com o apagamento da vogal, passando a constituir-se de duas sílabas: CVC.CCV. Já a estrutura silábica da segunda palavra é formada por duas sílabas: CV.CVC. Com o apagamento da fricativa glotal, permanecem duas sílabas, mas com ausência da consoante pós-vocálica: CV.CV. A transcrição fonética do apagamento dessa consoante é representada pelo símbolo [0] pelo sistema, para indicar que houve apagamento; já a transcrição das síncope de vogais não são contempladas pelo sistema.

Esse fenômeno ocorre com a fricativa glotal, especialmente no final de palavra e com as vogais.

Figura 6 - Regra de apagamento

$$\left[\begin{array}{c} C \\ +cont \\ +glot \end{array} \right] \rightarrow \emptyset \quad / \quad _ \#$$

Fonte: Figura adaptada de Cagliari (2002, p.101).

O fenômeno da *monotongação* acontece quando um ditongo passa a ser produzido como uma única vogal, ou seja, quando há o apagamento dos glides [j] ou [w] em ditongos crescentes ou decrescentes, segundo Silva (2011). As cartas fonéticas 7 e 32 do ALiPTG apresentam a realização desse fenômeno nas palavras <caixa> → [kaSa] e <tesoura> → [tesora], em que ocorrem em ditongos decrescentes. A autora desse atlas, Pereira (2008), acrescenta que os ditongos [ej], [aj], mesmo nos contextos considerados de redução, isto é, diante de [S] e [Z], apresentam a predominância de uso que tende à redução vocálica, com a realização [aj] em [a] nos mesmos contextos fônicos. A monotongação de ditongo crescente é mais frequente na variedade potiguar em posição final, como atesta a palavra <sandália> → [sa~"daLa], na carta fonética 29 desse atlas.

A regra abaixo estabelece que a sequência de vogal seguida dos glides deve se transformar em uma vogal simples. Isso ocorre quando se encontra uma fricativa coronal em posição final de sílaba.

Figura 7 - Regra de monotongação

$$\left[\begin{array}{c} V + V' \\ -cont \\ +sil \\ -alt \\ -bai \\ +ace \\ n \end{array} \right] \rightarrow V / \left[\begin{array}{c} - \\ +con \\ s \\ -sont \\ +con \\ t \end{array} \right]$$

Fonte: Seara, Nunes, Volcão (2015, p.149).

Segundo Silva (2011), o *fenômeno da ditongação*²² pode ocorrer de duas formas: (1) quando uma vogal simples em um hiato passa a ocorrer como um glide, perdendo a propriedade de ocupar o núcleo silábico e (2) quando ocorre uma inserção de um glide após uma vogal, ou seja, a transformação de um monotongo em um ditongo. O primeiro caso se apresenta nos hiatos formados por vogal posterior seguida de vogal média, constituindo um ditongo crescente. A carta fonética 57 do Atlas Linguístico do Centro-Oeste Potiguar mostra esse tipo de fenômeno na palavra <nódoa> → [nOdwa]. O segundo caso se manifesta, geralmente, nas palavras oxítonas terminadas por sibilantes [s] e [z] em posição pós-vocálica em sílabas tônicas finais, como em <paz> → [ˈpajs] e <arroz> → [aˈhojs] (Ver SILVA, 2012).

A regra a seguir prevê que haverá a inserção de um segmento glide, quando uma consoante, em posição final de sílaba, for antecedida por uma vogal acentuada.

Figura 8 - Regra de ditongação

$$\phi \rightarrow \# \left[\begin{array}{c} V' \\ -sil \\ - \\ cons \\ +alt \\ -rec \end{array} \right] / \left[\begin{array}{c} V \\ +acent \end{array} \right] \left[\begin{array}{c} +cons \\ +cor \\ +estr \end{array} \right]$$

Fonte: Figura elaborada pelo próprio autor.

O *fenômeno da epêntese* ocorre quando há acréscimo de um segmento

²² Utilizamos, aqui, o termo ditongação com base em Silva (2014), mas as autoras Seara, Nunes e Lazzarotto-Volcão (2015) descrevem o segundo caso como sendo um processo de reforço com a inserção da semivogal.

vocálico ou consonantal. Para essa autora, ocorre a inserção de consoante ou de vogal na pronúncia do português. O primeiro caso é mais raro e surge apenas em palavras criadas por derivação, como em *cafezal* que é formada de <café> mais o prefixo <al>. Esse caso de epêntese é irrelevante para a transcrição fonética automática, pois a palavra escrita já possui a letra que representa o som equivalente. A inserção da vogal epentética [i] é a mais recorrente no português, todavia, ocorre variação dessa vogal com a vogal média alta [e]. Esse caso de epêntese "se caracteriza pela inserção de uma vogal entre as consoantes em **encontro consonantais** que envolvam oclusivas, africadas, nasais ou fricativas." (SILVA, 2011, p.99, destaque da autora). Além desses dois casos, ocorre epêntese nas palavras estrangeiras, quando são produzidas por falantes nativos do PB, como em *link*, em que se acrescenta a vogal [i] no final da palavra. No entanto, esse caso não é contemplado pela transcrição fonética do Potigrafone.

Na carta fonética 33 do Atlas Linguístico do Centro-Oeste Potiguar, há um exemplo desse fenômeno na palavra <pneu> → [pinew] [penew], onde ocorre a inserção das vogais anterior [i] e [e], formando uma sílaba com a oclusiva bilabial sonora.

A regra a seguir formaliza a inserção das vogais nos encontros consonantais que envolvem oclusivas, nasais e fricativas.

Figura 9 - Regra de epêntese

$$\phi \rightarrow \left(\begin{array}{l} -\text{sil} \\ -\text{cons} \\ +\text{alt} \\ -\text{rec} \end{array} \right) \left(\begin{array}{c} V \\ +\text{acento} \end{array} \right) / \left(\begin{array}{l} +\text{cons} \\ +\text{cor} \\ +\text{estr} \end{array} \right)$$

Fonte: Figura elaborada pelo próprio autor.

3.2.2.2.3 Processo de neutralização

O **processo da neutralização**, para Seara, Nunes, Volcão (2015), ocorre quando os segmentos se fundem em um ambiente específico, como as vogais não acentuadas em final de palavra. O PB possui sete vogais em posição tônica e apenas três vogais átonas em final de palavra. Assim, as vogais finais não acentuadas <e> e <i> são pronunciadas como [i]. Como nas palavras <jure> →

["Zuri] e <júri> → ["Zuri], não distinguidas na transcrição fonética.

Este tópico mostra que os fenômenos fonéticos exigem cuidadosa utilização das regras fonológicas de transformação, de inserção e de cancelamento de segmentos, para que os processos sejam bem descritos e, conseqüentemente, o transdutor possa executar a transcrição automática com eficiência. As regras de transformação têm maior aplicação nos processos de assimilação e neutralização. Os primeiros exigem que os segmentos adquiram um traço de outros segmentos vizinhos na pronúncia pelo falante da língua. A inserção de segmentos está presente nos fenômenos da nasalidade, da ditongação e da epêntese e o cancelamento está presente nos fenômenos do apagamento e da monotongação.

O capítulo seis mostra uma aplicação dessas regras no desenvolvimento *Potigrafone* e no apêndice B exibimos o código-fonte escrito em *Foma*.

3.3 Estrutura silábica e tonicidade das palavras

Neste tópico, mostramos a estrutura silábica como unidade fundamental e organizacional sobre a qual os princípios da sonoridade, da distância mínima e do máximo *onset* devem ser considerados e destacamos a marcação da tonicidade por meio de um diacrítico do alfabeto SAMPA, utilizado para destacar os aspectos das sílabas e a distribuição do acento silábico.

3.3.1 Estrutura silábica do português

Como foi visto no tópico anterior, o processo de reestruturação silábica altera a distribuição dos segmentos, imprimindo nova estrutura silábica às formas de entrada. Numa transcrição fonética, há também outros mecanismos que não acarretam mudança silábica, mas são sensíveis à posição silábica. Exemplo disso é a consoante nasal, por exemplo, em posição pós-vocálica que não se realiza como segmento, pois esta consoante é eliminada após transmitir o traço nasal para a vogal precedente. A seguinte regra de transformação geraria formas estranhas à fala do português: as consoantes nasais devem ser transformadas em um traço de nasalização [~] após vogais; assim, a transcrição da letra <m> na palavra <campo> → ["ka~pu] estaria correta, porém, numa palavra como <cama> → ["ka~a] estaria incorreta. Dessa forma, a estrutura silábica contribui para a resolução de fenômenos

relacionados à escrita e aos processos fonológicos.

No tópico 6.3, que trata da arquitetura e do algoritmo do Potigrafone, o módulo beta-gerador é formado pelos princípios definidos neste tópico. Aqui, apresentamos a estrutura silábica como unidade fundamental e organizacional das palavras, com base no trabalho de Bird (2009), de Silva (2014) e de Hulden (2006), os quais fornecem os conceitos de sílaba, de princípios de sonoridade, da distância mínima e do máximo *onset* ou ataque silábico.

Bird (2009) diz que as sílabas são uma unidade fundamental e organizacional da Fonologia, em primeiro lugar, pelo fato de que, em muitas línguas, as alterações fonológicas são sensíveis à estrutura silábica. Ele exemplifica essa informação com base no fonema oclusivo alveolar /t/ que tem vários alofones em inglês e a escolha de um dos alofones depende do contexto fonológico. Em muitos dialetos do inglês, o /t/ é pronunciado como tepe [r] entre vogais como em <water>.

Mas também nas variantes <atlas> [et.les] e <cactus> [kek.t'es], onde a descrição fonética é dada entre colchetes e as fronteiras silábicas estão marcadas por um ponto, a sílaba padrão no inglês não pode começar com <tl> e, então, o /t/ de <atlas> é silabificado com a vogal precedente, ou seja, como coda. A sílaba final /t/ é regularmente glotalizada em inglês, enquanto o /t/ inicial é regularmente aspirado. Os padrões naturais para estes alofones, em termos de estrutura silábica, foi apresentado pelo autor para o inglês, porém, como podemos ver no próximo subtópico, essa sensibilidade ao contexto silábico também ocorre em português.

A letra <s> possui várias representações fonéticas, dependendo de sua posição na estrutura silábica. Quando ela está em posição intervocálica medial, é pronunciada como uma fricativa alveolar sonora [z]; quando está em posição de coda medial e precedida por um segmento oclusivo linguodental [t], é pronunciada como uma fricativa palatal surda [S].

Outra evidência apontada por esse autor diz respeito ao uso do padrão silábico através de empréstimos linguísticos. Quando uma palavra é emprestada de uma língua para outra, elas devem ser reajustadas conforme os padrões sonoros da língua receptora. A palavra do inglês <link> é produzida com epêntese por falantes nativos do PB, como [li~ki] ou [li~ke], de forma que a estrutura silábica, para essa palavra, é diferente da forma em inglês. Com foi dito, esse fenômeno epentético não foi contemplado pelo sistema.

A terceira fonte de evidência para o uso da estrutura silábica apontada

pelo autor vem da Morfologia. Ele mostra que a língua nicaraguense *Ulwa* possui a posição do infixo possessivo sensível à estrutura silábica. O padrão silábico da *Ulwa* é (C) V (V | C) (C) e qualquer consoante intervocálica é silabificada como *onset* de uma sílaba seguinte, um princípio universal chamado de maximização do onset. Apresentamos, a seguir, apenas um exemplo com duas palavras que foram retiradas do quadro 1.19, da obra desse autor.

Quadro 9 - Máximo onset na língua *Ulwa*

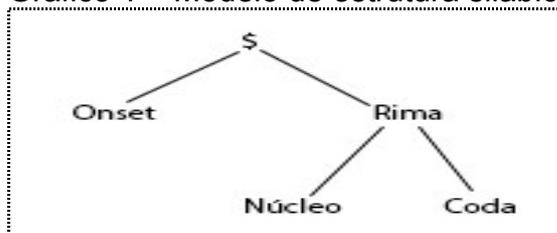
	Palavra	Possessivo	Tradução
a.	Baa	baa. ka	Excremento
b.	Bilam	bi.lam. ka	Peixe

Fonte: adaptado de Bird (2009, p.14).

Esse autor acrescenta que o infixo aparece entre os limites silábicos e podemos afirmar que a posição do infixo é sensível à estrutura silábica. Qualquer análise da posição do infixo deve considerar o peso silábico (*syllable weight*). As sílabas, como tendo uma vogal curta e não seguida de consoantes, são definidas para ser sílabas leves (*light syllable*), enquanto que as outras sílabas - as que têm duas vogais, ou um vogal longa, ou uma consoante final são definidas como sílabas pesadas (*heavy syllable*).

Silva (2014) diz que os *fonemas* se dispõem em padrões identificáveis, chamado de sílaba, que são maiores do que os fonemas e menores do que os morfemas. As sílabas, nas teorias lineares, são unidades fonológicas que se constituem de uma margem esquerda, chamado *onset* (ou ataque), um *núcleo* e uma margem direita, chamada de *coda*. Nas teorias não lineares, a sílaba é composta apenas de *onset* e *rima* e, esta, se divide em *núcleo* e *coda*. É o que podemos ver no gráfico abaixo.

Gráfico 1 – Modelo de estrutura silábica baseado em *onset*-rima



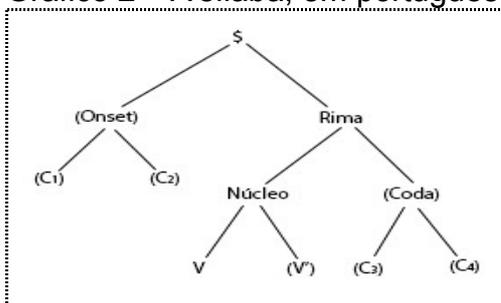
Fonte: adaptado de Silva (2014, p. 205)

Para Bird (2009), nesse modelo de *onset*-rima, as consoantes, vindo antes da primeira vogal, são lincadas para o nó do *onset* e o restante do material seguindo o nó da rima. Uma rima contém, obrigatoriamente, um núcleo e, opcionalmente, uma coda. Nesse modelo, a sílaba é tida como pesada, se e somente se a rima ou o núcleo são ramificados, ou seja, as sílabas são preenchidas com o núcleo e com a coda, ou com o núcleo e glide.

Silva (2014) diz que a silabação primária, apresentada no gráfico 1, deve conter as informações das regras particulares de cada língua, derivando-se uma representação superficial para as palavras da língua, ou seja, um dos constituintes silábicos presentes na estrutura, associa-se a uma ou mais posições da camada CV (C – consoantes, V – vogais).

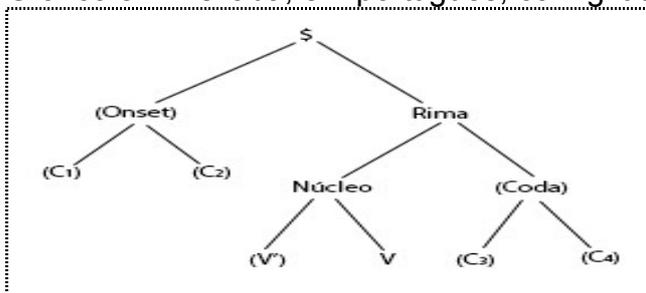
Os gráficos 2 e 3 ilustram a estrutura primária e as camadas CVs. Esses gráficos foram construídos com base no esquema apresentado por Silva (2014, p. 154).

Gráfico 2 – A sílaba, em português, com glide em posição pós-vocálica



Fonte: adaptado de Silva (2014).

Gráfico 3 – A sílaba, em português, com glide em posição pré-vocálica



Fonte: Adaptado de Silva (2014).

Nos gráficos 2 e 3, temos a letra C maiúscula representando os segmentos consoantes que são opcionais - os símbolos entre parenteses. A letra V maiúscula constitui o núcleo da sílaba, formado por segmentos obrigatórios. No gráfico 2, a estrutura silábica $C_1C_2VV'C_3C_4$ representa uma cadeia formada por

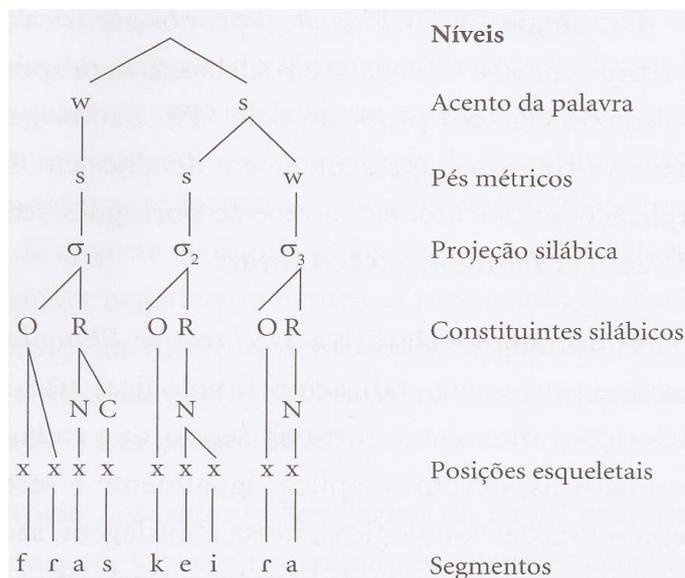
ditongo decrescente (vogal-glide) e, no gráfico 3, a estrutura silábica $C_1C_2VV'C_3C_4$ apresenta uma sequência de ditongo decrescente (glide-vogal).

A ideia apresentada anteriormente remete à estrutura fonológica da sílaba como sendo uma unidade dividida em subconstituintes, semelhante às representações arbóreas da sintaxe. Nessa estrutura, as vogais constituem os itens obrigatórios no sistema e as consoantes e os glides são acessórios, ou seja, a sílaba, em português, deve possuir pelo menos uma vogal para que seja bem formada. O número de segmentos que podem ser associados a determinado constituinte e a ordem com que tais segmentos ocorrem são definidos pelas restrições dos *princípios de sonoridade, da distância mínima e do máximo onset*.

Como se verá no tópico 4.3.2, a representação subjacente, chamada também de *representação lexical*, é organizada em níveis e

A relação entre as representações subjacentes e as representações fonéticas se dá por meio de processos de derivação. Derivações devem seguir os princípios que atuam em cada nível para as derivações finais sejam bem-formadas. (SILVA, 2014, p.205).

Gráfico 4 – Níveis da representação lexical



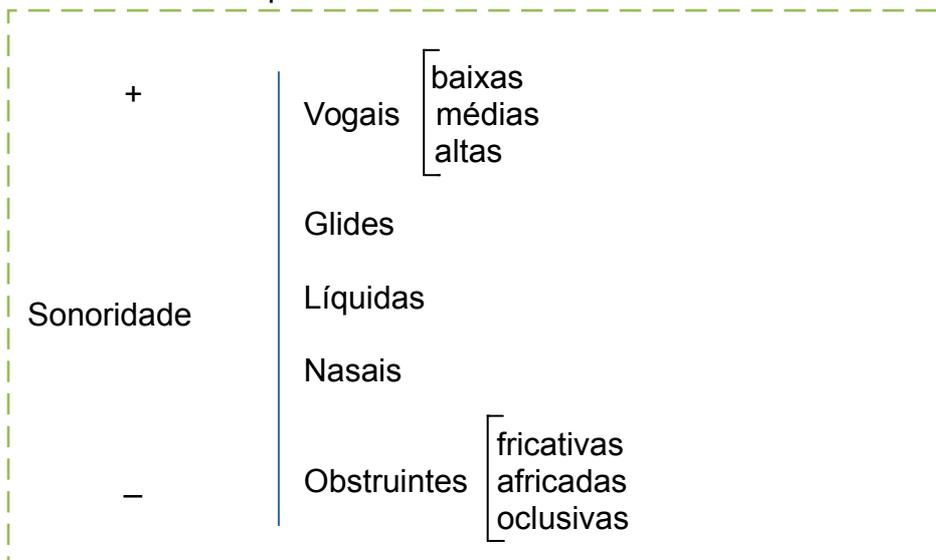
Fonte: Silva, 2011, p. 111.

Essa autora acrescenta que a relação entre as representações subjacentes e as representações fonéticas ocorrem por meio de processos derivacionais. As derivações seguem os princípios que atuam em cada nível para que elas sejam bem formadas, implicando que, pela violação de algum princípio, ocorrerá a geração de estruturas malformadas. “Os princípios são estabelecidos a

partir das evidências linguísticas e descrições estruturais das línguas naturais.” (SILVA, 2011, p. 111IDEM).

Para Silva (2014, p.207), o *princípio da sonoridade* é “uma gradação referente ao grau de abertura do trato vocal durante a produção dos sons e da quantidade de energia produzida durante a produção de um som.” Tal princípio expressa a tendência de que os segmentos mais sonoros ficam mais próximos do núcleo e os menos sonoros ocupam as posições periféricas. Em português, as líquidas são mais sonoras do que as nasais e do que as obstruintes e, por isso, elas estão mais próximas do núcleo num ataque complexo. As línguas que contêm conjuntos complexos de consoantes, como foi visto nos gráficos 2 e 3, são geralmente orientados na sua estrutura silábica pelo conceito de uma *hierarquia de sonoridade* que pode ser observada no gráfico abaixo.

Gráfico 5 – Hierarquia de sonoridade e a escala de sonância



Fonte: Silva (2014, p.207)

Essa escala permite a classificação dos segmentos em termos de sonoridade e a organização silábica, de modo que os elementos respeitem as seguintes condições: a) não é permitida a sequência de elementos no *onset* ou na coda de mesma escala sonância; b) o núcleo da sílaba é constituído pelo elemento mais sonoro da sequência de som e c) a escala cresce de sonância do *onset* para o núcleo e decresce do núcleo para coda, segundo Seara, Nunes e Volcão (2015).

Os segmentos que possuem o nível máximo de sonoridade na sílaba

ocupam o pico silábico ou a posição nuclear da sílaba. Em português, somente as vogais ocupam esta posição. O gráfico 5 ilustra a escala de sonoridade, pois as vogais e os glides estão associados a núcleos por apresentarem uma sonoridade alta, porém apenas as primeiras podem constituir núcleo silábico.

A posição periférica é ocupada pelas consoantes líquidas, nasais e obstruintes. O português apresenta uma *distância de sonoridade* menor ou igual líquida para a segunda consoante do ataque complexo. Na sequência [pl], como na palavra <placa>, há um aumento de sonoridade que sai de uma oclusiva [p] para uma líquida [l]. Todavia, quando a realização da sequência é, por exemplo, [pn], como na palavra <pneu>, em que a sonoridade aumenta de uma obstruinte para uma nasal, há a inserção de uma vogal entre as duas consoantes – [pinew] ou [penew].

Esse princípio explica a realização da epêntese e a formação do ataque e da coda complexos em português. Esse princípio foi considerado na constituição do sistema para a aplicação da inserção de segmentos fonéticos e na formação silábica.

Hulden (2006) apresenta uma aplicação desse princípio na fonologia de estados finitos. Ele afirma que os estados principais são os elementos mais sonoros e formam o núcleo da sílaba, no qual agrupa os segmentos menos sonoros. O *onset* de uma sílaba espelha a coda no nível de sonoridade, pois enquanto o primeiro cresce em sonoridade o segundo decresce, ou seja, o nível de sonoridade do onset vai das obstruintes às líquidas e o nível da coda vai das líquidas às obstruintes. A fim de ilustrar a abordagem no sistema *Foma*, consideramos a ilustração no quadro 10 que apresenta a constituição da sílaba na linguagem de programação

Quadro 10 – Constituição da sílaba por meio da linguagem Foma

```
define Onset [(Obstruintes) (Nasais) (Líquidas)];
define Coda [(Líquidas) (Nasais) (Obstruintes)];
define Rima [vogais Coda];
define sílaba [Onset Rima];
```

Fonte: Hulden, 2006, p.89.

As entradas (*input*) para as sílabas são sequências de consoantes e vogais. Uma sílaba é uma sequência de um início opcional em *onset*, seguido por

um núcleo e por uma coda opcional. Desse modo, a definição de sílaba consiste de *onset* e rima, o qual a coda é a imagem do outro na hierarquia de sonoridade. Outra generalização sobre a sílaba é que tem uma escolha através da filiação de uma consoante para uma vogal ou para um *onset*.

Devido a esse princípio, nem todas as consoantes, no caso da língua portuguesa, preenchem o início das palavras e nem todas têm o mesmo *status* de coda. Nem toda sílaba, no padrão silábico apresentado no gráfico 1, preenchem necessariamente esses constituintes, pois a única unidade obrigatória é o núcleo. Este é formado somente por vogais ou por vogal e glide e se constitui no pico silábico. Em outras línguas, entretanto, algumas consoantes podem preencher esta posição. Às margens do núcleo, estão o *onset* e a coda. Este pode ser preenchido pelas consoantes e aquele por uma consoante, quando ocorre ataque simples (ver ilustração nos gráficos 2 e 3) ou por duas consoantes, quando o corre ataque complexo.

3.3.2 Tonicidade das palavras

Nas palavras com mais de uma sílaba, sempre existe uma sílaba que é pronunciada com maior intensidade do que as outras, mas isso não ocorre na escrita. A marcação da tonicidade das palavras é uma tarefa de maior importância para o conversor gráfico fônico. Limita-se a destacar os aspectos das sílabas e a distribuição do acento silábico na palavra do PB.

A sílaba tônica é produzida com maior pulso torácico, portanto, audivelmente percebida como sendo mais alta e de duração mais longa. Ela é marcada por meio de aspa dupla antes da sílaba, como na palavra < sabia > [sa"bia] em que a aspa indica que a sílaba subsequente é tônica. "As vogais tônicas estão em oposição às vogais átonas. Vogais átonas podem ser pretônicas ou pós-tônicas. Vogais pretônicas antecedem o acentoônico e vogais postônicas sucedem o acentoônico." (SILVA, 2014, p.77).

A distribuição do acentoônico nas palavras do PB ocorre de forma bastante previsível. A acentuação só pode recair nas três últimas sílabas da palavra. Segundo Seara, Nunes e Volcão (2015), as palavras proparoxítonas geralmente têm origem estrangeira e apresentam uma forte tendência ao apagamento na pronúncia da sílaba pós-tônica, como acontece com a palavra < fósforo > produzida como

[ˈfɔsfru]. A maioria das palavras oxítonas é finalizada por consoante, pois "O acento em PB é sensível ao peso da sílaba: o acento tem por tendência cair na sílaba finalizada por consoante - as sílabas travadas ou pesadas." (SEARA, NUNES e VOLCÃO, 2015, p. 128). As palavras paroxítonas constituem maior quantidade no PB e são preferidas quando a última sílaba é terminada por vogal.

Os aspectos da estrutura silábica e a distribuição acentual na palavra permitem a construção de um sistema que tenta descrever a relação gráfica e fônica de forma mais integrada e interdependente. A integração ocorre na concatenação dos elementos na divisão silábica que serve tanto para aplicação das regras fonológicas quanto para a marcação da tonicidade. A interdependência se apresenta na própria aplicação das regras em cascata. Existem algumas regras que exigem a utilização das regras precedentes. Por exemplo, a regra da coda depende da separação silábica bem formada para que sua aplicação seja satisfatória na transcrição. Se a separação silábica da palavra < cama > for < ca.ma >, então, essa regra não se aplica a essa situação, mas se for < cam.a >, então aplica-se a transcrição, mas de forma errada.

Desse modo, atendendo ao contexto do grafema, o Potigrafone considera que, se uma sílaba contém uma vogal com o acento gráfico, então, a sílaba é identificada como sendo tônica, logo, recebe a marca de tonicidade. Se não há marca gráfica, então, o sistema analisa os casos particulares:

- a) se a palavra termina em <a>, <e> ou <o>, seguida ou não de <m|n|s>, então a sílaba <\$> anterior à que contém <a>, <e> ou <o> é uma sílaba tônica ($\$_{\text{tônica}}$). Exemplo: < casa >, < cartas >, < cantam >.
- b) se a palavra termina em <l>, <r>, <x> ou <z>, então <\$> anterior será uma <\$_{tônica}>. Exemplo: cantar, papel, capaz, duplex.
- c) se a palavra termina em <i> ou <u>, seguida ou não de <m|n|s>, então, <\$> que contém <i> ou <u> será uma <\$_{tônica}>. Exemplo: paris, comum, Jesus, delfim.

O sistema faz apenas a marcação da sílaba tônica primária. As palavras derivadas morfológicamente, como é o caso dos advérbios de modo cuja terminação é <mente>, como em <rápido> → <rapidamente> - que deveria receber a marcação do acento secundário.

3.4 Conceitos computacionais

Nos tópicos anteriores, deste capítulo, mostramos os principais conceitos linguísticos que orientaram a implementação do conversor Potigrafone. Neste tópico, no entanto, apresentaremos os aspectos computacionais e formais desse sistema, destacando os conceitos de língua formal, de expressão regular e de transdutores de estados finitos além da distinção entre a forma subjacente e a forma de superfície, cuja ideia está relacionada às formas de entrada e às formas de saída, vistas no início deste capítulo.

3.4.1 Língua formal, expressão regular e rede de estados finitos

O termo língua é usado aqui no sentido geral para se referir a um conjunto de cadeias de caracteres, que são chamadas de *string*. Isso implica dizer que "Uma *língua formal* é um conjunto formado por sequências resultantes da concatenação de elementos extraídos de um conjunto de símbolos, chamado *alfabeto* ou *sigma*". (ALENCAR, 2011, p.19). Para melhor compreensão desse conceito, suponham os elementos do conjunto $\Sigma = \{b, a, h, !\}$, que podem ser repetidos e concatenados entre si. A partir de Σ pode ser construída uma língua formal L_1 , a qual pode ter como palavras <baaah!> ou <bah!>, etc.

Alencar (2011, p.21) diz que uma língua formal é caracterizada pela "enumeração exaustiva de seus elementos" e pela "especificação de 'um critério de pertinência que é satisfeito por todos os elementos do conjunto e somente por esses elementos'". A língua L_1 pode ser definida por meio das expressões regulares $R_1 = \{b, a, a, h, !\}$ ou $R_2 = b a^+ h !$. Em R_1 , a língua é definida pela enumeração dos elementos e, em R_2 , por critérios de pertinência. Portanto, do conjunto formado pelos elementos, pode-se gerar um número de palavras que é aceito pela linguagem formal, ou seja, as regras para a linguagem são explicitamente declaradas em termos das cadeias de símbolos em que podem ocorrer.

A concatenação dos símbolos de L_1 não ocorre de qualquer forma. A palavra <ab!h>, por exemplo, não pertence a essa língua, mesmo que todos esses elementos estejam contidos em L_1 , pois foi definida uma estrutura gramatical para ela. O primeiro elemento da língua deve ser obrigatoriamente um , depois um ou mais <a>, com um <h> seguido de um sinal de exclamação <!>. Isso é chamado de

gramática de uma língua formal. "A gramática de uma língua formal é um mecanismo capaz de produzir todas e somente as cadeias pertencentes a essa língua", ou seja, são "dispositivos capazes de gerar uma língua". (ALENCAR, 2011, p.22). Na definição de uma língua formal, considera-se sempre a existência de uma gramática que a constitui e também os autômatos que são "os dispositivos capazes de reconhecer uma língua, isto é, determinar se uma dada sequência é elemento da língua." (ALENCAR, 2011, p. 22). Portanto,

Autômatos e gramáticas, não obstante constituírem dispositivos computacionais diferentes, estão estreitamente relacionados, porque, dado um autômato capaz de reconhecer uma dada língua formal, uma gramática capaz de gerar essa língua pode ser construída automaticamente, ou seja, de forma algorítmica. (ALENCAR, 2011, p. 23).

Esse autor acrescenta que a construção de um algoritmo capaz de gerar as cadeias pertencentes a uma língua é caracterizada por meio de uma gramática que consiste de quatro elementos ordenados, a saber $G = (\Phi, \Sigma, \Gamma, S)$. Temos, abaixo, a descrição de cada símbolo:

- a) Φ é um alfabeto de símbolos não terminais;
- b) Σ é um alfabeto de símbolos terminais, com a condição de que $\Phi \cap \Sigma = \emptyset$ (com a condição que a intersecção dos símbolos não terminais com os símbolos terminais seja vazia);
- c) $\Gamma = \Phi \cup \Sigma$, onde Γ é a união dos símbolos não terminais com os símbolos terminais;
- d) R é um conjunto de regras de reescrita do tipo $\langle \alpha, \beta \rangle$, ou seja, α é substituível por β ;
- e) S é o símbolo inicial com a condição de que $S \in \Phi$.

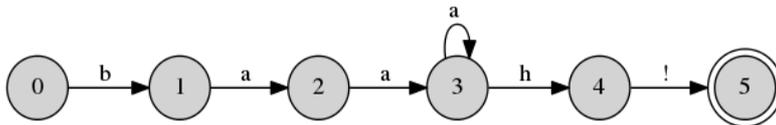
Os elementos que compõem o conjunto de símbolos terminais Σ da gramática são as vinte e seis letras do alfabeto latino, como foram mostradas no tópico sobre *formas gráficas da língua*, e está presente no módulo arquivo do sistema, no tópico 6.3. Os símbolos não terminais Φ são as letras dos *constituintes silábicos*. Não existe intersecção entre os elementos do primeiro alfabeto Σ com os símbolos do segundo Φ , como pode ser comprovado no módulo arquivo, no apêndice B. O conjunto de regras de reescrita é aplicado aos elementos de entrada que os substituem no processo derivacional.

Uma língua formal é declarada por meio de uma expressão regular. Uma *expressão regular* é uma fórmula numa linguagem especial que especifica uma

classe de *string*²³. Em outras palavras, ela "é uma composição de símbolos, caracteres com funções especiais, que, agrupados entre si e com caracteres literais, formam uma sequência, uma expressão." (JARBAS, 2012, p.19). Ela é interpretada como uma regra que indicará sucesso, se uma entrada (*input*) de dados qualquer unir-se com essa regra, ou seja, obedecer exatamente a todas as condições.

A língua L_1 foi denotada por duas expressões regulares R_1 ou R_2 , as quais são compiladas por meio de uma *rede de estados finitos*. A figura abaixo, também chamada de *grafo de transição*, mostra a rede de estados finitos - doravante *finite states transducer* (FST) - produzida por essas expressões regulares.

Figura 10 – Grafo de transição de L_1



Fonte: Elaborada pelo próprio autor.

Em uma FST, os círculos representam os estados e estes possuem um rótulo que é o nome da entrada para o estado. No Foma²⁴, os rótulos recebem como nome um número. O *estado inicial* está assinalado pelo círculo que contém o rótulo zero (0) e o *estado final* é representado com duplo círculo, neste caso, o círculo que contém o rótulo cinco (5). As setas (ou *arestas*) representam os *arcos* e indicam o caminho dos estados. O rótulo de um arco é um símbolo do alfabeto da língua, ou seja, um elemento do *sigma* que, nessa FST, é constituído de quatro elementos, $\Sigma = \{!, a, b, h\}$.

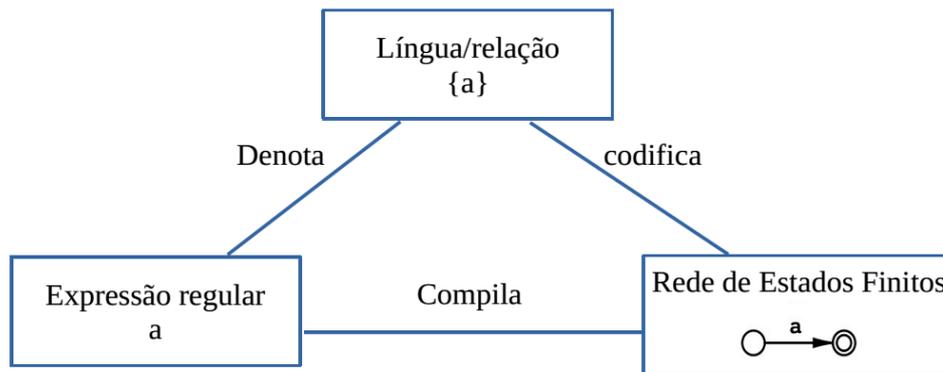
Beesley e Karttunen (2002, p. 44) resumizam dizendo que “uma expressão regular denota um conjunto de cadeias (isto é, uma linguagem) ou pares de cadeias (isto é, uma relação). Ela é compilada por uma rede de estados finitos que codificam a linguagem correspondente ou a relação que pode ser infinita”²⁵, como mostra a figura abaixo.

²³ Neste trabalho, utilizamos os termos *string*, em inglês, ou cadeia de caracteres - a tradução - para referirmo-nos à concatenação de um, dois ou mais elementos do alfabeto terminal e não terminal.

²⁴ No capítulo 6 mostramos com mais detalhes o uso da concatenação, união, intersecção de elementos e a formação das redes de estados finitos por meio da linguagem Foma.

²⁵ Texto original: A regular expression *denotes* a set of strings (i.e. a language) or string pair (i.e. a relation). It can be *compiled* into a finite-state network that compactly *encodes* the corresponding language or relation that may well be infinite.

Figura 11 - Língua formal, expressão regular e rede de estados finitos



Fonte: adaptado de Beesley e Karttunen (2002, p. 44).

Portanto, Beesley e Karttunen (2002, p. 46) afirmam que “Um dos resultados fundamentais da teoria da língua formal é a demonstração de que os estados finitos de uma língua são precisamente um conjunto de línguas que podem ser descritas por uma **expressão regular**.”²⁶ (Destaque dos autores). Esta inclui um conjunto de operadores lógicos que diferem de uma linguagem de programação para outra.

3.4.2 Formas subjacentes e formas de superfície

Os transdutores de estados finitos são dispositivos capazes de relacionar uma cadeia de entrada a uma cadeia de saída. No subtópico acima, foram vistas que as expressões regulares denotam um conjunto de cadeias ou um conjunto de pares ordenados de cadeias. A relação entre grafema e fone de uma língua natural é constituída por pares ordenados, em que um ou mais símbolos do alfabeto gráfico corresponde a um ou mais símbolo do alfabeto fonético.

Como pudemos ver no capítulo 3, os alfabetos ortográfico e fonológico do PB são conjuntos de símbolos que podem ser relacionados entre si. Essa relação se formaliza pelo fato de haver uma definição regular entre os dois alfabetos, como diz Kornai:

Nós definimos as relações regulares de forma análoga ao caso de linguagens regulares: dado dois alfabetos P e Q, uma relação $R \subset P \times Q$ é regular se e somente se for finitamente gerado a partir de conjuntos finitos pelas operações de união, concatenação, e fechamento Kleene. As relações regulares estão na mesma relação com FSTs como linguagens regulares

²⁶ One of the fundamental results of formal language theory is the demonstration that finite-state languages are precisely the set of languages that can be described by a **regular expression**.

são FSAs. Na verdade, é conveniente pensar em línguas como relações unários.²⁷ (KORNAI, 2008, p.40).

A letra P, nessa citação, representa um conjunto de símbolos gráficos e a letra Q²⁸, o conjunto de sons. A relação entre esses conjuntos constituem a língua L, se, e somente se, o produto cruzado de P em Q estiver contido em L.

Como pudemos ver no capítulo 3, o conjunto P foi subdividido em dois subconjuntos: o primeiro, contendo as cinco letras que representam os segmentos vocálicos e o segundo, contendo as dezessete letras que representam os segmentos consonantais. Há também um elemento sem correspondência fonológica, ou seja, um símbolo mudo, o *h*, usado por herança etimológica, e três de uso restrito, apenas com aplicação às palavras estrangeiras. O conjunto Q é constituído por trinta e três símbolos fonéticos do SAMPA e está subdividido em *sete* símbolos que representam as vogais tônicas, *cinco* que representam as vogais nasais, *duas* para representarem os glides e *dezenove* que representam as consoantes.

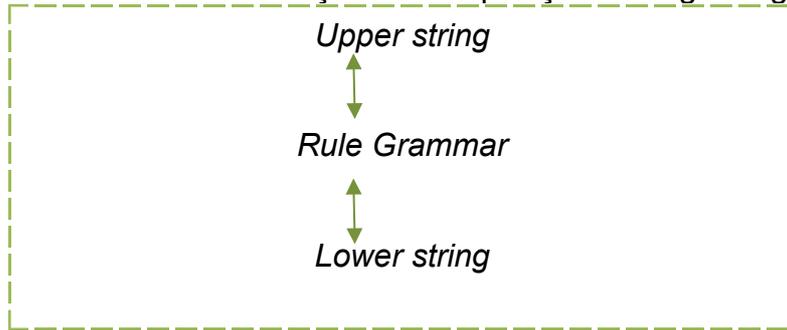
Considerando esses conjuntos, os pares ordenados das cadeias de caracteres formam um transdutor de estados finitos. Karttunen (2009) diz que o primeiro membro de um par da relação é chamado de cadeia superior (*upper string*) e o segundo é chamado de cadeia inferior (*lower string*). Cada caminho do transdutor representa um par de *string*: uma entrada com uma saída, ou seja, a forma escrita com a forma fonética.

A rede de estados finitos representa somente um subconjunto de todas as possibilidades da língua para a relação da entrada com a saída, como podemos ver na figura abaixo.

²⁷ We define regular relations analogously to the case of regular languages: given two alphabets P; Q, a relation $R \subseteq P \times Q$ is regular iff it is finitely generated from finite sets by the operations of union, concatenation, and Kleene closure. Regular relations are in the same relationship to FSTs as regular languages are to FSAs. In fact, it is convenient to think of languages as unary relations. (KORNAI, 2008, p.40).

²⁸ Seguimos a mesma notação que foi utilizada por Kornai (2008), todavia, o conjunto de símbolos gráficos poderia ser representado pelo símbolo matemático Σ , mostrado na alínea b do tópico anterior.

Quadro 11 - Visualização da composição de língua regular



Fonte: adaptado de Beesley e Karttunen (2002, p.281).

Para o transdutor gráfico e fônico, as representações gráficas são cadeias do nível superior e as representações fonéticas são cadeias do nível inferior. Em outras palavras, a primeira cadeia no transdutor Potigrafone é constituída pelas formas gráficas da língua e a segunda pelas formas fonéticas da variedade linguística potiguar.

Nessa relação, como bem mostra o quadro 11, acima, a regra de reescrita está na sequência intermediária e

consiste em uma sequência ordenada de regras de reescrita que converteram as representações fonológicas abstratas em formas de superfície através de uma série de representações intermediárias. Cada regra de reescrita tem a forma geral $\alpha \rightarrow \beta / \gamma _ \delta$, onde α , β , γ e δ podem ser arbitrariamente cadeias complexas ou matrizes de traços. (BEESLEY e KARTTUNEN, 2002, p.309).²⁹

As regras são aplicadas em cascata com estágios intermediários, gerando as formas de superfície. A tarefa no desenvolvimento de um transdutor de grafema para pronúncia é criar uma cascata de regras que mapeiam as cadeias ortográficas do português (*lexical side*) para as cadeias fonéticas (*surface side*) que representam a pronúncia da cadeia de entrada.

Schane (1975) reforça essa ideia acima, dizendo que, quando se trabalha com uma língua e se observam as alternâncias fonológicas governadas por regras, deve-se então determinar as representações subjacentes e as regras necessárias que derivem todas as variantes, a partir da representação superficial. As regras fonológicas atuam entre a *representação fonológica*, como uma cadeia de símbolos gráficos, e a *representação fonética*, como uma cadeia de símbolos fonéticos.

Beesley e Karttunen (2009) dizem que os linguistas têm descrito

²⁹ consisted of an ordered sequence of REWRITE RULES that converted abstract phonological representations into surface forms through a series of intermediate representations. Such rewrite rules have the general form $\alpha \rightarrow \beta / \gamma _ \delta$, where α , β e γ and δ can be arbitrarily complex strings or feature-matrices. (BEESLEY e KARTTUNEN, 2002, p.309).

alternâncias fonológicas e mudanças sonoras em termos de *formas subjacentes*, com regras de reescrita, as quais são aplicadas para determinada finalidade na saída. A saída de uma regra de reescrita é uma sequência intermediária, com a saída da regra final sendo uma *forma de superfície*.

Exemplificamos essa diferença por meio de dois exemplos a seguir. Estes exemplos serão retomados também para exemplificar a execução da transcrição fonética do Potigrafone, no próximo capítulo.

(1)

Representação gráfica	<caixa>	<teia>
Regra fonológica	Redução das vogais pós-tônica final.	
Representação fonética	["kaiS6]	["teia]

(2)

Representação gráfica	<caixa>	<teia>
Regra fonológica	Transformação de vogais altas	
Representação fonética	["kajS6]	["teja]

Os exemplos (1) e (2) mostram as representações subjacentes das palavras /caixa/ e /teia/³⁰, respectivamente, sobre as quais foram aplicadas duas regras de transformação. A regra fonológica de *redução vocálica pós-tônica final* que alteram as representações fonéticas da palavra /kaiSa/, mas não altera a representação fonética de /teia/, de modo que a representação fonética do fonema /a/ será diferente para as palavras. No sistema fonológico do português, está prevista a mudança de um segmento, que possui traço alto, para se transformar em uma vogal não alta, quando ocorrido após uma sílaba tônica em final de palavra.

O exemplo (2) apresenta a aplicação da regra de transformação que altera as representações fonéticas daquelas palavras. Como foi dito acima, podemos perceber que as regras se caracterizam pela forma simples, econômica e de caráter generalizador, servindo para todos os contextos em que ocorre o fenômeno.

Nesse capítulo, consideramos que os transdutores de estados finitos são um exemplo bem-sucedido, quando a informação fonológica é tratada como uma cadeia de símbolos, segundo Bird (2009). As cadeias superior e inferior estão relacionadas de modo que uma letra pode representar um ou mais sons, ou seja,

³⁰ Estes exemplos foram retirados do ALiPTG e do Atlas Linguístico do Centro-Oeste Potiguar. Todavia, para quem quer aprofundar o conhecimento sobre os conceitos, recomenda-se a leitura de Schane (1975).

uma forma gráfica pode representar um ou mais sons da variedade linguística potiguar. A letra <s>, por exemplo, muda de sonoridade, dependendo do contexto silábico: quando está em posição de ataque, ele é produzido como uma fricativa alveolar [s], e quando está em posição pós-vocálica, precedida por uma oclusiva alveolar, é emitido como uma fricativa alvéolo-palatal [ʃ], porém, permanece o mesmo valor fonológico, apesar da variação; não muda a palavra. Desse modo, as formas de entrada do sistema destacam algumas peculiaridades como por exemplo: o uso dos diacríticos e a utilização das letras diacríticas consonantais e vocálicas na escrita dos dígrafos.

A escolha das formas de saída apresentam que os símbolos fonéticos do SAMPA são menos conhecidos pela comunidade dos linguistas; todavia, a utilização do AFI exige, em muitos casos, a instalação da fonte fonética adequada, para que o computador possa reconhecer os caracteres. Isso não acontece com o SAMPA, uma vez que todos os símbolos já estão disponíveis no teclado.

Além disso, a exposição dos conceitos foi relacionada à variedade linguística potiguar, por meio das tabelas que listaram os sons consonantais, os sons vocálicos e os encontros vocálicos pelo alfabeto fonético SAMPA. Nesse ponto, destacamos que as unidades gráficas foram relacionadas aos *símbolos fonéticos* sem levar em consideração os fenômenos fonéticos, pois temos aqui a preocupação de constituir uma visão geral do inventário fonético potiguar para nortear a transcrição do conversor.

Acrescentamos que o conceito de estrutura silábica direciona o transdutor para o reconhecimento das sílabas que pertencem ao português, para a separação silábica das palavras e, conseqüentemente, para a marcação da sílaba tônica. Essa ideia fica mais clara no próximo capítulo, quando estivermos pontuando sobre os módulos do sistema.

4 METODOLOGIA, IMPLEMENTAÇÃO E ANÁLISE DO SISTEMA

Nos capítulos anteriores, apresentamos os sistemas gráficos e fônicos para o português; definimos as formas de entrada e saída que correspondem, no transdutor, às formas profundas e às formas de superfície, respectivamente e apresentamos os conceitos linguísticos e computacionais que norteiam a implementação do Potigrafone.

Neste capítulo, mostramos os recursos utilizados no desenvolvimento e na avaliação desse sistema computacional. Primeiro, fizemos a identificação e a separação dos padrões fonológicos presentes no Atlas Linguístico do Centro-Oeste Potiguar e no ALiPTG; segundo, elaboramos a transcrição desses *corpora* para o sistema fonético SAMPA, que se encontra no apêndice A deste trabalho; terceiro, desenvolvemos o conversor, baseado na fonologia de estados finitos, para a transcrição das unidades gráficas da língua em segmentos fônicos dessa variedade linguística e, por último, executamos o sistema em *corpus* escrito de língua portuguesa para verificar o desempenho do sistema.

Considerando esses procedimentos, definimos alguns critérios pertinentes ao desenvolvimento do sistema: as entradas (*input*) devem ser formas gráficas do português, como apresentamos no tópico 3.1.1; essas formas não devem conter caracteres maiúsculos nem o sinal de hifenização, “-”; e não podem ser dígitos. Esses critérios foram definidos com base nos atlas linguísticos mencionados acima. Esses atlas não apresentam palavras estrangeiras, nomes próprios, palavras compostas e nem dígitos.

A escolha desses critérios trazem algumas consequências na aplicação do sistema. O conversor não reconhece as palavras que não fazem parte da língua portuguesa, ou seja, as palavras estrangeiras. As letras <k>, <w> e <y> não fazem parte do alfabeto do sistema, como foi destacado na nota de rodapé nº 9. Elas fazem parte do alfabeto português, todavia, são utilizadas apenas em palavras estrangeiras e em nomes próprios. O transdutor não faz a transcrição dos símbolos árabicos, uma vez que esses símbolos representam palavras da língua. Além disso, o Potigrafone não transcreve as palavras com o uso do hífen, uma vez que esse tipo de palavra exige a utilização de segmentador de palavras.

Embora tais critérios reduzam o campo de capacidade de transcrição do sistema, contribuem para um sistema mais eficiente na execução, visto que o

sistema não exige o uso de um dicionário de estrangeirismo ou de uma lista adicional. Além disso, os critérios se adequam aos propósitos definidos pelos *corpora*, de modo que a transcrição (*output*) caracteriza as particulares linguísticas da fala potiguar.

Feitas essas ressalvas, passamos a apresentar os *corpora* de referência utilizados na transcrição fonética, ou seja, a base empírica da qual foram retirados os fenômenos fonéticos presentes na transcrição automática do sistema; a implementação dos módulos do transdutor; as métricas utilizadas na avaliação do sistema e o desempenho do sistema.

4.1 Base empírica

A base empírica do sistema computacional parte de dois importantes atlas linguísticos, os quais apresentam as transcrições fonéticas da pronúncia potiguar: o *Atlas Linguístico do Centro-Oeste Potiguar* e o *Atlas Geolinguístico do Litoral Potiguar* (ALiPTG) que fazem parte do projeto do Atlas Linguístico do Brasil (ALiB). A transcrição fonética desses atlas foram feitas por linguistas e contempla as normas definidas no Projeto ALiB.

Segundo Mota e Cardoso (1997), o ALiB³¹ nasceu e ganhou forma nos fins de 1986, durante o Seminário Caminhos e Perspectivas para a Geolinguística no Brasil, realizado no Instituto de Letras da Universidade Federal da Bahia (UFBA), em Salvador. A partir daí, inicia-se um conjunto de atividades que lhe confere um caráter nacional, tanto pela sua abrangência quanto pela distribuição espacial da equipe que o dirige.

Esse projeto tem os seguintes objetivos: descrever a realidade linguística do Brasil com enfoque prioritário na identificação das diferenças diatópicas (fônicas, morfossintáticas e léxico-semânticas); oferecer aos estudiosos da língua portuguesa (linguistas, lexicólogos, etimólogos, filólogos, etc.) subsídios para o aprimoramento do ensino e aprendizagem e para uma melhor interpretação do caráter multidialetal do Brasil; estabelecer isoglossas com vistas a traçar a divisão dialetal do Brasil, tornando evidentes as diferenças regionais, através de resultados cartografados em mapas linguísticos e oferecer aos interessados nos estudos linguísticos um imenso volume de dados que permitirá aos lexicógrafos aprimorarem os dicionários,

³¹ Para mais informações, acesse o sítio do ALiB: <https://twiki.ufba.br/twiki/bin/view/Alib/WebHome>.

ampliando o campo de informações.

Não está proposto como objetivo do ALiB, mas poderia ser acrescentado que os atlas oferecem condições suficientes para que os desenvolvedores de sistemas implementem programas com base em informações seguras da realidade linguística documentada pela pesquisa empírica.

Os dados apresentados pelos mapas linguísticos desses *corpora* contribuem na identificação e na quantificação dos fenômenos linguísticos, para cada ponto de inquérito. Além disso, esses dois atlas correspondem à descrição dos fenômenos da maior parte do Estado, como podemos ver abaixo.

Uma característica desses atlas linguísticos é o fato de que a recolha de dados se processa mediante a utilização de ficha de identificação do informante com três questionários que seguem integralmente o modelo do Projeto ALiB. Os modelos usados na pesquisa são o questionário semântico-lexical (QSL), com 202 perguntas que cobrem 14 áreas semânticas; o questionário Fonético-fonológico (QFF), com 159 perguntas e o questionário morfossintático (QMS), com 49 perguntas. (PEREIRA, 2008).

Todavia, foi dado enfoque aos fatos linguísticos apresentados nas cartas fonéticas dos atlas mencionados, pois essas cartas

registram as variantes de um fonema comprovadas nos pontos investigados, ou de vários fonemas correspondentes a um único fonema mais antigo ou também determinadas séries de fonemas quase encontram na mesma situação do ponto de vista histórico. (SILVA, 2012, p. 110).

Ressaltamos que os atlas a seguir apresentam a diversidade de usos em sua distribuição espacial, com resultados acrescidos de algumas notas e ilustrações que complementam as informações sociolinguístico-culturais. Eles não contêm interpretação de dados, mas apenas análise linguística.

4.1.1 Atlas Linguísticos do Centro-Oeste Potiguar

Segundo Silva (2012)³², o Atlas Linguístico do Centro-Oeste Potiguar foi elaborado a partir da identificação das variáveis extralinguísticas diastrática, diassexual e diageracional dos fenômenos fonéticos e lexicais e da descrição da realidade do português do Centro-Oeste Potiguar.

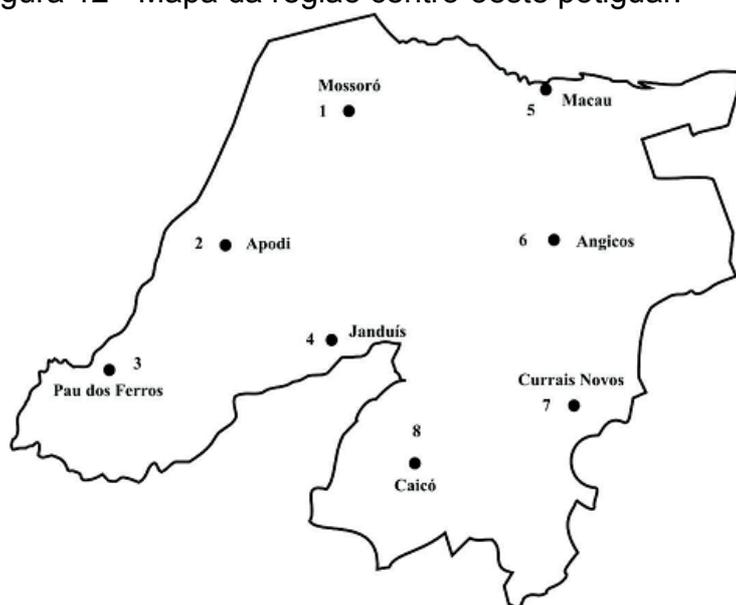
³² Neste subtópico, apresentamos as principais informações expostas por Silva (2012). Para mais informações sobre a pesquisa feita, a tese do autor deve ser consultada.

Considerando os dados socioeconômicos e culturais particulares da região pesquisada, esse autor selecionou oito pontos de inquéritos: quatro da mesorregião do oeste Potiguar (Mossoró, Apodi, Pau dos Ferros e Janduí) e quatro da mesorregião central Potiguar (Macau, Angicos, Currais Novos e Caicó), como mostra a figura abaixo.

O critério dessa escolha foi, primeiramente, o da importância dos aspectos demográficos, históricos, geográficos, políticos, econômicos e culturais e a influência deles sobre os outros municípios da região. Adotou-se, também, o critério da equidistância aproximada. Assim, todas as localidades foram distribuídas de maneira que abrangessem todo o centro-oeste Potiguar, com uma distância entre elas, de, pelo menos, 70 km.

Para a realização da pesquisa de campo, feita por Silva (2012), foram selecionados 32 informantes, levando em conta os seguintes aspectos: a) sexo - para cada ponto, foram entrevistados dois homens e duas mulheres (um homem e uma mulher de cada geração), totalizando 4 informantes por localidade; b) faixa etária - foi distribuída em duas gerações: jovens de 18 a 32 anos (G1) e adultos de 48 a 62 anos (G2). Em cada ponto, foram selecionados dois informantes da G1 e dois da G2; c) escolaridade - foram escolhidos os informantes com escolaridade igual ou inferior ao 9º ano do ensino fundamental.

Figura 12 - Mapa da região centro-oeste potiguar.



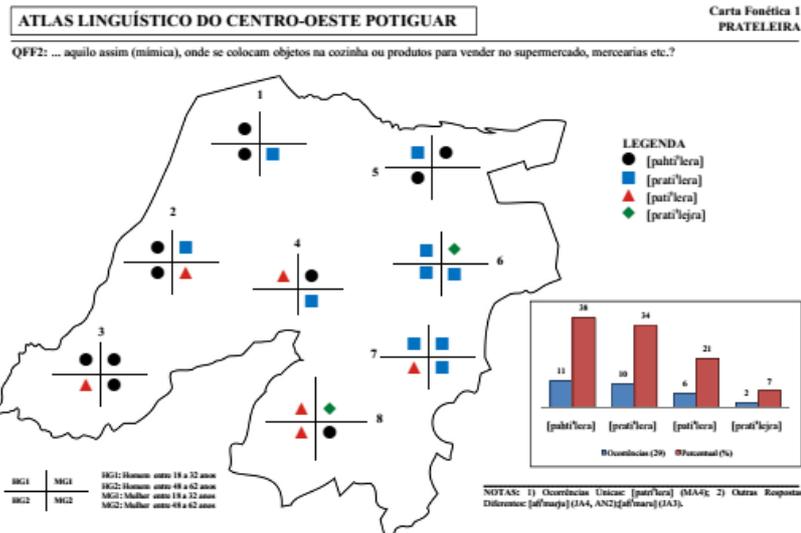
Fonte: Silva (2012, p. 139).

Para todos esses informantes, foram aplicados dois tipos de questionários: um fonético e um semântico e lexical, com o objetivo de coletar dados, possibilitando, assim, a elaboração das 147 cartas linguísticas (84 léxicas e 63 fonéticas) que formam o Atlas em foco e mostram a diversidade do falar da região pesquisada.

A figura 13 mostra o modelo de uma carta fonética do Atlas Linguístico do Centro-Oeste Potiguar e contém, segundo Silva (2012), os seguintes elementos:

- a) encontra-se o nome do Atlas na parte superior à esquerda;
- b) a identificação e o tema da carta e o respectivo número em algarismos arábicos estão à direita do nome;
- c) a indicação, juntamente com o seu número correspondente do tipo de questionário QFF, vem abaixo do título, alinhado à esquerda, seguido pela devida pergunta por extenso;
- d) no centro da carta, situa-se o mapa do centro-oeste potiguar com os pontos de inquérito enumerados de 1 a 8, que correspondem às cidades apresentadas na figura;
- e) em cada ponto, tem uma cruz, formando 4 espaços onde se registram as respostas dos 32 informantes;
- f) abaixo do mapa, alinhada à esquerda, encontra-se a legenda das variáveis sociais, em forma de cruz, referente à distribuição sequencial dos informantes nos pontos pesquisados;
- g) à direita do mapa, encontra-se a legenda com as variantes obtidas a partir da pronúncia do falante e transcritas com base no Alfabeto Fonético Internacional (AFI), com um número reduzido de sinais e diacríticos;
- h) os gráficos mostram a frequência das variantes em cada carta.
- i) por fim, quando for o caso, abaixo dos gráficos são apontadas algumas notas explicativas.

Figura 13 - Carta fonética do Atlas Linguístico do Centro-Oeste Potiguar



Fonte: Silva (2012, p.142).

4.1.2 Atlas Geolinguísticos do Litoral Potiguar

Pereira (2008)³³ diz que o ALiPTG apresenta dados para estudos da variedade linguística na forma da coleta de dados e no tratamento cartográfico de variantes faladas no litoral leste do Rio Grande do Norte. Nele podemos ver a contribuição no registro, na análise e na descrição das variantes linguísticas fonético-fonológicas, léxico-semânticas e morfossintáticas, tanto no ponto de vista diatópico e quanto no diastrático.

Essa autora diz que o estabelecimento da rede de pontos, sugerida por Antenor Nascentes para um Atlas Linguístico, foi feito segundo alguns critérios, tais como: (a) densidade demográfica, (b) a história do município, (c) aspectos geográficos e (d) importância econômica da localidade que representa o universo de pesquisa (ou pontos de pesquisa).

Considerando esses critérios, foi investigada a fala de 24 informantes em cinco municípios, dentre os 11 (onze) que constituem a rede de pontos do Projeto Atlas Linguístico do Rio Grande do Norte, pertencentes a uma área chamada convencionalmente de litoral potiguar. A figura 14 mostra a área que compreende os municípios de onde foram coletadas a pronúncia da fala potiguar e relaciona os

³³ Neste subtópico, apresentamos as principais informações expostas por Pereira (2008). Para maiores informações sobre a pesquisa feita, a tese da autora deve ser consultada.

números a cidades dessa mesorregião. A capital Natal corresponde ao número 1, Canguaretama ao 2, Touros ao 3, Macau ao número 4 e Areia Branca ao número 5.

Figura 14 - Mapa da região do litoral potiguar.

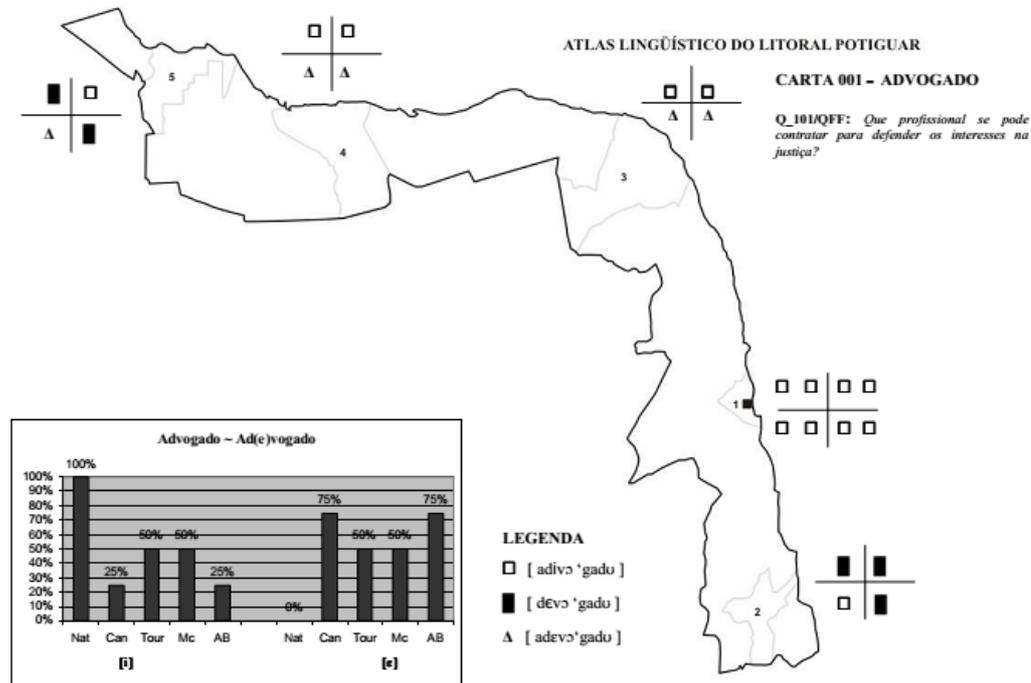


Fonte: Pereira (2008, p. 16).

Para Pereira (2008), os resultados do estudo dessa autora foram obtidos por meio de entrevistas *in loco* e posterior análise dos fatores que condicionam os possíveis traços de variação em níveis fonético, léxico e morfossintático que se pretendiam detectar. Os informantes foram selecionados com base em alguns critérios adotados pelo Projeto ALiB³⁴ e a coleta foi feita através dos questionários utilizados na recolha dos dados para a construção do *corpus*. O questionário apresenta 35 questões para o Questionário Fonético Fonológico (QFF), 35 questões que formou o Questionário Semântico Lexical (QSL) e 10 questões para o Questionário Morfossintático (QMS). Desses questionários, damos ênfase ao primeiro, ou seja, ao Questionário Fonético Fonológico (QFF).

³⁴ Mais informações sobre o projeto ALiB, acesse o sitio: <http://twiki.ufba.br/twiki/bin/view/AlibWebHome>.

Figura 15 - Carta fonética do Atlas Linguístico do litoral potiguar



Fonte: Pereira (2008, p. 20).

A figura 15 mostra o exemplo de uma carta fonética apresentada no ALiPTG e apresenta, segundo Pereira (2008), as seguintes características:

- as cartas linguísticas representam a diversidade dos falares compreendidos no litoral potiguar;
- as cartas fonéticas apresentam-se numeradas de 01 a 35;
- cada número no mapa do litoral representa um ponto de inquérito, como foi apresentado na figura acima;
- os dados postos à esquerda do cruzamento são relativos às respostas dos informantes do sexo masculino e à direita, relativos às respostas dos informantes do sexo feminino;
- a faixa etária é lida em linha vertical do cruzamento;
- acima do mapa, do lado direito, apresentam-se o número da carta, o tipo de questionário, o número e a questão de onde procede cada carta;
- baixo do mapa, à esquerda, encontram-se as legendas, transcritas pelos símbolos do Alfabeto Fonético Internacional (AFI), que correspondem à produtividade fonética pelo informante, localidade e os gráficos com base em índices percentuais de ocorrências;

h) os informantes ímpares pertencem ao gênero masculino e os informantes pares pertencem ao gênero feminino.

Após a escolha e identificação dos fenômenos fonéticos, os quais foram descritos no *tópico 4.1.2* deste trabalho, foi feita a conversão dos símbolos e diacríticos do alfabeto fonético internacional para o SAMPA. Todas as cartas dos questionários fonéticos de ambos os *corpus* estão no apêndice A.

Esse apêndice mostra a transcrição do *corpus* em duas tabelas paralelas: a primeira expõe as cartas fonéticas do Atlas Linguístico do Centro-Oeste Potiguar em ordem crescente e a segunda exhibe as cartas do ALiPTG de forma aleatória. A escolha dessa disposição, em paralelo, ajuda na comparação das transcrições de mesma palavra, mas diverge quanto ao número das cartas. O número na primeira tabela não corresponde necessariamente ao da segunda, todavia, apresenta a transcrição da mesma palavra.

Cada tabela está dividida em quatro colunas: a primeira apresenta o número da carta no *corpus* de origem; a segunda mostra a palavra escrita e as variantes fonéticas da palavra; a terceira exhibe o número de ocorrência para cada variante fonética e a somatória das ocorrências e a quarta explicita em percentuais as ocorrências, como mostra o exemplo na tabela na tabela 4.

Tabela 4 – Apresentação da transcrição fonética pelo SAMPA e as ocorrências nas cartas fonéticas dos atlas

Carta fonética		Ocorrência	%
1	PRATELEIRA		
	[pahti"lera]	11	37,9%
	[prati"lera]	10	34,5%
	[pati"lera]	6	20,7%
	[prati"leira]	2	6,9%
		29	
2	TELEVISÃO		
	[tEIEviza~w]	27	84,4%
	[televiza~w]	5	15,6%
		32	

4.2 Potigrafone - conversor gráfico fônico para a variedade linguística potiguar

No capítulo 2, apresentamos alguns conversores que utilizavam a transcrição fonológica. Um bom exemplo desses sistemas é o Grafone, desenvolvido pelo Instituto de Telecomunicação em Coimbra, Portugal. A transcrição fonológica

"não leva em conta as variações alofônicas" e "está muito próximo da transcrição ortográfica, uma vez que a ortografia do português é formalmente baseada nas representações fonológicas da língua." (SEARA, NUNES e VOLCÃO, 2015, p.113). No entanto, como vimos no tópico 3.2.1, os alofones são as variantes de um determinado fonema, quando sua posição não implica mudança de significado. A palavra <terra> pode ser pronunciada pelas variantes [ˈtɛxa], [ˈtɛra] e [ˈtɛha], porém essa diferença não implica mudança da palavra. Para essas três representações fonéticas, o Grafone apresenta apenas uma representação fonológica feita por meio do símbolo arquifonêmico /R/, /tˈER6/.

O Potigrafone é um conversor gráfico fônico para a variedade linguística potiguar, executando a transcrição fonética ampla das unidades gráficas da língua e considerando algumas variantes fonéticas presentes nessa variedade. O sistema procura fazer a transcrição fonética que expressa a forma linguística mais adequada do registro para a pronúncia potiguar, tomando como fundamento empírico os atlas linguísticos descritos no tópico anterior deste capítulo.

Nos subtópicos a seguir, apresentamos as peculiaridades da transcrição automática desse sistema, como o tipo de transcrição fonéticas, os sinais e diacríticos utilizados para a marcação de tonicidade, nasalização, palatalização e silabação. Mostramos a arquitetura, o algoritmo do sistema e os principais operadores da linguagem de programação Foma e, também, apontamos os principais módulos no sistema.

4.2.1 Transcrição fonética automática

Para Seara, Nunes e Volcão, (2015), a transcrição fonética é a representação dos sons emitidos por falantes de uma língua quando produz a fala. Isso é feito utilizando um alfabeto fonético. Como foi visto no tópico 4.1, Silva (2012) e Pereira (2008) utilizaram os símbolos do Alfabeto Fonético Internacional para fazer o registro da fala potiguar, mas, como foi justificado no capítulo 3 deste trabalho, o Potigrafone registra as variantes dessa fala por meio dos símbolos fonéticos do SAMPA.

As autoras, Seara, Nunes e Volcão, (2015), apresentam dois tipos de transcrição fonética: a transcrição ampla e a restrita.

Na transcrição restrita (detalhada), todos os detalhes fonéticos, mesmo aqueles que podem ser previsíveis pelo contexto, incluindo propriedades secundárias, são consideradas, e na transcrição ampla (aproximada) são explícitas apenas os aspectos mais gerais dos segmentos. (SEARA, NUNES e VOLCÃO, 2015, p.83).

As autoras exemplificam essa distinção com a palavra <quilo>, que pode ser transcrita de forma restrita [kʲilʷɔ] ou de forma ampla [kilɔ]. A primeira transcrição considera certas propriedades secundárias, detectadas em determinados segmentos previsíveis pelo ambiente a ser transcrito, como a palatalização da consoante velar diante de vogal alta anterior e o arredondamento da lateral diante de vogal alta posterior. A segunda não considera nenhum desses detalhes. Quanto mais fiel for à pronúncia, mais recursos fonéticos devem ser usados, exigindo-se maior número de sinais e diacríticos de um alfabeto.

Nesse trabalho, optamos pela transcrição fonética ampla e pela utilização dos sinais e diacríticos disponíveis no SAMPA. Nesse alfabeto, há vários elementos disponíveis para a transcrição da fala humana, todavia, utilizamos apenas quatro deles: o til (~) para a nasalização, a aspa dupla (") para marcação do acento tônico primário, a aspa simples (') para marcar a palatalização e o símbolo de igualdade (=) para a separação silábica, conforme descrição a seguir:

- a) a nasalização vocálica é transcrita por meio do acréscimo de um til (~) após o símbolo vocálico, como em <canto> [ka~tu] e <cama> [ka~ma], sendo que o primeiro marca a vogal nasal e o segundo destaca o fenômeno da nasalidade, como foi mostrado em 4.1.2.2.1;
- b) o acento tônico é marcado por meio do acréscimo da aspa dupla (") antes da sílaba tônica, como em <canto> [ˈka~tu];
- c) a palatalização é representada por meio do acréscimo de aspa simples ('), após consoante palatalizada como em <óleos> [ˈOl'us].
- d) a silabação é marcada por meio do símbolo de igualdade (=) posto entre as sílabas, como em <car=ro>. Este sinal não aparece na transcrição fonética, pois é um símbolo reservado ao beta-gerador (ver o tópico 6.3).

Segundo Wells (2014), a transcrição fonética feita com a utilização do SAMPA considera que as letras minúsculas do AFI possuem a mesma representação e, quando não há correspondência, usa-se a contrabarra (\) como modificador para criar um novo símbolo, por exemplo, o símbolo [|] representam

um som distinto do símbolo [\]; e quando não tem qualquer relação, os diacríticos do SAMPA seguem os símbolos que os modificam.

Além do uso dos sinais e diacríticos, a transcrição fonética contempla os processos fonológicos descritos no subtópico 3.2.2.2. A palavra <caixa>, por exemplo, apresenta duas formas fonéticas: uma transcrição com o fenômeno da monotongação, [ˈkaSa], e outra sem esse fenômeno, [ˈkajSa]. Essa opção está respaldada no fato de que, na fala potiguar, ocorrem essas duas variantes.

Essa forma de transcrição para essa palavra e para outros casos se fundamenta nas cartas fonéticas do Atlas Linguístico do Centro-Oeste Potiguar e do ALIPTG.

Assim, o Potigrafone faz a transcrição fonética da seguinte forma: recebe como entrada (*input*) uma palavra escrita graficamente, conforme a ortografia oficial e retorna como saída (*output*) uma ou mais formas fonéticas da fala potiguar, ou seja, um *input* pode gerar mais de um ou mais *outputs*, como exemplificamos no próximo tópico.

Todavia, nem todos os fenômenos presentes nas cartas fonéticas desses atlas foram considerados na implementação do sistema, apenas aqueles que estão descritos no tópico 3.2 deste trabalho. Por exemplo, a carta fonética da palavra <prateleira> mostra quatro transcrições diferentes. Nas cinquenta e seis entrevistas, ocorreram 37,5% para a pronúncia [pahtiˈlera], 26,7% para a pronúncia [pratiˈlera], 23,3% para a pronúncia [patiˈlera] e 5,4% para a pronúncia [pratiˈlejra]. Esse tipo de fenômeno é o apagamento da vogal nas palavras proparoxítonas, como na palavra <fósforo> que é pronunciada por [ˈfOsfru], o sistema não apresenta a transcrição, mas exibe apenas duas saídas e uma delas, obrigatoriamente, representa a forma mais próxima da entrada, ou seja, a transcrição [pratiˈlejra], que foi a menos frequente nos *corpora*. No entanto, é a mais alinhada com a entrada da palavra.

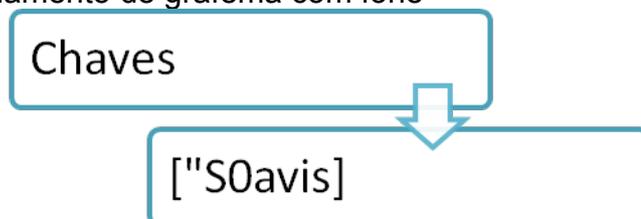
Além desse fato relativo à escolha das transcrições, há outros fatos importantes que merecem destaques. Alguns grafemas do português possuem correspondências unívocas com os sons da variedade linguística potiguar, situação na qual o sistema faz a conversão de modo direto. É o caso das letras <p, b, t, d, f, v> que são convertidos pelos símbolos [p, b, t, d, f, v], respectivamente.

Outros grafemas apresentam correspondências dependentes de fatores contextuais e posicionais. O grafema <g> se realiza como uma fricativa alvéolo palatal sonora [ʒ] ou como um oclusiva velar sonora [g], dependendo do contexto. A

saída do grafema <r> é realizada como uma fricativa glotal surda [h] ou como uma vibrante simples [r], dependendo da posição silábica em que a letra se encontra; quando está na posição de coda medial ou no ataque inicial de palavra, é uma consoante glotal; quando está em coda final ocorre o apagamento, sendo, portanto, transcrita como símbolo nulo; quando é a segunda consoante do ataque complexo ou quando está entre vogais é uma consoante vibrante.

Somam-se a esses o fato de que os dígrafos consonantais são um conjunto de grafemas para representar apenas um som da língua. No Potigrafone, esse fenômeno gráfico está diretamente relacionado ao alinhamento de grafema para fone. Para um sistema dessa natureza, Beesley e Karttunen (2002) dizem que o alinhamento é o primeiro passo que se deve fazer. Um símbolo gráfico da forma lexical é associado a um símbolo da superfície de um para um. Nesse caso, dígrafos têm representações diferentes das letras, pois é acrescentado um símbolo nulo [0] que corresponde à letra diacrítica para as consoantes, ver o tópico 3.1.1. Para exemplificar essa forma de transcrição, mostramos a figura 16 que ilustra a transcrição alinhada do dígrafo <ch>, mas que essa forma também ocorre com os outros dígrafos.

Figura 16- Alinhamento de grafema com fone



Fonte: Elaborado pelo autor.

A transcrição fonética de alguns dígrafos depende do contexto de escrita e da construção silábica, mais precisamente, da fronteira silábica. Nas palavras <crescer> e <casca>, as letras <sc> formam um dígrafo e um encontro consonantal, respectivamente, sendo que a transcrição fonética da letra <c>, na primeira palavra, tem como valor fonético o símbolo nulo [0], pois essa letra é diacrítica - veja o tópico 3.1.1 - e, na segunda palavra, a letra <c> tem valor fonético de oclusiva velar surda [k], ou seja, essa letra representa uma consoante em posição de ataque simples. Essa consoante pode ser produzida entre vogais, no início de palavra e/ou em posição de coda, mas o grafema <sc> só ocorre entre vogais, o que “condiciona” o

uso restrito da forma fonética. Observe que essas letras se apresentam no mesmo contexto de escrita - entre vogais, mas possuem valores fonéticos diferentes. A distinção de uma ou de outra forma fonética depende da construção de um separador silábico que delimite o início e o fim das sílabas. Por esses e outros casos, o Potigrafone separa as sílabas da forma gráfica, para que uma regra não entre em conflito com outras.

A transcrição automática do Potigrafone considera essas peculiaridades da escrita para a fala, levando em conta as características da forma escrita, da forma falada e do alinhamento dessas formas.

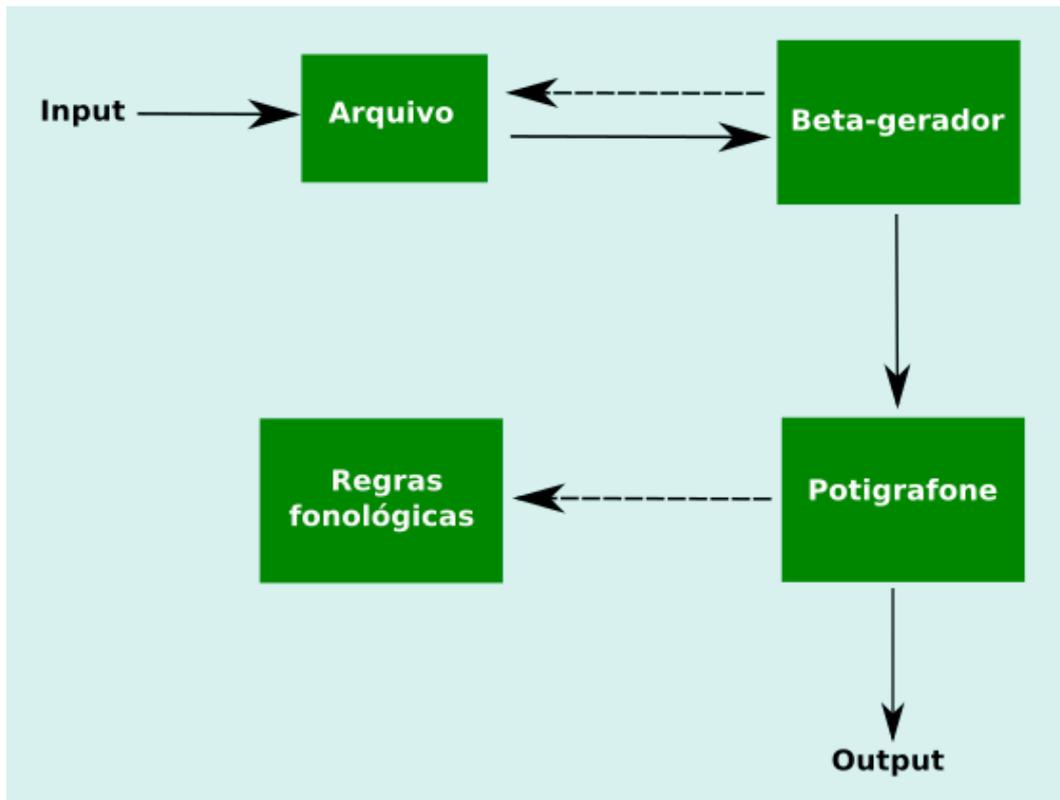
4.2.2 Arquitetura e algoritmo do sistema

O Potigrafone faz a transcrição automática utilizando uma abordagem de uma máquina de estados finitos, ou seja, esse sistema é um transdutor de estados finitos. Os transdutores são constituídos por duas "fitas": uma de entrada e uma de saída, traduzindo o conteúdo da entrada para o conteúdo da saída.

O processo de desenvolvimento desse sistema está descrito no tópico 4.2.3 e os módulos estão dispostos no apêndice B, os quais mostram a forma de nomeação e a estrutura interna dos arquivos. A arquitetura utilizada para representar a estrutura do transdutor se baseia em quatro módulos fundamentais: o arquivo, o beta-gerador, o das regras fonológicas e o Potigrafone. Esses elementos estão distribuídos em arquivos que contêm a implementação dos módulos do programa.

A figura 17 ilustra a arquitetura do sistema. O módulo *arquivo* corresponde ao subsistema no qual os elementos do alfabeto ortográfico são agrupados em variáveis. Nesse módulo, os símbolos de entrada foram enumerados dentro das variáveis de modo que o Potigrafone reconhece somente os caracteres pertencentes à língua portuguesa e, conseqüentemente, exclui os outros símbolos que não pertencem, como aqueles que foram mostrados nos critérios de entrada do sistema, no tópico 4.1. Além disso, esse módulo apresenta a expressão regular para a formação da estrutura silábica da língua portuguesa. Logo, esse arquivo contém os símbolos terminais para a concatenação das formas gráficas, e os símbolos não terminais, os quais agrupam as consoantes e as vogais para a constituição da sílaba. O arquivo desse módulo foi nomeado com o título "Alfabeto.foma" e está disponível no apêndice B.

Figura 17 - Arquitetura do sistema Potigrafone



Fonte: Elaborado pelo próprio autor.

O módulo *beta-gerador* verifica se todos os símbolos da cadeia de entrada pertencem à língua, ou seja, se são elementos que pertencem ao módulo do alfabeto. Se ao menos um deles não pertencer, esse módulo retorna o *output* <?+>, indicando que a palavra não faz parte da língua. Se a cadeia pertence a essa língua, então, o módulo reconhece a padronização silábica dos caracteres e gera as palavras, separando-as em sílabas bem formadas para o português. O *output* desse módulo é a concatenação dos símbolos gráficos com o símbolo de igualdade (=) que delimita as sílabas. Por exemplo, se entrar com a palavra <carro>, o módulo reconhece os símbolos e a retorna com a separação silábica <car=ro>. Nesse caso, o beta-gerador é um protossilabificador para essa língua, gera as palavras com a separação silábica. Para além dos motivos mostrados no tópico 3.3, optamos por esse método de construção do sistema, pois reduz o número de regras de transformação. O arquivo desse módulo foi nomeado com o título “Beta-gerador.Foma” e está disponível no apêndice B.

O módulo das *regras fonológicas* está disposto em um conjunto de arquivos separados em três tipos de arquivos relacionados à formalização das três

principais regras fonológicas: transformação de segmentos, inserção de segmento e cancelamento de segmento, como foi apresentado no tópico 3.2.2.1 e descrito a seguir:

- a) o arquivo que tem por título “Tonicidade.foma” contém todas as regras de inserção do símbolo tônico (“), para a marcação da sílaba tônica da palavra;
- b) os arquivos que têm por títulos “Consoantes.foma” e “Nucleo.foma” apresentam todas as regras de transformação das consoantes e do núcleo silábico para a língua portuguesa;
- c) o arquivo que tem por título “Fenomenos.foma” mostra as regras utilizadas para os fenômenos de ditongação, de epêntese, de nasalidade, de cancelamento, de monotongação, de apagamento, de vozeamento, de harmonia vocálica, de lenição, de palatalização e de neutralização.

Esse módulo é acessado pelo Potigrafone para a aplicação das regras na palavra gerada pelo módulo do beta-gerador.

O *Potigrafone* acessa o *output* do beta-gerador e aplica as regras fonológicas, relacionando os símbolos de entrada com os símbolos fonéticos do SAMPA e retorna a transcrição fonética para a variante linguística potiguar. Esse sistema é executado numa interface gráfica na linguagem html 5 e será disponibilizado na web de modo que o usuário poderá utilizá-lo com muita facilidade.

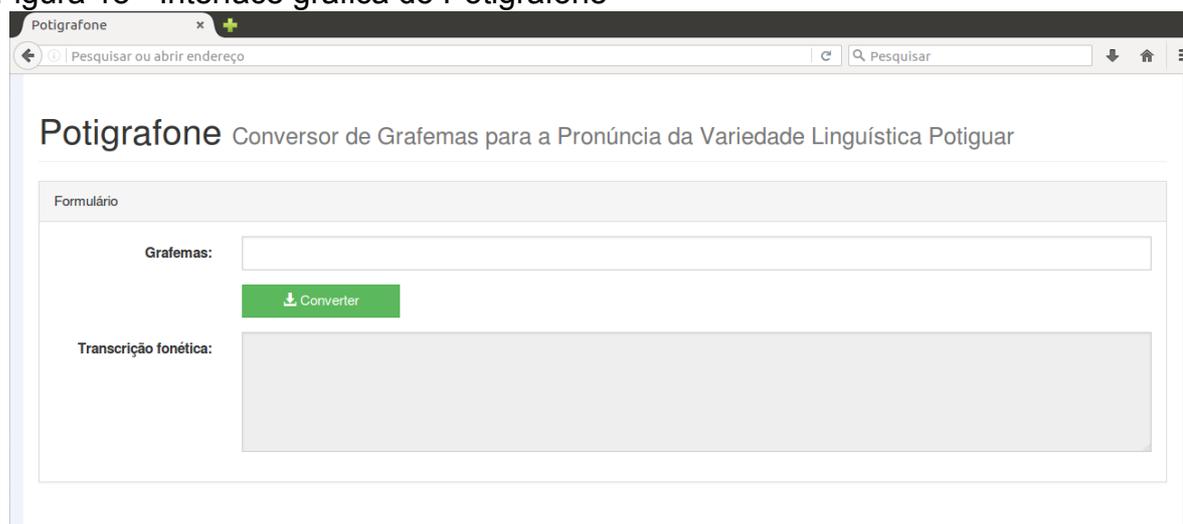
Como mostra a figura 18, o uso do sistema é muito intuitivo, tem apenas o campo de preenchimento com o grafema do português e o campo de saída onde apresenta a transcrição fonética. Porém, para melhor compreensão do uso desses módulos no sistema, utilizaremos o exemplo da palavra <caixa>, apresentado no tópico 3.4.2. Ressaltamos que o uso dos colchetes angulares indica que a palavra é um conjunto de grafemas da língua portuguesa.

O *input* <caixa>, escrito no campo da entrada de dados “grafemas”, é relacionado aos elementos do alfabeto da língua portuguesa existentes no arquivo de entrada e o sistema reconhece cada letra da palavra como pertencente ao alfabeto da língua formal. Esse processo é percebido no arquivo Alfabeto.foma. Se um dos símbolos de entrada não pertence ao símbolo de entrada, o sistema responde ao usuário com o conjunto de caracteres “???” , indicando que a palavra é estranha à linguagem.

Após o reconhecimento dos caracteres, os dados são acessados pelo

módulo *beta-gerador*, o qual executa a separação silábica da palavra, inserindo o símbolo de igualdade. Como vimos acima, a separação silábica é marcada por meio desse símbolo no SAMPA. Esse sinal não aparece na transcrição fonética, todavia, o potigrafone utiliza-o para a aplicação das regras, ou seja, ele serve como elemento de fronteira silábica. Assim, a transcrição da palavra acima por meio do beta-gerador fica dividida por meio de duas sílabas, <cai=xa>.

Figura 18 - Interface gráfica do Potigrafone



Fonte: Elaborada pelo próprio autor.

Feita a divisão silábica, o Potigrafone acessa o conjunto de regras fonológicas e aplica as regras de acentuação tônica e as regras de correspondência ao caractere de entrada. Para essa situação, o sistema executa a regra das palavras paroxítonas descritas no módulo das REGRAS DE TONICIDADE DAS PALAVRAS – o apêndice B deste trabalho – e atribui à sílaba tônica o diacrítico do SAMPA que marca a sílaba tônica, a aspa dupla “ ”. Com a marca da tonicidade, a palavra <cai=xa> fica assim descrita: <"cai=xa>.

Feita a marcação tônica, o sistema passa à execução das regras de transformação e dos fenômenos fonéticos. A letra <c> é transformada no som oclusivo [k], pois o conversor atribui a essa consoante a regra da consoante oclusiva

```
c -> k || _[vpos | vmed | liq | "="];
```

onde se lê: a letra <c> será transformada em oclusiva velar desvozeada [k], quando estiver antecedida das vogais médias e posteriores ou das consoantes líquidas ou do sinal de separação silábica. Desse modo, essa regra é válida para esse contexto

e também para outros, como nas palavras <co=po>, <cri=a>, <cac=to>.

Depois da transformação da oclusiva, o conversor aplica a regra do núcleo silábico ao par de vogais <ai>. Nesse contexto, a letra <a> não recebe nenhuma transformação, porque ela é núcleo silábico do ditongo, mas a letra <i> é transformada em um glide [j], como explicita a seguinte regra

`i -> j || [vogal]_ ;`

onde se lê: a letra <i> será transformada em glide [j], quando estiver seguida de uma vogal. Logo, a transcrição para esse ditongo é a vogal baixa [a] seguida de glide [j], [aj].

Após a transformação da vogal alta anterior em glide, o Potigrafone faz a transformação do grafema <x> em uma fricativa alveolar desvozeada [S], aplicando a seguinte regra

`x -> S || [.#. | .#. "" | [vogal vogal] "="]_ ;`

onde se lê: a letra <x> será transformada em consoante fricativa, quando estiver antecederida por uma fronteira de palavra, ou seja, no início de uma palavra, ou pelo início de uma palavra seguida de um sinal de tonicidade, ou pelo ditongo seguido de um separador silábico. Assim, essa regra aplica-se a outros contextos para esse grafema, como na palavra <xícara>.

Por último, o sistema executa a transformação do grafema <a> em uma vogal média átona, apresentada pelo símbolo <6> do SAMPA. Para isso, o Potigrafone executa a regra seguinte que mostra a redução vocálica para essa vogal

`a->6 || _(s) .#. ;`

onde se lê: a vogal <a> será reduzida, quando estiver numa sílaba pós-tônica final antecederida ou não do grafema <s>.

Apenas com a utilização dessas regras, a saída do sistema ficaria da seguinte forma: ["kaj=S6]. Nesse caso, teríamos uma saída para a transcrição da palavra <caixa> e com a separação silábica. No entanto, na fala potiguar, temos outra variante para essa palavra, a pronúncia com a forma monotongada. Para a aplicação do fenômeno da monotongação, o Potigrafone executa a seguinte regra

`j->[j | 0] || _ "=" ;`

onde se lê: a semivogal [j] receberá duas saídas: a semivogal [j] e a ausência de representação sonora [0], quando estiver precedida pelo sinal de separação silábica. Desse modo, a saída do sistema ficaria da seguinte forma: ["kaj=S6], ["ka=S6].

Feito isso, o Potigrafone executa a última regra que é a de cancelamento

do sinal de separação silábica,

```
["="] -> 0;
```

onde se lê: os sinais de separação silábica devem ser cancelados. Assim, finaliza o processo de transcrição fonética automática da palavra <caixa>, retornando ao usuário duas formas fonéticas da variedade linguística potiguar: ["kajS6], ["kaS6].

Além da arquitetura do sistema, apresentamos a seguir o algoritmo do Potigrafone que simula a execução do transdutor Potigrafone. Este algoritmo foi construído por França (2016) em português e se encontra disponível no Anexo E deste trabalho.

Variáveis

```
elemento      : caractere;
simbolos      : listaCaractere;
silaba        : listaCaractere;
palavra       : listaCaractere;
transcricao   : listaCaractere;
```

Funções

```
função existeElemento(elemento: caractere): inteiro;
função verificaSimbolos(): inteiro;
função transcrevePalavra(palavra :listaCaractere ): listaCaractere;
função regraTranscricao(sil :listaCaractere ): listaCaractere;
função principal;
```

Função existeElemento - verifica se os símbolos pertencem ao conjunto de símbolos de entrada

```
função existeElemento(elemento: caractere): inteiro
var
    aux      :inteiro;
    posicao   :inteiro;
inicio
    aux = 1;
    posicao = 0;
```

```

    enquanto posicao < tamanho(simbolos) faça
    inicio
        if( elemento==simbolos[posicao] ) então
        inicio
            aux = 1;
        fim if
        posicao++;
    fim enquanto
retorna aux;
fimfunção

```

Função verificaSimbolos - verifica se todos os caracteres da palavra são válidos

```

função verificaSimbolos(): inteiro
var
    aux :inteiro;
    posicao :inteiro;
inicio
    aux = 1;
    posicao = 0;
    enquanto posicao < tamanho(palavra) faça
    inicio
        if( existeElemento(palavra[i])==0 ) então
        inicio
            aux = 0;
            retorna aux;
        fim if
    fim enquanto
retorna aux;
fimfunção

```

Função transcrevePalavra - Implementa os passos da transcrição

```

função transcrevePalavra(palavra :listaCaractere ): listaCaractere
var

```

```

        auxTranscricao      : listaCaractere;
        silaba              : conjunto de listaCaractere;
        posicao :inteiro;
    inicio
        silaba = divisaoSilabica (palavra);
        posicao = 0;
        enquanto posicao < numeroElementos(silaba) faça
            inicio
                auxTranscricao      =      auxTranscricao      +
regraTranscricao(silaba[posicao]);
            fim enquanto

        retorna auxTranscricao;
    fimfunção

```

Função regraTranscricao - Implementa os passos da transcrição fonética

```

função regraTranscricao(sil :listaCaractere ): listaCaractere
var
    auxTranscricao      : listaCaractere;
inicio
    /* implementações das regras de transcrição p/diversos casos */
    retorna auxTranscricao;
fimfunção

```

Função principal

```

principal()
inicio
    escreva("Digite a palavra");
    leia (palavra);
    if (verificaSimbolos(palavra) ) então
        inicio
            transcricao = transcrevePalavra();
        fim

```

```

senão
início
    imprime "palavras com caracteres inválidos"
fim
    imprime transcriçao;
fim

```

Nesse tópico, apresentamos a arquitetura e o algoritmo do sistema. No próximo tópico, apresentamos os comandos e operadores utilizados nas expressões regulares da linguagem de programação Foma e na implementação dos módulos do sistema.

4.2.3 *Implementação do sistema pela linguagem Foma*³⁵

O tópico anterior destacou a arquitetura e o algoritmo do Potigrafone, mostrando a importância das funções de cada módulo. Este tópico considera a linguagem de programação: os comandos e os operadores básicos, e o módulo da constituição da formação silábica.

4.2.3.1 *Foma: comandos e operadores*

Segundo Hulden (2009), o *Foma*³⁶ é um compilador, uma linguagem de programação e uma biblioteca em C para a construção de autômatos e transdutores de estados finitos para muitas aplicações de processamento de linguagem natural, como a produção de analisadores morfológicos e fonológicos. A interface do *Foma* é semelhante à interface do *Xerox Finite State Transducer* (XFST) e suporta a maioria dos comandos e da sintaxe de expressão regular em XFST.

O *Foma* é uma linguagem que não faz distinção precisa das operações dos *autômatos de estados finitos* e dos *transdutores de estados finitos*. Suponhamos que, usando alguns dos operadores complexos para definir a língua que modela a regra de ortografia, queiramos transforma um grafema <i> em um <e>, (exceto

³⁵ O código-fonte do Potigrafone está disponível no **Apêndice C** e a forma de execução em linhas de comando pelo terminal do Foma está no **Apêndice D**.

³⁶ Parte desse subtópico é uma tradução e uma adaptação do apêndice da tese de Hulden (2009, p. 272 - 278) o qual está na referência.

depois da letra <c>) pode ser formalizado pela seguinte expressão.

```
Foma[1]: regex ~ $[\c e i|c i e];
```

Nesse caso, tem-se um autômato que aceita todas as palavras que segue a regra e rejeita todos os elementos que não o fazem. Se, agora, aplicar uma palavra para esse autômato, este retorna a palavra, se estiver de acordo com a regra. Ex.:

```
Foma[1]: down friend
friend
```

Isso ocorre porque, nesse caso, o autômato atua como um transdutor de estado finito, em que cada etiqueta de entrada é idêntica à saída do rótulo. Se aplicar uma palavra estranha à linguagem, o *Foma* apresenta erro. Ex.:

```
Foma[1]: down weird
???
```

A única marca distintiva entre um autômato e um transdutor é o número de *aridade*³⁷ da rede, exibido em *Foma* por meio do comando *print net*: se o número for 1, o sistema é um autômato, se apresentar 2 o sistema é um transdutor.

Outra característica do *Foma* é o fato de que, por padrão, essa linguagem minimiza os autômatos construídos. A minimização é a redução do número de estados de um autômato, reduzindo o tempo de resposta para uma palavra. Os transdutores de estados finitos possuem essa mesma característica, pois os pares de símbolos são considerados símbolos únicos.

Para descrever um transdutor simples que muda todos os símbolos **a** para **b**, e os símbolos **b** para **a**, deixando todo o resto inalterado, digita-se:

```
regex [?-a-b | a:b | b:a]*;
```

O *Foma* tem, também, suporte extensivo para a construção de transdutores complexos que reescrevem sequências de *strings* em outras cadeias e,

³⁷ Na matemática, a **aridade** de uma função ou operação é o número de argumentos ou operandos tomados. A aridade de uma relação é o número *n* de elementos que compõem as *n*-uplas ordenadas pertencentes à relação. (disponível em: <https://pt.wikipedia.org/wiki/Aridade>).

o exemplo acima pode ser construído com mais facilidade usando o operador de reescrita. Ex.:

```
regex a -> b , b -> a;
```

Um outro aspecto do *Foma* se concentra no fato de ela usar três símbolos especiais sobre as transições das redes de estados: Epsilon (0), Identidade (@) e Unknown (?). No entanto, apenas dois deles têm um significado especial em expressões regulares: o primeiro e o último. Essa discrepância entre símbolos especiais nos rótulos de arco e das expressões regulares é um artefato do caminho *Foma*. Do ponto de vista das expressões regulares, a semântica é simples: 0 é a *string* vazia e o ? significa qualquer símbolo. Quanto aos outros dois símbolos especiais, @ e ?, são descritos da seguinte forma: o símbolo @ é interpretado como a relação de identidade de qualquer símbolo, exceto o alfabeto da rede, ou seja, a:a, b:b, etc, onde se assume que *a* e *b* não fazem parte do alfabeto. O símbolo especial ? visa igualmente a qualquer símbolo, não no alfabeto da rede, no entanto, que ocorre apenas em um dos lados de uma etiqueta, por exemplo, a:?. O símbolo de a:? é traduzido por qualquer *a* símbolo, não no alfabeto.

Além desses aspectos peculiares à linguagem, destacamos, também, os *comandos* e *operadores* utilizados nas expressões regulares. Os comandos são palavra da linguagem que contêm instruções para determinada função. Apresentamos, a seguir, os comandos de definição de variáveis e compilação de uma expressão regular, os comandos de entrada utilizados no recebimento dos dados digitados pelo usuário e os comandos de saída de dados utilizados para mostrar informações sobre o transdutor.

O comando *regex* serve para a entrada das expressões regulares básicas e para compilá-las em uma máquina de estados finitos. Digitando-se o comando abaixo

```
Foma[0]: regex [a | b]*;
```

o *Foma* denota que os elementos contidos entre colchetes pertencem a uma língua formal e a compila numa rede de estados finitos, dando como resposta informações da rede, como quantidade de *bits* **271 bytes**, o número de estados **1 state** e de arcos, **2 arcs**, e o tipo da rede estados finitos, **Cyclic**.

Para se obter mais informações sobre essa *regex*, utiliza-se o comando

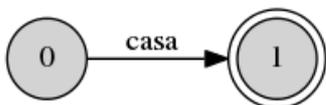
print net (ou simplesmente **net**) que imprime todas as informações sobre a *Finite-State Machine* (FSM). Ex.:

```
Foma[1]: net.
Sigma: a b
Size: 2.
Net: 66334873
Flags: deterministic pruned minimized epsilon_free
Arity: 1
Sfs0: a -> fs0, b -> fs0.
```

Essa FSM pode ser visualizada por meio do comando **view net** que gera o grafo de transição para a expressão regular que está na pilha.

Acrescentamos que, por padrão, o *Foma* une as cadeias de caracteres (*strings*) dentro de uma expressão regular como se fossem apenas um carácter. Isso significa que uma expressão regular, como *Foma[0]: regex casa*; gera a rede apresentada no gráfico 6.

Gráfico 6 - Grafo de transição de palavra não concatenada

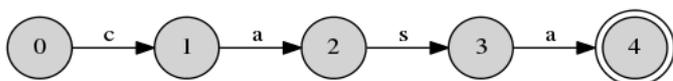


Fonte: Elaborado pelo próprio autor.

Provavelmente não é o que se pretende, uma vez que a cadeia <casa> é tratada como um único **símbolo**. Se o que se desejava era obter a concatenação dos quatro caracteres <c>, <a>, <s> e <a>, uma maneira de fazer isso seria utilizando o espaço em branco ou as chaves. Ex.:

```
Foma[0]: regex c a s a;
Foma[0]: regex {casa};
```

Gráfico 7 - Grafo de transição de palavra concatenada



Fonte: Elaborado pelo próprio autor.

Como as expressões regulares são inseridas com o comando *regex*, o *Foma* armazena os dados em uma pilha interna. O número na linha refere-se ao número de redes armazenadas na pilha. Muitas operações de autômatos são

aplicadas, por padrão, para a última rede definida, ou seja, a rede que estiver em cima da pilha.

Para verificar se um transdutor está executando as informações codificadas, usamos os comandos *down* e *upper*. Esses comandos são de entradas de dados pelo usuário. O primeiro comando exige que o usuário entre com palavras da forma lexical, a forma escrita da palavra (forma subjacente) e retorna a forma fonética, a forma transcrita (forma de superfície). Ex.:

```
Foma[1]: down casa
["kaz6]
```

O segundo comando executa a função inversa do comando anterior. Resumindo: se a intenção é testar a produção de várias palavras e da FSM, digitamos os comandos *upper* ou *down* para introduzir uma subcamada onde as palavras podem ser digitadas uma após a outra, aplicando-as para o topo da rede na pilha, quer na forma descendente, quer na forma ascendente, como no exemplo abaixo.

```
Foma[1]: up ["kaz6]
casa
```

Além disso, as expressões regulares podem ser compiladas em redes de estados finitos por meio de *definições de variáveis* e, então, marcadas para reutilizá-las em expressões posteriores. Por exemplo, definem-se duas expressões regulares:

```
Foma[0]: define ContainsA ?*A ?*;
defined ContainsA: 2 states, 4 arcs, Cyclic.
Foma[0]: define ContainsB ?*B ?*;
defined ContainsB: 2 states, 4 arcs, Cyclic.
```

Em seguida, elas podem ser utilizadas em outras definições ou expressões regulares, como por exemplo: *Foma[0]: regex ContainsA & ContainsB;* Obtemos, então, a língua que contém pelo menos um A e pelo menos um B.

A seguir, apresentamos os principais operadores dessa linguagem.

Os operadores são sinais que armazenam valores para os elementos,

tanto para a entrada, quanto para a saída. Os operadores utilizados são os de concatenação, união, intersecção, complementação, *Kleene plus* e *Kleene star*, para o desenvolvimento de um transdutor nessa linguagem. Os símbolos utilizados no quadro 12 são apresentados com mais detalhes por Hulden (2009).

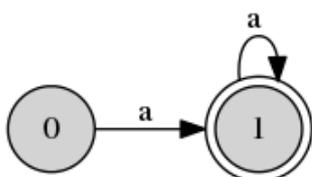
Quadro 12 - Principais símbolos utilizados no Foma

Nº	Operador	Descrição
1	A^+	Uma ou mais vezes
2	A^*	Zero ou mais vezes
3	AB	Concatenação (espaço em branco)
4	(A)	Opcionalidade
5	$A B$	União
6	$A . o . B$	Composição

Fonte: Hulden (2009).

O operador *Kleene plus*, item 1 da tabela, aplica a concatenação de um ou mais símbolos de uma língua L , ou seja, uma ou mais concatenação de uma ou mais string de L . Esse operador armazena qualquer elemento do alfabeto, excluindo o zero. Suponhamos o seguinte: conjunto $A = \{a\}$. Então, se $a \in L$ e L é definida pela expressão regular L^+ , então, as palavras pertencentes à L são formadas por a uma ou mais vezes. A FSM deixa claro que as palavras dessa língua são compostas de pelo menos um elemento a .

Gráfico 8 - O uso do operador *Kleene plus* numa língua formal



Fonte: Elaborado pelo próprio autor.

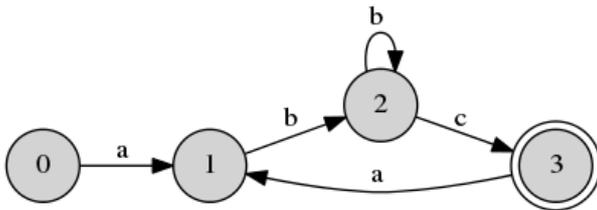
O operador *Kleene star*, item 2, aplica a concatenação de zero ou mais símbolos da língua L . Esse operador pega qualquer elemento do alfabeto, incluindo o zero. Suponhamos o seguinte: conjunto $A = \{\epsilon, a\}$. Então, se $a \in L$ e L é definida pela expressão L^* , logo, a é uma palavra pertencente à L , é formada por zero ou mais elementos de a . Veja o gráfico abaixo da FSM de L^* .

Gráfico 9 - O uso do operador *Kleene star* numa língua formal

Fonte: Elaborado pelo próprio autor.

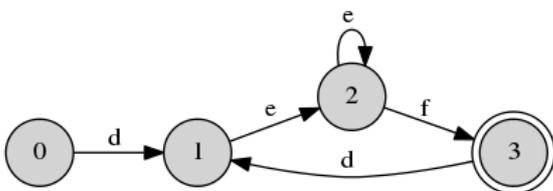
A concatenação de duas FSM A e B, item 3, escrita como A B, tem uma seta ligando cada estado final do primeiro FSM para o estado inicial da segunda. O estado inicial ou final em AB depende de onde A ou B aceita a *string* vazia ϵ . Por exemplo, suponhamos que se queira reconhecer a língua AB onde $A = \{a, b, c\}$ e $B = \{d, e, f\}$ mostrados pelas FSMs descritas abaixo separadamente pelos gráficos 10 e 11.

Gráfico 10 - Concatenação do conjunto A



Fonte: Elaborado pelo próprio autor.

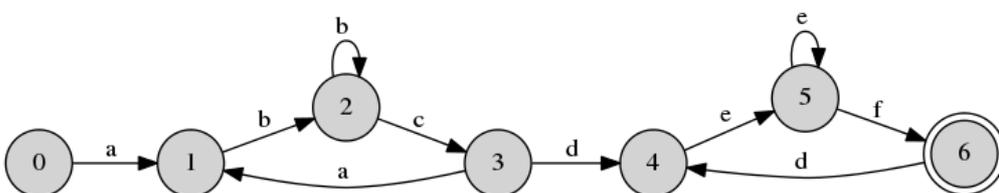
Gráfico 11 - Concatenação do conjunto B



Fonte: Elaborado pelo próprio autor.

Ao ligar o estado final de A com o estado inicial de B, temos a concatenação de A e B.

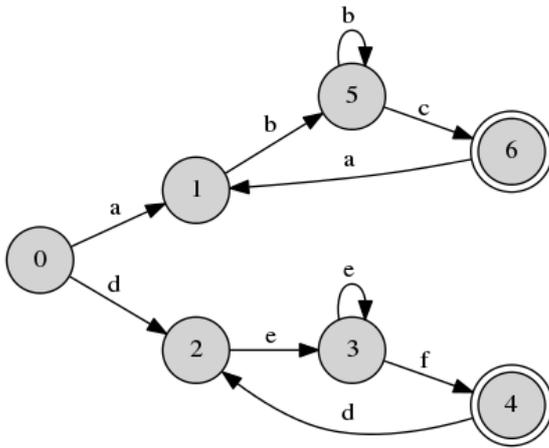
Gráfico 12 - Concatenação dos transdutores A e B



Fonte: Elaborado pelo próprio autor.

A união (ou disjunção) de uma FSM é similar à concatenação. A diferença é que, ao invés de executar na sequência, a FSM é executada em paralelo. A união de duas FSM A e B, escrito $A \cup B$, aceita a *string* s se, somente se, ou A ou B, ou ambos, aceita(m) s . A união de A e B é expressa diagramaticamente da seguinte maneira:

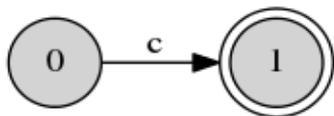
Gráfico 13- União de transdutores $A \cup B$



Fonte: Elaborado pelo próprio autor.

A intersecção de duas FSM A e B, escrita como $A \cap B$, aceita uma cadeia s se, e somente se, s for elemento que pertença a A e a B. Por exemplo, consideremos a intersecção entre dois transdutores que reconhecem as línguas formais do conjunto $A = \{a, b, c\}$; e $B = \{c, d, e\}$. Dois estados são compatíveis se, e somente se, os conjuntos de etiquetas têm uma intersecção não vazia. A intersecção desses autômatos é formada tomando todos os pares de estados que são compatíveis e ligando-os com arcos sempre que ambas as projeções dos pares estão ligadas. A intersecção dos autômatos acima é mostrada abaixo, pois o elemento c pertence ao conjunto A e ao conjunto B.

Gráfico 14 - Intersecção de transdutores $A \cap B$



Fonte: Elaborado pelo próprio autor.

O Foma é uma linguagem de programação destinada ao processamento

de linguagem natural, utilizando a tecnologia de estados finitos. Ela constrói os transdutores e autômatos de forma integrada por meio de expressões regulares, que exigem a utilização de comandos e operadores fundamentais aos desenvolvimento dos sistemas. O próximo subtópico apresenta a aplicação desses comandos e operadores no desenvolvimento dos elementos silábicos do Potigrafone.

4.2.3.2 Módulos do sistema

O produto final da compilação de uma gramática linguística é relativamente simples, quando é desenvolvido por meio de um transdutor de estado finito. Todavia, há vários estágios intermediários envolvidos, compreensíveis apenas quando se observam os processos pelos quais foram constituídos. Desse modo, detalhamos a construção dos módulos, a partir dos comandos e operadores do Foma. Estes módulos estão disponíveis no apêndice B.

Como foi destacado no tópico 6.3, a construção do transdutor foi feita em arquivos intitulados com a função dos códigos do sistema, ficando assim distribuídos: i) um destinado aos símbolos terminais e não terminais; ii) um contendo o silabificador e iii) cinco pertencem às regras fonológicas e um para o transdutor. Cada arquivo foi nomeado de modo que o título identifica a função dos códigos no sistema. Além disso, a estrutura interna do código reflete os seguintes parâmetros: (1) um cabeçalho contendo o comando *source* - responsável por executar um código-fonte em Foma; (2) as definições das variáveis por meio do comando *define* e (3) a construção da rede de estados finitos por meio do comando *regex*.

O arquivo *Alfabeto.foma* contém as funções declaradas das vogais e das consoantes e da estrutura silábica. As funções de expressões regulares podem ser definidas de forma semelhante. O formato para as funções declaradas é feito da seguinte forma: *define NomeDaFunção ExpressãoRegular*. O comando *define* é usado na definição de uma variável que contém um conjunto de elementos da língua formal. O uso desse comando é utilizado abaixo para definir as vogais e as consoantes da língua. Ex.:

```
Foma[1]: define Vogais [a | e | i | o | u];
```

Em posição de ataque simples, há dezenove consoantes. No entanto,

esse submódulo terá apenas 18 grafemas como entrada. No Foma, essa relação foi construída nesse arquivo contendo as definições das entradas, ou seja, as formas gráficas do conjunto das vogais e do conjunto das consoantes, que podem vir em posição de ataque ou de coda. Observado o conjunto de grafemas, foram definidas as seguintes consoantes em ataque simples:

Foma[1]: define Consoantes [p | b | t | d | g | c | f | v | s | z | x | j | m | n | r | l | ç | h];

Os grafemas <ç> e <h> foram também incluídos como ataque simples. O primeiro não faz parte do alfabeto da língua portuguesa, mas representa o fonema /s/ na escrita. O segundo não possui representação fonológica, no entanto, etimologicamente, é um grafema em muitas palavras do português (honestidade, homem, etc.) e faz parte de alguns dígrafos.

No português, admite-se o máximo de duas consoantes na posição de ataque complexo, sendo que a primeira deve ser uma obstruinte [p, b, t, d, k, g, f, v] e a segunda uma líquida [l, r]. Como foi visto no tópico 4.2.2, a língua portuguesa apresenta distância de sonoridade menor do que as líquidas ou igual a elas. Na concatenação de <p> com <r>, ocorre aumento de sonoridade de uma oclusiva para uma líquida; já na sequência de <p> e/ou <n> e <s>, ocorre inserção de vogal entre elas na fala. No ataque complexo, o sistema obedece aos princípios de sonoridade, pois não é possível a inserção de uma consoante não líquida em posição de segunda consoante. O quadro, a seguir, mostra os grupos de ataque complexo permitidos no (PB)

Quadro 13: Ataque complexo

Grupos de /r/	Entrada	Saída	Grupo de //	Entrada	Saída
Pr	prato	["pratu]	PI	Pleno	["ple~nu]
Br	branco	["bra~ku]	BI	Emblema	[e~"ble~ma]
Tr	trator	[tra"toh]	TI	Atleta	[a"tlet6]
Dr	drácula	["drakul6]	DI	dlacula*	???
Kr	creme	["kre~mi]	KI	Cliente	[kli"e~ti]
Gr	grave	["gravi]	GI	Glória	["glOria]
Fr	fricativa	[frika"tiv6]	FI	Flauta	["flawt6]
Vr	livro	["livru]	VI	livlo*	???

Fonte: adaptado de Silva (2014, p. 156).

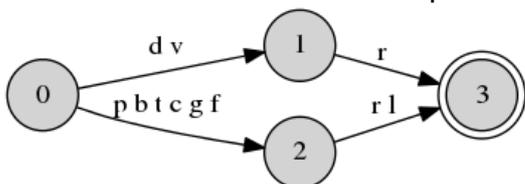
* Simulação da voz de uma criança para testar o sistema.

Esse princípio foi aplicado definindo, no arquivo, as entradas do ataque complexo formado pela concatenação das consoantes obstruintes e das líquidas.

```
define C1 [p|b|t|d|c|g|f|v];
define C2 [r|l];
define Ataquecomplexo [[C1 C2]-[d l v l]];
```

Essas definições computacionais foram inscritas dentro de um arquivo para que o sistema beta-gerador pudesse acessá-las e executá-las como *script* do programa. A concatenação das consoantes em C1 e C2 gera 16 possibilidades de ataques complexos, incluindo aí <vl> e <dl>. No entanto, aplicamos uma regra de restrição para que nessas não houvesse a possibilidade de esse ataque complexo acontecer. Podemos ver no grafo abaixo que mostra 4 estados, 11 arcos, 14 caminhos.

Gráfico 15 – Autômato do ataque complexo do português



Fonte: Elaborado pelo próprio autor.

A formação desse autômato impede que haja a possibilidade de ter alguma entrada que não possa ser da língua. Uma como <ptanta> não pode ser entrada para o sistema.

Além do ataque, há os dígrafos (emprego de duas letras para a representação gráfica de um som da língua) que são constituídos por meio de letras diacríticas. Em português, há seis letras diacríticas para os dígrafos consonantais (<h, r, s, c, ç, u>) originando os seguintes pares de grafemas: ch, lh e nh e as germinadas rr e ss, xc, sc, sç; duas para as vogais nasais (<m> e <n>), fenômeno da nasalização e, por fim, gu e qu. Podemos ver no quadro 14 abaixo os dígrafos consonantais.

Os dígrafos têm como saída um símbolo nulo, o zero fonético [0]. O

grafema <x>, quando representar o dífono [ks], será representado pelo símbolo maiúsculo [K]. No vocábulo <tórax>, o <x> representa os pares de sons na fala: a oclusiva e a fricativa [ks], logo, o *output* pelo sistema será de [ˈtOraK].

Quadro 14: Dígrafos consonantais

Dígrafos	Entrada	Saída	Dígrafos	Entrada	Saída
ch	chaves	[ˈʃOavis]	Rr	carro	[ˈkaʁOu]
lh	coalhada	[koaˈLɔad 6]	Ss	Cassado	[kaˈsɔadu]
nh	montanha	[moˈ˜ta˜J 06]	Sc	Nascer	[naˈsɔe0]
			Sç	Nasço	[ˈnas0u]
			Xc	exceto	[eˈs0etu]
			Xs	Exsudar	[es0uˈda0]*

Fonte: Elaborado pelo próprio autor.

* Todos esses exemplos foram executados no sistema.

Os dígrafos não foram desenvolvidos por meio da concatenação de consoantes como acontece no ataque complexo, mas pelas relações entre o conjunto dos símbolos dentro das cadeias. O dígrafo <rr> é a concatenação da coda <r> com o ataque simples <r>. A transformação do segundo "erre" em fone [0] é feito por meio de regras de reescrita, de modo que o sistema distingue a palavra <nascer> de <casca>, por exemplo, diferenciando quando a letra <c> é um dígrafo e quando é ataque silábico: <nascer> -> [naˈs0eh]; <casca> -> [ˈkaska].

Outro elemento da estrutura silábica a destacar é a rima, que se constitui de núcleo e coda. Há sete grafemas do português que podem aparecer em posição de coda silábica: <n>, <m>, <l>, <s>, <z>, <x> e <r>. Os grafemas <n> e <m> marcam apenas a nasalização da vogal anterior, logo, na transcrição, elas não têm valor de consoante, mas de um traço de nasalização, por isso, o símbolo que as representam é o diacrítico til [˜]. Assim, todas as consoantes nasais, nessa posição, são transformadas nesse traço por meio da regra de reescrita.

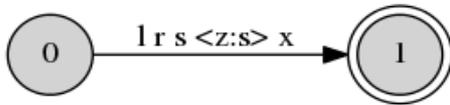
Na fala potiguar, o <l> é semivocalizado [w], ou seja, o <l> torna-se um som muito próximo de [u]. Na forma fonética, essa letra representa uma semivogal. Em alguns casos, não é possível distinguir se a semivogal foi produzida a partir dessa letra ou de uma vogal. A forma fonética [ˈmaw] constitui-se de duas formas

gráficas <mal> e <mau>.

As letras <s>, <z> e <x> têm como *output* o som fricativo surdo [s], a letra <r> tem como representação a fricativa glotal [h] e, em final de palavra, essa letra é representada pelo zero fonético [0].

No sistema *Foma*, a coda é construída com base em um arquivo que contém as definições de entrada formadas pelo conjunto grafemático, conforme representado no gráfico abaixo.

Gráfico 16 – Autômato da coda silábica

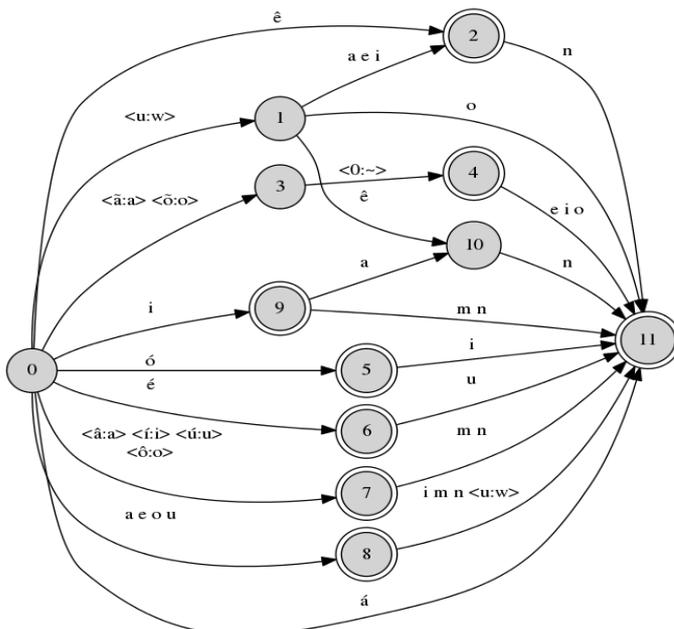


Fonte: Elaborado pelo próprio autor.

As vogais constituem o núcleo silábico e são os únicos elementos obrigatórios na sílaba. Elas podem vir ou não acompanhadas de uma outra vogal com o som foneticamente mais fraco, chamado de *glide*, formando um ditongo ou tritongo. O núcleo da sílaba pode ser precedido ou sucedido de uma ou mais consoantes.

O núcleo silábico é composto por vogais individuais ou formando ditongos e tritongos. Ele é construído pelas definições de entradas que são formadas pelo conjunto grafemático que vem nessa posição. O gráfico abaixo mostra a formação do núcleo silábico por meio de uma rede de estados finitos.

Gráfico 17 – Autômato da rima



Fonte: Elaborado pelo próprio autor.

Este capítulo apresentou uma visão geral da constituição do transdutor para a variedade linguística potiguar, destacando i) o tipo de transcrição e os diacríticos de marcação da tonicidade, da nasalização e da palatalização das palavras; ii) a relação da forma gráfica para a forma fonética com suas respectivas peculiaridades; iii) a constituição do algoritmo e arquitetura do sistema e iv) a implementação dos módulos, utilizando os comandos e operadores do Foma.

4.3 Avaliação e análise do sistema

Para Bird, Klein e Loper (2009), a avaliação de um modelo computacional tem como objetivo principal decidir se o sistema classifica com precisão um determinado padrão linguístico. O resultado da avaliação é importante para compreender até que ponto o modelo é confiável, para que fins utilizá-lo e, também, para guiar o desenvolvedor na busca de melhorias futuras para o modelo.

Pensando assim, o processo de avaliação do Potigrafone foi dividido em duas fases: a constituição do *corpus* de referência e a utilização do sistema automático de avaliação. A primeira fase constituiu-se de dois passos: o primeiro foi a pesquisa de palavras no *corpus* CETENFolha e o segundo a transcrição automática deste *corpus* e a correção manual dos erros para usá-lo como arquivo *gold* ou *corpus* de referência. A segunda fase consistiu na avaliação por meio das métricas de *Exact Macht Ratio* e *Labelling Fscore*.

4.3.1 Corpus de referência

O *corpus* de referência foi constituído a partir de uma amostra de 1.500 palavras do *corpus* CETENFolha/Nilc São Carlos³⁸. Há duas formas de acesso a esse *corpus*: (1) consultando por meio do ACDC no endereço eletrônico: <http://www.linguateca.pt/ACDC/>. Esse programa acessa os *corpora* da linguatca, dentre os quais o "NILC/São Carlos". Através do ACDC, é possível procurar palavras e expressões dentro do CETENFolha, usando a restrição [classe="JOCF"]. Todavia, a pesquisa feita nesse sistema retorna ao usuário até novecentas e noventa e nove ocorrências. (2) cadastrando-se com o nome e e-mail para o qual será enviado um

³⁸ O anexo C mostra os dados referentes ao *corpus* CENTENFolha/Nilc São Carlos. Acesso pelo *link* <http://www.linguateca.pt/acesso/contabilizacao.php#listaPosSAOCARLOS>.

usuário e uma senha de acesso e, assim, fazer o download do CETENFolha, a fim de ter acesso ao texto completo, para efeitos de processamento computacional complexo.

Optamos pela segunda forma de acesso e separamos o número de palavras mencionadas acima. Observamos que as palavras e os símbolos estão organizados em ordem decrescente de ocorrência no *corpus*, facilitando a separação dos termos mais frequentes para os menos frequentes e contribuindo para que o sistema fosse analisado em um conjunto de palavras que não se repetia. Além disso, antes da execução do segundo passo, removemos todos os nomes próprios, as palavras estrangeiras, palavras que continham erros de digitação, números, palavras com letras maiúsculas, as palavras separadas por hífen e as palavras monossilábicas, pois algumas destas não apresentam acentuação tônica, como a palavra <de>.

Dado esse passo, prosseguimos com a primeira transcrição automática dessa amostra por meio do sistema, objetivando a constituição do *corpus* de referência e iniciamos a correção dos erros de transcrição para utilizá-lo como *corpus* padrão ouro na avaliação do sistema.

4.3.2 Medidas de avaliação do sistema

Após a transcrição automática daquela amostra, fizemos a avaliação automática do Potigrafone por meio das métricas *Exact Match Ratio* e *Labelling FScore*. A avaliação automática faz uso da comparação entre os textos para determinar a qualidade na transcrição fonética das palavras. Uma ou mais transcrições feitas por humanos são utilizadas como referências, ou seja, como padrão *gold*.

A métrica de avaliação comumente usada na transcrição fonológica é a acurácia. Essa métrica foi utilizada para a avaliação do Grafone, apresentado no tópico 2.1.4, e no conversor probabilístico de máxima entropia, no tópico 2.1.3. A acurácia é "A métrica mais simples que pode ser usada para avaliar um classificador e medir os percentuais de entrada no conjunto de teste que o classificador rotulou corretamente."³⁹ (BIRD, KLEIN e LOPER, 2009, p. 239). Ela calcula uma pontuação

³⁹ The simplest metric that can be used to evaluate a classifier, accuracy, measures the percentage

para um modelo, comparando as entradas em um conjunto de teste com os símbolos corretos para essas entradas, ou seja, a tarefa fundamental desse método é a classificação binária dos resultados obtidos.

Não optamos por essa métrica, pois o Potigrafone faz a transcrição das palavras considerando que uma entrada (uma forma gráfica) pode ter uma ou mais saída(s), uma ou mais forma(s) fonética(s). Por isso, esse método retorna apenas os dados relativos a uma saída do sistema, desconsiderando as demais transcrições para uma palavra. Logo, os dados apresentados pelo sistema não seria falseados, uma vez que pontuaria apenas as ocorrências corretas ou incorretas, ou seja, uma avaliação binária.

Por isso, as métricas *Exact Match Ratio* e *Labelling FScore* são medidas que fazem a classificação multitópica e avaliam através de correspondência de rótulo e de classe parcial ou total, considerando todas as classes e seus rótulos como equivalentes. Além disso, essas medidas, assim, contam com a identificação correta ou incorreta dos rótulos, independentemente de sua ordem ou *rank*. (SOKOLOVA e LAPALME, 2009, p. 430)⁴⁰.

Apresentamos, a seguir, as fórmulas matemáticas expressas no trabalho de Sokolova e Lapalme (2009) que serviram de fundamento para a construção da ferramenta desenvolvida por França e apresentada por Souza *et alia* (2015).

Tabela 5 - Classificação das medidas multitópicas para avaliação do sistema

Medida	Fórmula	Foco de avaliação
Exact Match Ratio	$\frac{\sum_{i=1}^n I(L_i^d = L_i^c)}{n}$	A classificação média exata por transcrição.
Labelling Fscore	$\frac{\sum_{i=1}^n \frac{2 \sum_{j=1}^l L_i^c[j] L_i^d[j]}{\sum_{j=1}^l (L_i^c[j] + L_i^d[j])}}{n}$	A classificação média da transcrição com resultados parciais.

Fonte: Sokolova e Lapalme (2009, p. 430).

of inputs in the test set that the classifier correctly labeled.

⁴⁰ The quality of multi-topic classification is assessed through either partial or complete class label matching; the latter is often referred to as exact matching. We consider all classes and their labels as being equivalent. These measures thus count correct or incorrect label identification independently of their order or rank.

A primeira métrica faz a relação de correspondência exata dos grafemas com a transcrição fonética, ou seja, o sistema de avaliação automática, por meio dessa métrica, conta apenas as correspondências exatas entre o grafema previsto e o fone verdadeiro. Considere o seguinte exemplo hipotético:

- (1) <caixa> - > ["kajSa]
- (2) <caixa> - > ["kaSa]
- (3) <caixa> - > ["kaiSa]

Está previsto, na transcrição fonética do Potigrafone, que o ditongo decrescente <ai> deve ser monotongado. Assim, tem-se duas saídas para essa palavra que correspondem às transcrições (1) e (2). Por meio da medida *Exact Match Ratio*, o cálculo se faz, como diz Souza *et alia* (2015), atribuindo o valor *um* a cada acerto e *zero* a cada erro, somando-os e, em seguida, dividindo-os pela quantidade total de ocorrência na amostra. Dessa forma, obtém-se a medida de acerto por correspondência entre grafema e sons da fala potiguar.

A segunda métrica, a medida *Labelling FScore*, avalia o desempenho médio da transcrição, tendo em conta a correspondência parcial de acerto, ou seja, a avaliação é feita considerando os acertos parciais. Considerando o exemplo acima, a palavra <caixa> recebeu três transcrições das quais cada transcrição possui um valor de 33,3%, ou seja, o sistema acertou apenas 66,6% na transcrição dessa palavra e errou 33,3%. Com a utilização dessa medida, consegue-se mensurar e conhecer os erros e os acertos parciais na transcrição fonética.

4.3.3 Avaliação e resultados do sistema

Como foi dito acima, o sistema foi avaliado automaticamente através de um programa⁴¹ que avalia estatisticamente o desempenho, utilizando as medidas *Exact Match Ratio* e *Labelling FScore*. Partindo do *corpus* de referência intitulado *dicio_corpusbr.txt*, construímos quatro dicionários *gold*, que estão nomeados e descritos no quadro 15, e quatro transcrições automáticas dos modelos apontados no quadro 16, com seus respectivos arquivos transcritos. Esses dados estão estruturados da seguinte forma: na primeira coluna as formas gráficas e na segunda, as formas fonéticas, separadas por um espaço em branco na horizontal e por uma

⁴¹ O código do sistema está disponível em <https://github.com/mardoniofranca/Multi-labelled-Suite>.

linha em branco na vertical, como podemos ver ilustrado na figura 19, a seguir.

Figura 19 – Alinhamento padrão dos dicionários e das transcrições fonéticas

casa	"kaz6
réptil	"hEptiw
réptil	"hEpetiw
réptil	"hEpitiw
caixa	"kajS6
caixa	"kaS6

Fonte: Elaborado pelo próprio autor.

Esses dados foram alinhados dessa forma para que o programa executasse o cálculo estatístico nas métricas apresentadas acima e, com os dados em mãos, verificássemos a interferência dos módulos da tonicidade e dos fenômenos fonéticos sobre o desempenho do sistema.

Quadro 15 - Conjunto de *corpus* ouro para a avaliação do sistema

Nº	Corpus padrão ouro	Descrição
1	dicio_corpusbr_GOLD	Contém as formas gráficas e fonéticas.
2	dicio_corpusbr_GOLD_SF	Contém as formas sem os fenômenos linguísticos.
3	dicio_corpusbr_GOLD_STF	Contém as formas sem os fenômenos linguísticos e sem as regras de tonicidade.
4	dicio_corpusbr_GOLD_ST	Contém as formas sem as regras de tonicidade.

Fonte: Elaborado pelo próprio autor.

Cada dicionário do quadro 15 pode ser assim descrito:

- O (1) apresenta o padrão ouro para a transcrição das palavras considerando a forma de "uma-para-uma" e de "uma para várias", ou seja, de uma forma gráfica para uma forma fonética ou de uma forma gráfica para várias formas fonéticas, como foi ilustrado na figura 18.
- O (2) mostra o padrão ouro para a transcrição das formas gráficas sem os fenômenos fonéticos, mas com a acentuação silábica das

palavras.

- O (3) elenca o padrão ouro para a transcrição das formas "puras", sem a marcação da tonicidade e sem os fenômenos fonéticos.
- O (4) dispõe sobre o padrão ouro para a transcrição das formas gráficas sem a acentuação fonética, mas com os fenômenos fonéticos.

Após a constituição desses dicionários para a avaliação, analisamos o sistema considerando quatro partes distintas, verificando separadamente o desempenho do Potigrafone, primeiro contendo todos os módulos e depois com a ausência dos módulos responsáveis pelo processamento dos fenômenos fonéticos, da tonicidade ou de ambos os módulos. Ressaltamos que esses módulos foram apresentados no tópico sobre a implementação do sistema e está disponível no apêndice B deste trabalho.

Essa divisão tentaram responder às seguintes indagações:

- Qual dos dois módulos mais influencia o desempenho do transdutor?
- Qual é o nível de exatidão na transcrição das palavras sem a interferência desses módulos?

Para respondê-las, executamos cada parte do programa separadamente na amostra *dicio_corpusbr.txt*. Como podemos ver no quadro 16, executamos os transdutores nesse dicionário e obtivemos os seguintes: o Potigrafone - contém todos os módulos e gerou o arquivo *dicio_corpusbr_POTIGRAFONE*; o PotigrafoneSF - não contém o módulo dos fenômenos fonéticos e gerou o arquivo *dicio_corpusbr_POTIGRAFONE_SF*; o PotigrafoneSTF - não contém os módulos dos fenômenos fonéticos e da tonicidade das palavras e forneceu o arquivo *dicio_corpusbr_POTIGRAFONE_STF* e, por fim, o PotigrafoneST - não tem as regras de tonicidade e forneceu o arquivo *dicio_corpusbr_POTIGRAFONE_ST*.

Quadro 16 - Execução do sistema na amostra *dicio_corpusbr.txt*

Modelos	Arquivos
Potigrafone	<i>dicio_corpusbr_POTIGRAFONE</i>
PotigrafoneSF	<i>dicio_corpusbr_POTIGRAFONE_SF</i>
PotigrafoneSTF	<i>dicio_corpusbr_POTIGRAFONE_STF</i>
PotigrafoneST	<i>dicio_corpusbr_POTIGRAFONE_ST</i>

Fonte: Elaborado pelo próprio autor.

A avaliação automática foi feita alinhando os *corpus* de padrão ouro apresentados na figura 19 com os arquivos transcritos pelo sistema mostrado no quadro 16, seguindo a ordem presente nas tabelas. O alinhamento no programa estatístico foi executado por meio dos seguintes passos: (1), na linha 25 do arquivo *main.java*, escrevemos o título do dicionário padrão ouro, por exemplo, *dicio_corpusbr_GOLD* do quadro 15, e (2), na linha 26, do arquivo *main.java*, digitamos o título do arquivo transcrito pelo sistema correspondente ao modelo que queríamos avaliá-lo, que nesse exemplo específico, o arquivo *dicio_corpusbr_POTIGRAFONE* do quadro 16, como podemos ver ilustrado esse processo na figura 20. Esses procedimentos foram feitos para os demais modelos e obtivemos os resultados que estão apresentados no gráfico 20 para as duas métricas utilizadas.

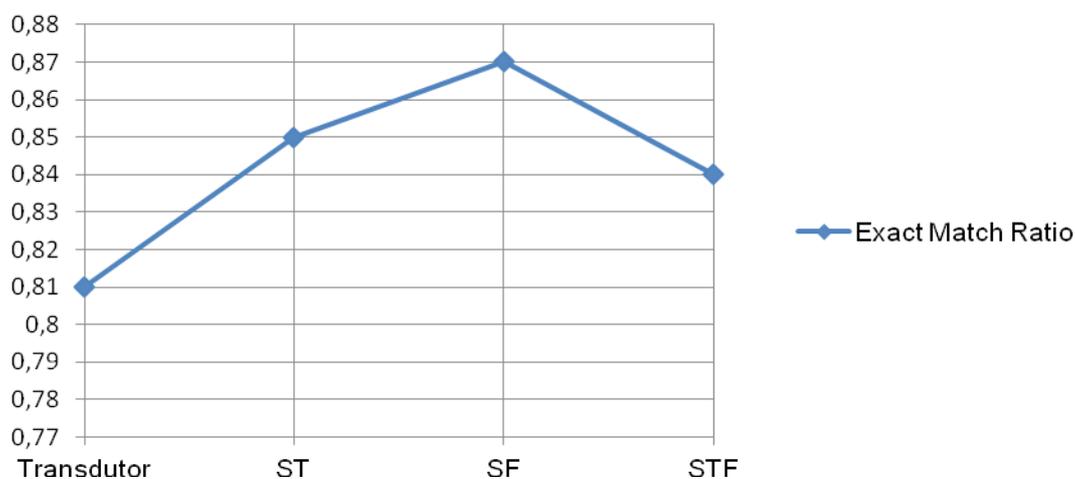
Figura 20 – Obtenção dos dados pelo sistema estatístico

```
22 public class Main {
23
24     public static String PATH = "data/test/";
25     public static String filesG[] = {"dicio_corpusbr_GOLD"};
26     public static String filesT[] = {"dicio_corpusbr_POTIGRAFONE"};
```

Fonte: Elaborado pelo próprio autor.

No gráfico 18, apresentamos o resultado da métrica *Exact Match Ratio* que contribuiu para avaliar a relação de correspondência exata dos grafemas com a transcrição fonética, ou seja, a avaliação automática conta os valores correspondentes entre as formas gráficas e fonéticas. Quando os valores do arquivo transcrito não correspondem exatamente aos valores presentes no dicionário padrão ouro, o sistema calcula como um erro. Por exemplo, se a transcrição da palavra <caixa> for ["kaiS6] e o padrão ouro para essa palavra for [kajS6], então, o sistema calcula como um erro.

Gráfico 18 – Medida Exact Match Ratio

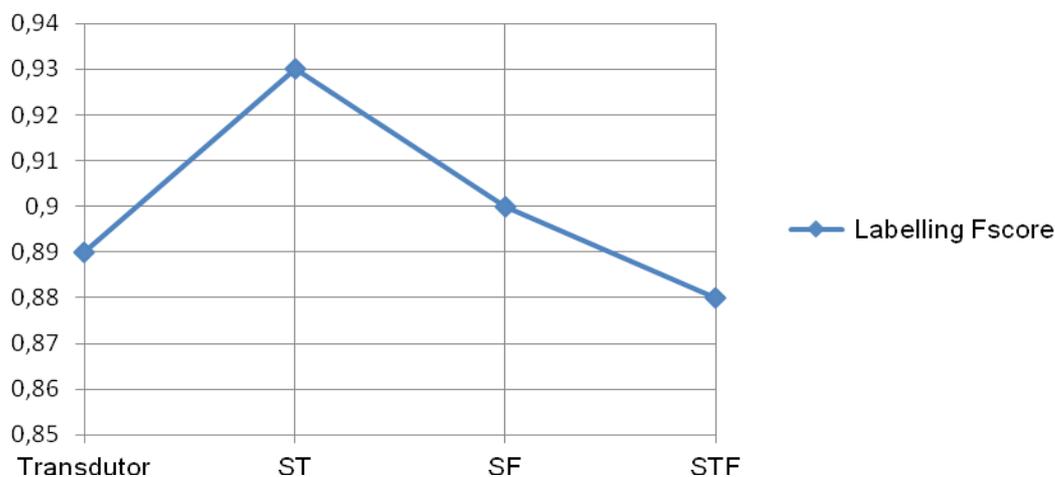


Fonte: Elaborado pelo próprio autor.

Esse gráfico expõe os resultados dessa medida para os sistemas. Nele é fácil observar que o modelo mais exato é o sem a transcrição dos fenômenos (SF) com 87%; depois vem o modelo sem a marcação de tônica (ST) com 85%; em terceiro, o modelo sem os fenômenos e sem a tonicidade (STF) com 84% e, por último, o modelo integrado com todos os módulos com 81% de exatidão na transcrição das palavras.

Partindo dessas informações, observamos que o sistema aumenta a exatidão com a ausência dos módulos da tonicidade e dos fenômenos fonéticos, pois eles interferem na exatidão da transcrição fonética, cerca de 4% e 6%, respectivamente. No entanto, quando o sistema é avaliado com a ausência desses módulos, ele apresenta 84% de exatidão, 3% a mais do que o sistema com todos os módulos. Então, podemos dizer que o módulo da marcação tônica interfere menos no resultado do sistema do que o módulo dos processos fonéticos; porém, há outros fatores sistemáticos que interferem na exatidão do sistema e esses fatores estão presentes em outros módulos.

O gráfico 19 mostra a segunda medida, *Labelling FScore*, que avalia o desempenho médio da transcrição, tendo em conta a correspondência parcial de acerto, ou seja, a avaliação é feita considerando os acertos totais e parciais, sendo que, para cada acerto total, o cálculo é obtido sobre 100% e, quando é parcial, o cálculo é obtido pela média simples dos valores.

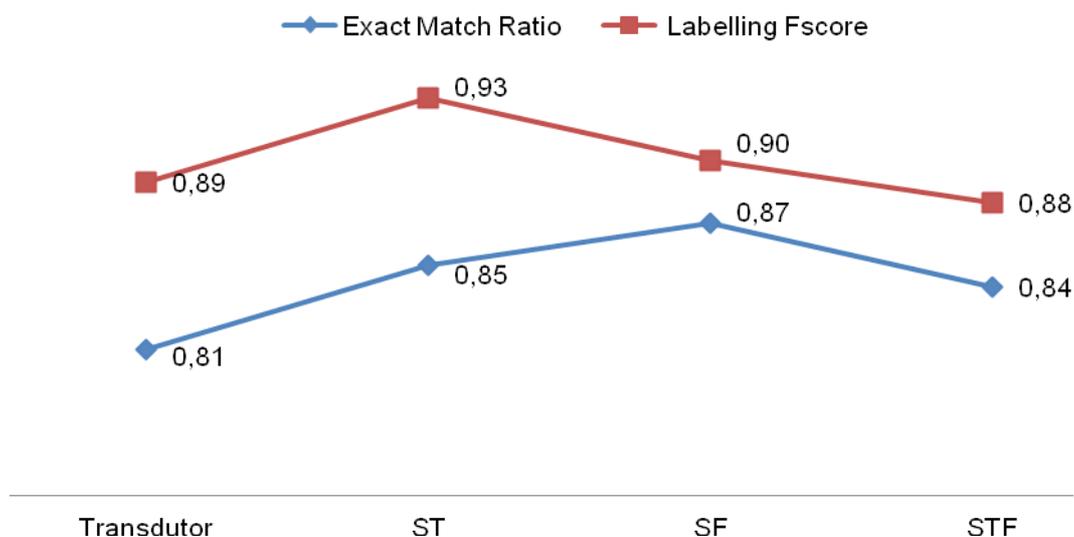
Gráfico 19 – Medida *Labelling Fscore*

Fonte: Elaborado pelo próprio autor.

Nesse gráfico, é possível verificar que o modelo ST apresenta o melhor desempenho nos acertos parciais e totais na transcrição com 93%; em segundo, o modelo SF, com 90%; em terceiro, o modelo integrado com todos os módulos com 89% e, por último, o modelo STF com 88%. Com base nesses resultados, observamos que ocorre aumento de 4% em acertos totais e parciais quando o Potigrafone é executado sem o módulo de tonicidade, reduzindo em 3% desse valor, quando é transcrito sem os fenômenos fonéticos e diminuindo esse valor em 2%, quando não utiliza nenhum desses dois módulos.

O gráfico 20 exibe os resultados das duas métricas: *Exact Match Ratio* e *Labelling Fscore*. Podemos verificar dois fatos em relação ao desempenho do sistema: primeiro, o desempenho do Potigrafone aumenta em exatidão e nos acertos totais e parciais, quando a transcrição é feita sem a marcação tônica e sem os fenômenos fonéticos e, segundo, outros módulos também interferem no desempenho, uma vez que os valores com todos os módulos e sem esses dois módulos se assemelham, tanto para a medida *Exact Match Ratio* como para a medida *Labelling Fscore*.

Gráfico 20 – Dados da avaliação do sistema



Fonte: Elaborado pelo próprio autor.

Assim, podemos mencionar que esses resultados servem para verificar o desempenho do Potigrafone e analisar os principais erros cometidos por esse sistema. Também contribuíram para apontar alguns erros referentes aos três modelos e, conseqüentemente, ao programa de uma forma geral.

Com a marcação de tonicidade, ocorreram dois tipos de erros: a duplicação de tonicidade e ausência de marcação. O primeiro caso ocorreu com as palavras grafadas com os dígrafos <sc> e <ss> que apresentaram duas marcações de tônicas, como nas palavras <crescimento> -> [kre"s0i"me~tu], <necessidade> [nese"s0i"dadi] e o segundo caso aconteceu nas palavras com o <lh>, como em <mulher> [muL0e0]. No entanto, o conjunto de letras <qu> e <gu> não apresentaram erros nos contextos de ditongo, como em <qualidade> [kwali"dadi], e nem na condição de dígrafo <questão> [k0es"ta~w].

A marcação do fenômeno da nasalidade apresentou erros pontuais na transcrição nas palavras terminadas em <mente>. Na variedade potiguar, não ocorre nasalidade na fronteira com essa desinência, como em <constantemente> [ko~sta~ti"me~ti].

A maior parte dos erros está presente no núcleo silábico. A representação das letras <a>, <e> e <o> em final de palavra, como nas palavras cas[6]s, pel[i]s e tod[u]s, respectivamente, mostra erro apenas no grafema <a> quando ocorre no final de palavra, precedido por uma vogal, como na palavra <havia> [0a"vj6], em que

deveria ser transcrita como duas vogais, ou seja, um hiato. Porém, a maioria dos erros são provocados pela confusão gerada pelo grau de abertura dos sons que devem corresponder aos grafemas <o>, <e> e <a>, os quais podem ser pronunciados como [o] ou [O], [e] ou [E], [ɔ] ou [a], como em <f[O]rmação>, <[E]la> e <padari[a]>. O ditongo nasal <ém >, em final de palavra, está sendo transcrito de forma errada, como em <além> [a"lE~], onde o símbolo correto deveria ser [e~] fechado.

A classificação desses erros norteará a constituição de novas regras para a correção e o melhoramento do desempenho do sistema, tanto na exatidão quanto na correspondência total e parcial do sistema.

5 CONCLUSÃO

A língua portuguesa está situada entre as línguas de ortografia razoavelmente fonêmica, como diz Bechara (2005). Todavia, nem sempre é verdadeira a afirmação de que cada sequência ortográfica corresponde a uma sequência de fonema. Os sistemas computacionais que se propõem a tarefa de transcrição fonética não executam uma atividade automática trivial. Eles contribuem para o aperfeiçoamento no pré-processamento do sistema de síntese de fala e reconhecimento de voz e também auxiliam os linguistas, os foneticistas e os lexicógrafos no processo de transcrição fonética.

Os conversores mostrados no capítulo 2 foram desenvolvidos tendo como foco uma aplicação específica. Por um lado, o Grafone é um sistema híbrido que faz a transcrição fonológica da língua sem, contudo, considerar uma fala específica; por outro, o *Petrus* foi desenvolvido numa abordagem de regras e executa a transcrição fonética dos grafemas para a variedade paulista. O Potigrafone apresenta semelhanças e diferenças dos demais conversores, pois é um transdutor de estados finitos, mas diverge pelo fato de fazer a transcrição fonética e não fonológica dos grafemas. Ele apresenta semelhança com *Petrus*, pois ambos fazem a transcrição fonética, mas são distintos quanto às variedades linguísticas e, conseqüentemente, quanto à variante transcrita. Assim, o Potigrafone não é uma cópia dos demais gráficos e fônicos existentes, uma vez que sua aplicação específica se dá para a variedade linguística potiguar e, ressalte-se, apresenta a transcrição com alguns fenômenos fonéticos, o que não ocorre com os outros sistemas.

Concluindo o trabalho, destacamos, a seguir, alguns fatos relevantes sobre o desenvolvimento desse transdutor.

Primeiramente, ressaltamos que os conversores desenvolvidos para o PE e para o PB, apresentados no capítulo 2, são sistemas que apresentam bom desempenho na execução de transcrição de grafema para fonema. No entanto, tais sistemas não executam a transcrição fonética para a fala potiguar e nem consideram a possibilidade de mais de uma saída para uma mesma entrada. A única exceção a esses processos automáticos é o sistema de transcrição fonética *Petrus*, no entanto, este se restringe a duas saídas para as formas escritas que podem apresentar o fenômeno da alternância vocálica, como a palavra <gosto> que recebe duas transcrições: [gostu] e [gOstu], em que a mesma apresenta duas saídas, uma para a

vogal fechada [o] e outra para a vogal aberta [O].

Em segundo lugar, tratamos as formas gráficas (formas de entrada) da língua portuguesa enfatizando algumas peculiaridades, como o uso de diacríticos e de letras diacríticas consonantais e vocálicas que exigem maior atenção na implementação de um sistema gráfico e fônico. Muitos erros do sistema estão localizados nessas peculiaridades da escrita e, para elas, o Potigrafone apresentou alguns diferenciais quanto à substituição das letras diacríticas consonantais com o símbolo fonético “0”, pois mantém o alinhamento do grafema para fone. Essa forma de relação foi exemplificada na figura 16 – página 92. Quanto às letras diacríticas vocálicas, o transdutor marca a transcrição fonética das palavras que contém vogais nasais e as que apresentam o fenômeno da nasalidade. O diacrítico do SAMPA “~” foi utilizado para marcar esse características da fala potiguar.

Em terceiro lugar, mostramos que a transcrição do grafema para a pronúncia da variedade linguística potiguar é feita utilizando o alfabeto fonético SAMPA – um alfabeto fonético pouco conhecido pelos linguistas, porém é mais adequado para uma aplicação computacional. A escolha desse alfabeto fonético e da linguagem de programação foi fundamental na criação do transdutor. Por um lado, o alfabeto definiu os caracteres fonéticos que o conversor apresenta como forma de saída para o usuário e os diacríticos que marcam a sílaba tônica e os fenômenos da nasalidade e da palatalização; por outro lado, a linguagem de programa contribuiu para a constituição do algoritmo e da arquitetura do sistema e a implementação dos módulos por meio dos comandos e operadores do Foma.

Acrescentamos a isso que a transcrição fonética automática do Potigrafone contempla alguns fenômenos fonéticos presentes na fala potiguar. Alguns desses fenômenos estão presentes em outras variedades linguísticas, como o fenômeno da nasalidade, porém, como partimos de um corpus de transcrição da fala potiguar, registramos que os fenômenos descritos pelo sistema são dessa variedade linguística. Além disso, o uso dos atlas linguísticos fundamenta a transcrição automática do programa e a torna o mais próximo possível da fala potiguar. Por isso, dependendo da palavra, o sistema apresenta uma ou mais transcrição para o mesmo vocábulo.

Em quarto lugar, vemos que os conceitos advindos da Fonologia, de modo geral, e da Fonologia Computacional, de forma particular, se fundem nas técnicas de aplicação e no modo com que o transdutores relacionam as informações

fonológicas através das cadeias superiores com as cadeias inferiores. Isso ocorre de forma tão intrínseca na transcrição que o usuário final só consegue compreender se conhecer os mecanismos linguísticos utilizados na aplicação computacional. Para o desenvolvimento do sistema, esses conceitos podem se perder, quando não há bases empíricas consistentes que sejam capazes de auxiliar na construção do sistema e na definição de critérios que conduzam a implementação de forma orientada. Nesse caso, os atlas linguísticos deram subsídios necessários à implementação do sistema.

Em quinto lugar, optamos por não avaliar o sistema com as medidas de avaliação dos outros sistemas, ou seja, avaliar o nível de acurácia ou de precisão, por serem medidas binárias. Essas medidas avaliam apenas o desempenho na transcrição de um grafema para um fonema ou fone e o Potigrafone pode apresentar uma ou mais formas de saída para uma mesma palavra. Por isso, utilizamos duas medidas adequadas para a avaliação do sistema: as medidas de *Exact Match Ratio* e *Labelling FScore*. Essas métricas foram consideradas as mais adequadas para a avaliação do sistema, pois elas apresentam resultados mais precisos para os programas com mais de uma saída. Essas medidas fazem a classificação multinível e avaliam, quer através da correspondência de rótulos e de classes parcial ou total, considerando todas as classes e quer por meio de rótulos equivalentes.

Com o uso dessas medidas, os resultados mostraram que o desempenho do Potigrafone foi de 81% para a medida *Exact Match Ratio* e 89% para a medida *Labelling Fscore*. Desse modo, o transdutor apresentou boa performance na exatidão e nos acertos totais e parciais, no que se refere à transcrição sem a marcação tônica e sem os fenômenos fonéticos. Outro resultado importante foi que os módulos de transformação fonéticas também interferiram no desempenho, uma vez que os valores mostrados com a presença e com a ausência de todos os módulos se assemelharam em valores (ver gráfico 20).

Essas conclusões nos levam a considerar que a classificação dos erros norteará a constituição de novas regras que façam a correção e, conseqüentemente, o melhoramento do sistema, tanto na exatidão da transcrição quanto na correspondência total e parcial do sistema. Além disso, essas conclusões já nos apontam alguns caminhos que devem ser trilhados para aumentar o desempenho do sistema. O primeiro deles é a marcação de tonicidade para a qual ocorreram dois tipos de erros: a duplicação de tonicidade e ausência de marcação. O primeiro caso

ocorreu em palavras grafadas com os dígrafos <sc> e <ss> que apresentaram duas marcações de tônicas, como nas palavras <crescimento> -> [kre"s0i"me~tu], <necessidade> [nese"s0i"dadi]. O segundo caso se deu nas palavras com o <lh>, como em <mulher> [muL0e0]. No entanto, o conjunto de letras <qu> e <gu> não apresentaram erros nos contextos de ditongo, como em <qualidade> [kwali"dadi] e nem na condição de dígrafo <questão> [k0es"ta~w].

O segundo é a marcação do fenômeno da nasalidade, que apresentou erros pontuais na transcrição de palavras terminadas em <mente>. Na variedade potiguar, não ocorre nasalidade na fronteira com essa desinência, como em <constantemente> [ko~sta~ti"me~ti]. A maior parte dos erros está presente no núcleo silábico. A representação das letras <a>, <e> e <o> em final de palavra, como nas <casas>, <peles> e <todos>, respectivamente, mostra-se erro apenas no primeiro grafema em ditongo no final, como em <havia> [0a"vj6], não distinguindo a transcrição quando é hiato e quando é ditongo crescente. Porém, a maioria dos erros são provocados por confusão gerada pelo grau de abertura dos sons que devem corresponder aos grafemas <o>, <e> e <a>, os quais podem ser pronunciados como [o] ou [O], [e] ou [E], [6] ou [a], como em <f[O]rmação>, <[E]la> e <padari[a]>. O ditongo nasal <ém >, em final de palavra, está sendo transcrito de forma errada, como em <além> [a"lE~], onde o símbolo correto deveria ser [e~] fechado.

As conclusões e as considerações acima mencionadas nos conduzem a possíveis trabalhos diretamente relacionados ao Potigrafone, a saber: (1) numa ação mais urgente, implementar o conversor para o Alfabeto Fonético Internacional (AFI) como forma opcional na transcrição fonética, uma vez que os símbolos desse alfabeto é mais conhecido pela comunidade de linguistas e foneticistas; (2) comparar o desempenho desse sistema com os outros apresentados no capítulo 2, pois, dessa forma, teremos real compreensão da transcrição do transdutor em relação aos outros sistemas. Lembramos que tal comparação só poderia ser feita após o desenvolvimento do sistema. Todavia, ressaltamos que não é nosso objetivo compará-lo com outros sistemas semelhantes, uma vez que a comparação já exige um longo trabalho de pesquisa que deve ser feito após o desenvolvimento do sistema. (3) integrar esse sistema com o *POS-tagger* (etiquetador morfossintático), contribuindo para a distinção na transcrição de palavras que apresentam alternância vocálica, como <ac[o]rdo> (substantivo) e <ac[O]rdo> (verbo) e também para

aplicação em transcrição fonética de sentenças. Os conversores mencionados no capítulo 2 e o Potigrafone não fazem a transcrição fonética de sentença, mas apenas de palavras isoladas e, por fim, (4) usar o sistema no ensino da variação linguística, da fonologia da língua e da relação da fala com escrita. Esse trabalho não estaria diretamente relacionado ao desenvolvimento ou ao melhoramento do sistema, mas à contribuição fora do escopo das aplicações tecnológicas e das pesquisas linguísticas do sistema.

Dessa forma, cremos que um sistema como esse pode ser utilizado para a reflexão linguística da fala potiguar e dos fenômenos fonéticos que estão presentes nas variedades linguísticas do país.

REFERÊNCIAS

ALENCAR, Leonel Figueiredo de. (2011) Línguas formais, gramáticas e autômatos no processo automático das palavras, In: Alencar, L. F e Othero, G. de A. **Abordagens computacionais da teoria da gramática**. Campinas-SP: Mercado de Letras, p. 13-76.

BARBOSA, Felipe *et alia*. Grapheme-Phone Transcription Algorithm for a Brazilian Portuguese TTS, in: **Proceedings of PROPOR 2003**, Junho, Faro, Portugal, 2003, p. 23 - 30. Disponível em: <http://www.ime.usp.br/~cpaz/downloads/algorithm-portuguese.pdf>. Acesso em: 10 de agosto de 2015.

BARROS, Maria João; WEISS, Christian. Maximum Entropy Motivated Grapheme-To-Phoneme, Stress and Syllable Boundary Prediction for Portuguese Text-to-Speech, **IV Jornadas en Tecnologías del Habla**, 2006, p. 177-182. Zaragoza, España. Disponível em: http://lorien.die.upm.es/~lapiz/rtth/JORNADAS/IV/finals/4jth_127.pdf. Acesso em: 09 de agosto 2015.

BECHARA, Evanildo. **Moderna gramática portuguesa**. 38. ed. Rio de Janeiro: Lucerna, 2005.

BEESELEY, Kenneth R. e KARTTUNEN, Lauri. **Finite-State Morphology: Xerox Tools and Techniques**, 2002.

BISOL, Leda. **Introdução a estudos de fonologia do português brasileiro**. 4. ed. Porto Alegre-RS: EDIPUCRS, 2005. Disponível em: http://books.google.com.br/books?id=TFzWAq-S710C&pg=PA33&lpg=PA33&dq=metastase+na+fonologia&source=bl&ots=w9JIVWSfsV&sig=PFTtpcwbaFafFwPxGvpF3yEcBfE&hl=en&sa=X&ei=B5aEU6riOOvMsQTR8YLABQ&redir_esc=y#v=onepage&q=metastase%20na%20fonologia&f=false

BIRD, Steven. Phonology. In: MITKOV, Ruslan. **The oxford handbook of computational linguistics**. New York: Oxford University Press, 2009, p. 3 -24.

BIRD, Steven, KLEIN, Ewan e LOPER, Edward. Learning to classify text. In: _____. **Natural language processing with python**. United States of America: O'Reilly, 2009, p. 221 - 257. Disponível em: <http://www.nltk.org/book/>. Acesso em: 05/2012.

BLAC, Alan W.; LENZO, Kevin; PAGEL, Vincent. Issues in Building General Letter to Sound Rules, in: **3rd ESCA Workshop on Speech Synthesis**, Jenolan Caves, Australia, p. 77-80, 1988. disponível em https://www.cs.cmu.edu/~awb/papers/ESCA98_lts.pdf. Acesso em: 11 08 2015.

BRAGA, Daniele.; COELHO, Luís; RESENDE Jr., Fernando Gil Viana. A Rule-Based Grapheme-to-Phone Converter for TTS Systems in European Portuguese, **VI Int. Telecommunications Symposium**, Fortaleza-CE, Brazil, 2006, p. 976 – 981.

BOUMA, Gosse. A Finite-State and Data-Oriented Method for Grapheme to Phoneme Conversion, in: **1st Meeting of the North American Chapter of the**

Association for Computational Linguistics, Seattle, USA, 2000, p. 303 -310, disponível em: <http://ucrel.lancs.ac.uk/acl/A/A00/A00-2040.pdf>, acessado em: 07 de agosto 2015.

CAGLIARI, Luiz Carlos. **Análise fonológica**: introdução à teoria e à prática, com especial destaque para o modelo fonêmico. Campinas-SP: Mercado de Letras, 2002.

CASEIRO, D. *et alia*. Grapheme-to- Phone Using Finite-State Transducers, in: **IEEE Workshop on Speech Synthesis**, USA. 2002. disponível em: <http://www.inesc-id.pt/pt/indicadores/Ficheiros/183.pdf>. Acesso em: 09 de agosto 2015.

FRANÇA, Mardônio. Algoritmo do sistema [mensagem pessoal]. Mensagem recebida por <mardfranca@gmail.com> em 02 jun. 2016.

GOMES, Fábio Rogério Rodrigues. **Cartografia linguística e educação na Amazônia: um estudo semânticolexical da fala na/da microrregião Marabá/Pará**. 2007. 287 f. Dissertação (Mestrado em Educação) – Universidade do Estado do Pará, Belém, 2007.

HULDEN, Mans. **Finite-state machine construction methods and algorithms for phonology and morphology**. 2009. 309 f. Tese (doutorado em Linguística) - Arizona State University, Department of Linguistics, 2009.

_____. Finite-State Syllabification. In: _____. YLI-JYRÄ, A.; KARTTUNEN, L.; KARHUMÄKI, J. **FSMNL 2005**, LNAI 4002, 2006, p. 86–96.

_____. Foma: a finite-state compiler and library. **Proceedings of the EACL**, Athens, Greece, 3 April, p. 29–32.

Jargas, Aurélio Marinho. **Expressões Regulares** - Uma abordagem divertida. 4. ed. São Paulo: Novatec, 2012.

KARTTUNEN, Lauri. Finite-State technology. In: MITKOV, Ruslan. **The oxford handbook of computational linguistics**. New York: Oxford University Press, 2009, p. 340 - 357.

KORNAI, András. Phonology. In: _____. **Mathematical Linguistics**. London: Springer, 2008, p. 23 - 48.

MARQUIAFÁVEL, Vanessa; BOKAN, Alessandro e ZAVAGLIA, Claudia. **PETRUS: A rule-based grapheme-to-phone converter for Brazilian Portuguese**. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/propor-demo/2014/004.pdf>. Acesso: 09 de agosto de 2015.

MARQUIAFÁVEL, Vanessa e ZAVAGLIA, Claudia. Transcrição fonética automática para lemas em verbetes de dicionários do Português do Brasil, in: **Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology**, Cuiaba - MT, Brasil, outubro 2011, p. 154-158. disponível em: http://nilc.icmc.sc.usp.br/til/stil2011_English/stil/artigos/Short/STIL2011_SP1.pdf. Acesso em: 10 de agosto de 2015.

MASSINI-CAGLIARI, Gladis; CAGLIARI, Luiz Carlos. Fonética. In: MUSSALIM, Fernanda; BENTES, Anna Christina. **Introdução à Linguística: domínios e fronteiras**. 2. ed. São Paulo: Cortez, 2011.

MASIP, Vicente. **Fonologia, fonética e ortografia do português**. 1. ed. Rio de Janeiro: E.P.U, 2014.

MATTE, Ana Cristina Fricke; MEIRELES, Alexsandro; FRÁGUAS, Cecílio Cosac. SIL Web – analisador fonológico silábico-acental de texto escrito. In: **Revista de Estudos da Linguagem**, v.14, n.1, 2006, p.31-50. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/4976/4514>. acesso em: 10 de agosto de 2015.

MATZENAUER, Carmen Lúcia Barreto. Parecer para qualificação de doutorado [mensagem pessoal]. Mensagem recebida por <carmen.matzenauer@gmail.com> em 01 out. 2015.

MOTA, Jacyra Andrade; CARDOSO, Suzana Alice Marcelino A construção de um Atlas Linguístico do Brasil: o percurso do ALiB. In: **SIGNUM: Est. Ling.**, Londrina, v. 12, n. 1, p. 237-256, jul.2009. p. 237-256. Disponível em: <http://www.uel.br/revistas/uel/index.php/signum/article/viewFile/4243/4603>. Acesso em: 15/08/2015.

PEREIRA, Maria das Neves. **Atlas geolinguístico do litoral potiguar - ALiPTG**. 2008. 136 f. volumes I e II, Tese (doutorado em Linguística) - Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2008.

_____. **Atlas geolinguístico do litoral potiguar - ALiPTG**. 2008. 136 f. volumes II, Tese (doutorado em Linguística) - Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2008.

SCHANE, Sanford A. **Fonologia gerativa**. Trad. Alzira S. da Rocha, Helena M. Camacho, Junéia Mallas. Rio de Janeiro: Zahar Editores, 1975.

SEARA, Izabel Christine; NUNES, Vanessa Gonzaga; LAZZAROTTO-VOLCÃO, Cristiane. **Para conhecer fonética e fonologia do português brasileiro**. São Paulo: Contexto, 2015.

SILVA, Denilson *et alia*. A rule-based grapheme-phone converter and stress determination for Brazilian Portuguese natural language processing, in: **VI International Telecommunications Symposium**, Setember, 2006, Fortaleza-CE, Brazil, 2006, p. 992 – 996.

SILVA, Moisés Batista da . **Atlas linguístico do centro-oeste potiguar**. 2012. 327f. Tese (doutorado em Linguística) –Universidade Federal do Ceará, Centro de Humanidades, Departamento de Letras Vernáculas, Fortaleza, 2012.

SILVA, Thaís Cristófar. **Fonética e fonologia do português**. 10. ed. São Paulo: Contexto, 2014.

_____. **Dicionário de fonética e fonologia**. São Paulo:Contexto, 2011.

SHIBATA, Danilo Picagli. **Tradução grafema-fonema para a língua portuguesa baseada em autômatos adaptativos**. 2008. 104 p. Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais.

SOKOLOVA, Marina; LAPALME, Guy. A systematic analysis of performance measures for classification tasks. In: **Information Processing and Management**, nº 45, 2009, p. 427–437. disponível em: <http://atour.iro.umontreal.ca/rali/sites/default/files/publis/SokolovaLapalme-JIPM09.pdf>. Acesso em: 18/11/2015.

SOUZA, Jéssica de Oliveira. Uma avaliação de analisadores morfológicos do português. disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/tilic/2015/011.pdf>. Acesso em: 04/12/2015.

Teixeira, A.; Oliveira, C., Moutinho, L. On the Use of Machine Learning and Syllable Information in European Portuguese GraphemePhone Conversion, **Proc. PROPOR'2006**, 2006, p.212-215.

VASILÉVSKI, Vera. **Construção de um sistema computacional para suporte à pesquisa em fonologia do português do Brasil**. 2008. 166f. Tese de doutorado - Pós-graduação em Linguística da Universidade Federal de Santa Catarina, 2008.

_____; SCLIAR-CABRAL, Leonor; ARAÚJO, Márcio José. Phonologic and Syllabic Patterns of Brazilian Portuguese Extracted from a G₂P Decoder-Parser. In: **International Journal of Advanced Computer Science**, Vol. 3, No. 8, Pp. 416-427, Aug., 2013.

VEIGA, Arlindo; CANDEIAS, Sara; PERDIGÃO, Fernando. Conversão de Grafemas para Fonemas em Português Europeu – Abordagem Híbrida com Modelos Probabilísticos e Regras Fonológicas. **Linguamática**, Vol. 3 Nº 1, 2 - Dezembro 2011 - Pág. 39–51.

WELLS, J.C. **Computer-coding the IPA**: a proposed extension of SAMPA. University College London. disponível em: <https://www.phon.ucl.ac.uk/home/sampa/ipasam-x.pdf>. Acesso em: 30 agosto 2014.

WELLS, J.C. SAMPA computer readable phonetic alphabet. In Gibbon, D., Moore, R. and Winski, R. (eds.). **Handbook of Standards and Resources for Spoken Language Systems**. Berlin and New York: Mouton de Gruyter. Part IV, section B. 1997. Disponível: <http://www.phon.ucl.ac.uk/home/sampa/index.html>. Acesso em: 08/2014.

APÊNDICE A - TRANSCRIÇÃO DOS CORPUS PELO ALFABETO FONÉTICO SAMPA

Atlas Linguístico do Centro-Oeste Potiguar

Carta fonética		Ocorrência	%
1	PRATELEIRA		
	[pahti"lera]	11	37,9%
	[prati"lera]	10	34,5%
	[pati"lera]	6	20,7%
	[prati"lejra]	2	6,9%
		29	
2	TELEVISÃO		
	[tEIEviza~w]	27	84,4%
	[televiza~w]	5	15,6%
		32	
3	CAIXA		
	[kajSa]	19	59,4%
	[kaSa]	13	40,6%
		32	
4	TESOURA		
	[ti"zora]	18	58,1%
	[te"zora]	13	41,9%
		31	
5	TRAVESSEIRO		
	[travi"seru]	29	93,5%
	[trabi"seru]	2	6,5%
		31	
6	LÂMPADA		
	["la~pada]	20	64,5%
	["la~pida]	9	29,0%
	["la~pa]	2	6,5%
		31	
7	TORNEIRA		
	[toh"nera]	30	93,8%
	[toh"nejra]	2	6,3%
		32	
8	FÓSFORO		
	[fOsfOru]	10	31,3%
	[fOsku]	9	28,1%
	[fOskuru]	6	18,8%
	[fOsfEru]	4	12,5%
	[fOfuru]	3	9,4%
		32	
9	PÓLVORA		
	["pOvura]	11	36,7%
	["pOva]	5	16,7%
	["pOvara]	4	13,3%
	["pOwvora]	3	10,0%
	["pOvra]	3	10,0%
	["pOvOra]	2	6,7%

Corpus AliPTG

Carta fonética		Ocorrência	%
27	PRATELEIRA		
	[pahti"lera]	4	17,4%
	[prati"lera]	11	47,8%
	[pati"lera]	7	30,4%
	[pahte"lejra]	1	4,3%
		23	
31	TELEVISÃO		
	[tEIEviza~w]	22	95,7%
	[televiza~w]	1	4,3%
		23	
7	CAIXA		
	[kajSa]	9	37,5%
	[kaSa]	15	62,5%
		24	
32	TESOURA		
	[ti"zora]	10	43,5%
	[te"zora]	12	52,2%
	[te"zowra]	1	4,3%
		23	
8	CANOA		
	[ka~"noa]	14	87,5%
	[ka"noa]	2	12,5%
		16	
19	LÂMPADA		
	["la~pada]	20	83,3%
	["la~pda]	3	12,5%
	["la~pa]	1	4,2%
		24	
11	COSTA		
	["koStas]	21	87,5%
	["kostas]	3	12,5%
		24	
14	FÓSFORO		
	["fOsfOru]	3	12,5%
	["fOsfu]	5	20,8%
	["fOSfuru]	1	4,2%
	["fOsfOru]	1	4,2%
	["fOfuru]	14	58,3%
		24	
3	BEIJAR		
	[bej"Za]	7	31,8%
	[be"Za]	15	68,2%
		22	
4	BONITO		
	[bu"nitu]	11	45,8%
		11	45,8%

	["pOvEra]	2	6,7%
		30	
10	ALMOÇO		
	[aw"mosu]	30	93,8%
	[a"mosu]	2	6,3%
		32	
11	GORDURA		
	[guh"dura]	22	71,0%
	[goh"dura]	9	29,0%
		31	
12	PENEIRA		
	[pe~"nera]	24	75,0%
	[pe"nera]	5	15,6%
	[pe~"nejra]	3	9,4%
		32	
13	COLHER		
	[ko"L0E]	14	43,8%
	[ku"L0E]	9	28,1%
	[kuj"L0E]	7	21,9%
	[kOj"L0E]	2	6,3%
		32	
14	GARFO		
	["gafu]	18	56,3%
	["gahfu]	7	21,9%
	["gawfu]	7	21,9%
		32	
15	FERVENDO		
	[fE"ve~nu]	8	25,8%
	[feh"ve~du]	7	22,6%
	[fEh"ve~nu]	6	19,4%
	[fE"ve~du]	3	9,7%
	[feh"ve~du]	3	9,7%
	[fE"ve~du]	2	6,5%
	[frE"ve~du]	2	6,5%
		31	
16	TÁBUA		
	["tawba]	20	62,5%
	["tabua]	10	31,3%
	["tawbua]	2	6,3%
		32	
17	CEBOLA		
	[se"bola]	22	68,8%
	[si"bola]	10	31,3%
		32	
18	TOMATE		
	[to"mati]	17	54,8%
	[to"mati]	4	12,9%
	[to~"mati]	4	12,9%
	[tO"matSi]	4	12,9%
	[tu"mati]	2	6,5%
		31	
19	ABÓBORA		
	[a"bObOra]	15	75,0%

	[bo"nituj]	2	8,3%
		24	
15	FUMAÇA		
	[fu"masa]	20	83,3%
	[fu~masa]	4	16,7%
		24	
16	GORDURA		
	[guh"dura]	20	83,3%
	[goh"dura]	4	16,7%
		24	
25	PENEIRA		
	[pe~"nera]	19	79,2%
	[pe"nera]	4	16,7%
	[pe~"nejra]	1	4,2%
		24	
10	COLHER		
	[ko"L0E]	3	15,0%
	[ku"IEh]	1	5,0%
	[kuj"IE]	14	70,0%
	[kOj"IE]	2	10,0%
	[ku"IE]	4	20,0%
		20	
18	JOELHO		
	[Zu"eL0u]	6	25,0%
	[Zu"eL0u]	18	75,0%
		24	
13	FERVENDO		
	[feh"ve~du]	13	56,5%
	[fEh"ve~nu]	6	26,1%
	[feh"ve~du]	2	8,7%
	[fE"ve~du]	2	8,7%
		23	
23	PEITO		
	[pejtSu]	12	52,2%
	[pejtu]	11	47,8%
		23	
9	CEBOLA		
	[se"bola]	18	75,0%
	[si"bola]	6	25,0%
		24	
33	TOMATE		
	[to"mati]	14	63,6%
	[tO"matSi]	1	4,5%
	[tu"mati]	7	31,8%
		22	
6	BRAGUINHA		
	[bra"g0iJ0a]	3	12,5%
	[bah"g0iJ0a]	4	16,7%
	[bEh"g0iJ0a]	5	20,8%

		32	
63	SALOBRA		
	[saˈloba]	16	76,2%
	[saˈlobra]	5	23,8%
		21	

APÊNDICE B – CÓDIGO FONTE DO POTIGRAFONE

1) Arquivo **Alfabeto.Foma**

```
clear
define ocl          [p | b | t | d | c | g];
define fri          [f | v | s | z | x | j];
define nas          [m | n];
define liq          [r | l];
define esp          [h | ç | q];
define vogal        [a | e | i | o | u];
define vant         [e | i | é | í | ê];
define vpos         [o | ô | ó | u | ú];
define vmed         [a | á | â];
define vacentuada   [á | â | é | í | ó | ú | ê | ô];
define vnasal       [ã | õ];
define gli          [i | u];
define C1           [ocl | fri | nas | liq | esp];
define C2           [ [ ocl | f | v ] [ liq ]-[d | v | l] ];
define Co           C1.r -esp;
define dig          [[c | l | n] h];
define nasalizacao [vogal | [vacentuada-[ á | é | ó ] ]];
define N1           [vogal | vacentuada] ;
define N2           [[vogal | é | ó] gli] | [gli vogal];
define N3           [vnasal [e | i | o]];
define N4           [u [vogal-u] [i | u]] | [u vnasal [o | e]];
define nucleo       [N1 | N2 | N3 | N4 | nasalizacao] ;
define rima         nucleo (Co);
```

2) Arquivo do **Beta-gerador.Foma**

```
clear
source Alfabeto.Foma
define silaba       [(C1 | C2) rima];
define gerador silaba+;
define separador    gerador @> ... "=" || _gerador;
define Rapaga["="]-> 0 || [c | l | n]_[h],
                    vogal_[vogal| vant]      ,
                    vnasal_[o | e s]          ;
define Rinsere [..]-> "=" ||_[ [ c | l | n ] h]      ,
                    [a | u]_[i z] .#.          ,
                    [e | i]_[a [r | l]] .#.      ,
                    [u | a]_[i r] .#.          ,
                    [vogal]_[e | o | a | [a | e] s | e m] .#. ,
                    i_i , u_u , a_a           ;
```

3) Arquivos das regras de reescrita

REGRAS DA TONICIDADE DAS PALAVRAS

```
clear
define acentuada [...] -> "" || [#. | "=" ]_(C1 | C2| dig) vacentuada ;
define paroxitona [...] -> "" || [#. | "="]_[[C1| C2| dig] vogal ] "=" silaba.#. ,
[#. | "="]_[(C1 | C2) vogal (Co) ] "=" [silaba | dig vogal].#.,
[#. | "="]_[[C1| C2| dig] vogal vogal (Co)] "=" silaba .#. ,
[#. | "="]_silaba "=" [C1 vogal n s] ;
define oxitona [...] -> "" || ["=" | .#. ]_(C1|C2 | dig ) vogal (vogal) [r | l | x | z | i s].#. ,
[#. | "=" ] _[(C1 | C2| dig) vnasal [o | e] (s)].#. ,
.#._(C1| C2| dig) vogal vogal (s | r) .#. ,
.#._(C1| C2| dig) i (m) .#. ;
define limpar [""]-> 0 || vacentuada (n | Co) "="_ ,
[#. | "="]_(C1|C2 | dig) vogal (vogal) (Co) "=" "" (C1|C2 | dig)
vogal (vogal) [r | l | x | z | i s].#. ,
[#. | "="]_(C1|C2 | dig) vogal (vogal) (Co) "=" "" (C1) vnasal [o
| e] (s).#. ,
[#. | "=" | s]_ [C1 """] ;
```

REGRAS DAS CONSOANTES

```
clear
define Rcodas h -> "0" || [#. ]_ , [c | l | n]_ ,,
r->h || _["=" | .#. ] , [#. """]_ ,,
l->w , nas->"~" || _[#. | "="] ,,
n->"~" || _s ,,
x -> K || _.#. ,,
x-> s || [#. e ]_ ;
define Roclusiva c -> k || _[vpos | vmed | liq | "="] ;
define Rfricativa s -> z || [vogal | vacentuada] ["=" ("")]_ [vogal | vnasal] ,,
c -> s , g -> Z || _[vant] ,,
ç -> s ,,
j -> Z ,,
c -> S || _"0" ,,
x -> z || [#. ]_ [[e "=" """] | ["" ê "="]]_ ,,
x -> S || [#. ]_ [#. """]_ ,
[vogal vogal] "=" ,
[["~" | m [e|é]] "=" ("")]_ ;
define Rliquida n -> J , l -> L || _"0" ;
define Rdiacritica s -> "0" || [s "=" ("")]_ ,,
r -> "0" || [h "=" ("")]_ ;
define Rcorrecao [""]-> [] || _"0" ;
define Rinclusao [...] -> ["""] || _[[s | h] "=" "0"] ;
```

REGRAS DO NÚCLEO SILÁBICO

```
clear
define Rditongo      u -> w || _["=" vogal] .#.
                    o -> w , [e | i] -> j || [C1 | "0"]_["="] [a | o | e] (s).#.
                    u->"0" || [g | q]_[vant]
                    u->w || [g | q]_[vmed| vpos]
                    q -> k
                    i -> j, u -> w || [vogal]_ , [é |ó] ("=") _
                    [ã]-> [a "~"], õ -> [o "~"]
                    o ->w, e -> j || [ã | õ]_(s).#.
                    ["="]-> "~" || [ m u i]_ t o .#.
                    [.] -> j || [e | ê | é] "~"_(s).#.
                    [.] -> w || [a "~"]_#.
define Rcorrecao1   [""]-> 0 || g u "="
                    ["="]->0 || g u _""
                    [.] -> "" || _g u "=" "" , "~" s_t;
define Rreducao     o->u, a->6, e->i || _(s) .#. ;
                    #e->i || v_ (""") s "=" 0;
define Racentuada   [á | â] ->a , é->E, ê->e, í->i, ó->O, ô -> o, ú->u;
```

REGRAS DOS PROCESSOS FONOLÓGICOS

```
clear
define Rnasalidade  ["="]-> "~" || [vogal]_ (""") [m | n | J];
define Rvozeamento s -> [z | h] || _["=" (""") m] ;
define Rharmonia    [s | x] ->S || vogal_ "=" (""") t ;
define Rpalatalizacao t -> [t | tS] || [e j "="] _u .#.
                    j -> [L | j] || [t e]_ "=" 6 .#. ;
define Rlenicao      L ->[j] [""] L || "=" _"0" ,
                    L ->[l | j | L] || ["""]_ ;
define Rapagamento h -> ["0"] || _.#.,
                    h -> [h | w | 0] || .#. [""] (C1) vogal]_ ,
                    d -> [n | d] || ["~" "="] _u .#. ;
define Rditongacao  [.] -> j || _z .#. ,
                    z ->s || _.#.
                    6->a || j "=" _ .#. ;
define Rhiato       w-> o || "" [g | r | n]_["=" 6] .#.;
define Rmonotongo   j->[j | 0] || [l]_ "=" [a] , [r] _ "=" [i | u] .#. , [a | e]_C1 , [e |
a]_ "=" ,
                    w -> [w | 0] || o_ , _ "=" 6 , .#. "" C1 O _ ;
define Rvocalismo   o -> [o | u] || _[[h "="] | ["~"]] "" , [k | b]_ "=" "" ,
                    e -> [e | i] || _["=" "" z] , .#._["~" | S] "=" , .#. [s]_ , .#.
[r]_[s |z],
                    e -> [e | E] || _[h "=" """] ;
define Ralternancia o -> O || .#. "0"_(s) ("=") "" (h | r) ;
define Repentese    [.] -> i || p_s , p_t
                    [.] -> [i | e] || p_n, d_["=" v o] ,
```

```

"~" -> m i || .#. a_[""] n E "=" ;
define Rajuste [""]-> 0 || _[s "=" "0" u].#. ,
[""] silaba_[s "=" "0"] ,
_[[k | m] vogal] [""] [h "=" "0" e w] ,,
["~"]-> 0 || .#. a_[""] m];
define AlterTonicaInicial e -> E || [""] (n) _ s "=" "0" [6] (s) .#. ,
.#. [""] [b | n | m | t | v | s | p] _ "=" [b | t | l] [i | 6] (s)
.#. ,,
o -> O || [""] (n) _ "=" [v] [6] (s) .#. ,
.#. [""] [t | k | b] _ (h) "=" [s | t | l] [i | 6] (s) .#. ,,
o -> O || [""] [v | s]_(w) "=" [t] [i | 6] (s) .#. ,,
o -> O || _[h | S] "=" [l | t | p | v | L "0" | "0" | s | g] u s ,
("=") ([""]) pr _ "=", .#. ([""]) [b] _ [h] "=";

```

4) Arquivo do **Potigrafone.Foma**

```

clear
source Betagerador.Foma
source Rtonicidade.Foma
source Rconsoantes.Foma
source Rnucleo.Foma
source Rfenomenos.Foma
source Ralternancia.Foma
define betagerador   separador   .o.
                    Rapaga       .o.
                    Rinsere      ;
define tonicidade   acentuada    .o.
                    paroxitona   .o.
                    oxitona.o.
                    limpar       ;
define consoante    Rcoda        .o.
                    Roclusiva    .o.
                    Rfricativa   .o.
                    Rliquida     .o.
                    Rdiacritica  .o.
                    Rcorrecao    .o.
                    Rinclusao    ;
define nucleo1 Rditongo .o.
                    Rcorrecao1   .o.
                    Rreducao     .o.
                    AlterTonicaInicial .o.
                    Racentuada   ;
define fenomeno     Rnasalidade  .o.
                    Rvozeamento .o.
                    Rharmonia    .o.
                    Rpalatalizacao.o.
                    Rlenicao       .o.
                    Rapagamento .o.

```

```

Rditongacao .o.
Rhiato .o.
Rmonotongo .o.
Rvocalismo .o.
Ralternancia .o.
Repentese .o.
Rajuste ;
define cancelando ["="] -> 0
["'"] -> 0 || _h "=" "0" [vogal].# ,
["'"] -> 0 || _[ " " " m e "~" t];
regex
betagerador .o.
tonicidade .o.
consoante .o.
nucleo1 .o.
fenomeno .o.
cancelando ;

```

APÊNDICE C – EXECUTAR O CONVERSOR POTIGRAFONE

Há três maneiras de executar o conversor Potigrafone, duas formas pelo terminal do Linux e pela web.

Pelo terminal do Linux faz necessário alguns requisitos básicos:

- Baixar o arquivo **Foma** no site do googlecode no endereço:
<https://code.google.com/p/Foma/downloads/list> ;
- Descompactar o arquivo do Foma em uma pasta.
- O arquivo do Foma contém: 2 pastas (include e lib), 3 arquivos executáveis (Foma, flookup, cgflookup) e 3 arquivos no formato .txt (Readme, copying e Changlog).

1. Execução por meio de comandos do Foma

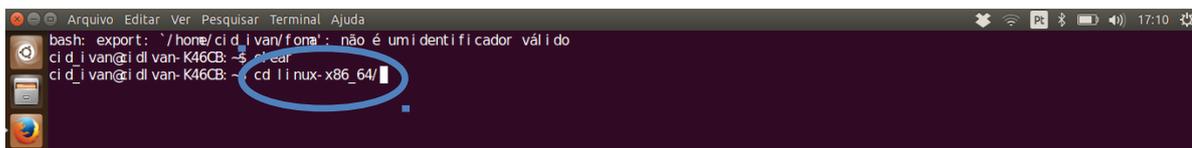
1º passo:

Na pasta do arquivo Foma, deve inserir todos os módulos que estão na pasta do Potigrafone.

2º passo:

Abrir a pasta que está os arquivos do Foma e do Potigrafone pelo terminal do Linux. Supondo a pasta está no diretório /home, executa-se os seguintes comandos:

cd PastaDescompactada

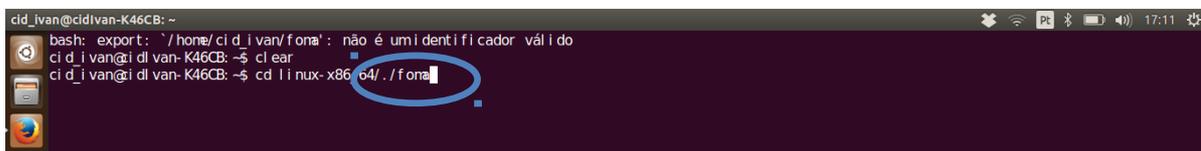


```
bash: export: `/home/cidivan/foma': não é um identificador válido
cidivan@cidivan-K46CB: ~$ clear
cidivan@cidivan-K46CB: ~$ cd /home/cidivan/foma/
```

3º passo:

Nessa pasta, acesso o **Foma** por meio do seguinte comando:

./Foma



```
cidivan@cidivan-K46CB: ~$ ./Foma
```

Após apertar a tecla **enter**: a seguinte informação:

```
Foma, version 0.9.17alpha
Copyright © 2008-2012 Mans Hulden
This is free software; see the source code for copying conditions.
There is ABSOLUTELY NO WARRANTY; for details, type "help license"
```

```
Type "help" to list all commands available.  
Type "help <topic>" or help "<operator>" for further help.
```

```
Foma[0]:
```

4º passo:

Onde se encontra o prompt, digite os seguintes seguinte comando:

- **source Potigrafone**
- **enter**
- O sistema compila todo o código na tela e apresenta a opção:

Foma [1]:

- onde você deve digitar o seguinte comando:

Foma [1]: down

Apply down> Nesse espaço, escreve uma palavra e aperte na tecla enter.

2. Execução por meio de flookup

1º passo:

Acessa a pasta onde está o flookup

2º passo:

digita o comando:

```
cid_ivan@cidIvan-K46CB:~/Foma$ ./flookup
```

- Após o enter aparece:

```
Usage: flookup [-h] [-a] [-i] [-s "separator"] [-w "wordseparator"] [-v] [-x] [-b] [-l <#/#k|  
#m|f>] [-S] [-P] [-A] <binary Foma file>  
cid_ivan@cidIvan-K46CB:~/Foma$
```

3º passo:

No prompt, digite os seguintes comandos:

```
echo "casa" | ./flookup -i Potigrafone.fst  
casa kaz6  
cid_ivan@cidIvan-K46CB:~/Foma$
```

4º passo:

Para quem quer executar um arquivo de texto.txt, basta executar o seguinte comando:

```
cat arquivoteste | ./flookup -i Potigrafone.fst  
torneira tohnejr6
```

```
torneira      tohner6
peneira              pe~nejr6
peneira              pe~ner6
teia    tej6
teia    te6
peixe  pejSi
peixe  peSi
vassoura  vas0owr6
vassoura  vas0or6
tesoura   tezowr6
tesoura   tezor6
caixa    kajS6
caixa    kaS6
cid_ivan@cidIvan-K46CB:~/Foma$
```

ANEXO A - CORRESPONDÊNCIA ENTRE OS SÍMBOLOS DO SAMPA COM OS SÍMBOLOS DO ALFABETO FONÉTICO INTERNACIONAL

SAMPA		IPA		Unicode		label and exemplification
symbol	ASCII			hex	dec.	
Vowels						
A	65	ɑ	script a	0251	593	open back unrounded, Cardinal 5, Eng. <i>start</i>
{	123	æ	ae ligature	00E6	230	near-open front unrounded, Eng. <i>trap</i>
6	54	ɐ	turned a	0250	592	open schwa, Ger. <i>besser</i>
Q	81	ɔ	turned script a	0252	594	open back rounded, Eng. <i>lot</i>
E	69	ɛ	epsilon	025B	603	open-mid front unrounded, C3, Fr. <i>même</i>
@	64	ə	turned e	0259	601	schwa, Eng. <i>banana</i>
3	51	ɜ	rev. epsilon	025C	604	long mid central, Eng. <i>nurse</i>
I	73	ɪ	small cap I	026A	618	lax close front unrounded, Eng. <i>kit</i>
O	79	ɔ	turned c	0254	596	open-mid back rounded, Eng. <i>thought</i>
2	50	ø	o-slash	00F8	248	close-mid front rounded, Fr. <i>deux</i>
9	57	œ	oe ligature	0153	339	open-mid front rounded, Fr. <i>neuf</i>
&	38	æ	s.c. OE lig.	0276	630	open front rounded
U	85	ʊ	upsilon	028A	650	lax close back rounded, Eng. <i>foot</i>
}	125	ʉ	barred u	0289	649	close central rounded, Swedish <i>sju</i>
V	86	ʌ	turned v	028C	652	open-mid back unrounded, Eng. <i>strut</i>
Y	89	ʏ	small cap Y	028F	655	lax [y], Ger. <i>hübsch</i>
Consonants						
B	66	β	beta	03B2	946	voiced bilabial fricative, Sp. <i>cabo</i>
C	67	ç	c-cedilla	00E7	231	voiceless palatal fricative, Ger. <i>ich</i>
D	68	ð	eth	00F0	240	voiced dental fricative, Eng. <i>then</i>
G	71	ɣ	gamma	0263	611	voiced velar fricative, Sp. <i>fuego</i>
L	76	ʎ	turned y	028E	654	palatal lateral, It. <i>famiglia</i>
J	74	ɲ	left-tail n	0272	626	palatal nasal, Sp. <i>año</i>
N	78	ŋ	eng	014B	331	velar nasal, Eng. <i>thing</i>
R	82	ʀ	inv. s.c. R	0281	641	vd. uvular fric. or trill, Fr. <i>roi</i>
S	83	ʃ	esh	0283	643	voiceless palatoalveolar fricative, Eng. <i>ship</i>
T	84	θ	theta	03B8	952	voiceless dental fricative, Eng. <i>thin</i>
H	72	ɥ	turned h	0265	613	labial-palatal semivowel, Fr. <i>hui</i>
Z	90	ʒ	ezh (yogh)	0292	658	vd. palatoalveolar fric., Eng. <i>measure</i>
?	63	ʔ	dotless ?	0294	660	glottal stop, Ger. <i>Verein</i> , also Danish <i>stød</i>
Length, stress and tone marks						
:	58	:	length	02D0	720	length mark

			mark			
"	34	'	vertical stroke	02C8	712	primary stress *
%	37	'	low vert. str.	02CC	716	secondary stress
`	96		(see note 1)			falling tone
'	39		(see note 1)			rising tone

Note 1: The SAMPA tone mark recommendations were based on the IPA as it was up to 1989-90. Since then, however, the IPA has changed its symbols for falling and rising tones. These SAMPA tone marks may now be considered obsolete, having in practice been superseded by the SAMPROSA proposals.

Diacritics (shown with another symbol as an example)

=n	60	◌̩	inf. stroke	0329	809	syllabic consonant, Eng. <i>garden</i> (see note 2)
O~	126	◌̃	sup. tilde	0303	771	nasalization, Fr. <i>bon</i>

Note 2: At the time SAMPA was established it was assumed that the syllabicity diacritic should **precede** the base character. More recently, ISO and Unicode have established that all diacritics should **follow** the base character, and this principle should be applied in future work.

Fonte: SAMPA - computer readable phonetic alphabet. disponível em: <http://www.phon.ucl.ac.uk/home/sampa/index.html>. Acesso em: agosto 2014.

ANEXO C - SÍMBOLOS FONÉTICOS USADOS PELO SISTEMA DE ESCRITA FONÉTICA SAMPA PARA O PORTUGUÊS

Tabela 2: Símbolos fonéticos usados pelo sistema de escrita fonética SAMPA para o português.

Símbolo	Palavra	Transcrição		
Oclusivas				
p	pai	paj		
b	barco	"barku		
t	tenho	'teju		
d	doce	"dos@		
k	com	ko~		
g	grande	"gr6nd@		
Fricativas				
f	falo	"falu		
v	verde	"verd@		
s	céu	sEw		
z	casa	"kaz6		
S	chapéu	S6"pEw		
Z	joia	"ZOj6		
Nasais				
m	mar	mar		
n	nada	"nad6		
J	vinho	"viJu		
Líquidas				
l	lanche	"l6nS@		
L	trabalho	tr6"baLu		
r	caro	"kru		
R	rua	"Ru6		

i	vinte	"vint@	lápiz	"lapiS
e	fazer	f6"zer		
E	belo	"bElu		
a	falo	"falu		
6	cama	"K6m6	madeira	m6"d6jr6
O	ontem	"Ont6~j~		
o	lobo	"lobu		
u	jus	ZuS	futuro	fu"turu
@	felizes	f@a'liz@s		
i~	fim	fi~		
e~	emprego	e~"pregu (ou em-)		
6~	irmã	ir"m6~		
o~	bom	bo~		
u~	um	u~		
aw	mau	maw etc.: iw, ew, Ew, (ow)		
aj	mais	majS etc.: ej, Ej, Oj, oj,		
6~j	têm	t6~j~ etc.: e~j~, o~j~, u~j~		

Outros símbolos	
"	Acento tônico primário (posto antes da sílaba tônica)
%	Acento tônico secundário (posto antes da sílaba tônica)
•	Separador silábico

Fonte: Seara, Nunes e Lazzarotto-Volcão (2015, p. 28).

ANEXO D - MATRIZ DE TRAÇOS DAS VOGAIS E CONSOANTES DO PB

(3) Matriz fonética

	p	b	t	d	k	g	tʃ	dʒ	f	v	s	z	ʃ	ʒ	h	m	n	ɲ	l	ʎ	r	i	e	ɛ	a	ɔ	o	u	ɪ	ə	ʊ
consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	
silábico	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	
soante	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
contínuo	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	
solt. retardada	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
nasal	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	
lateral	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	
anterior	+	+	+	+	-	-	-	-	+	+	+	+	-	-	-	+	+	-	+	-	+	-	-	-	-	-	-	-	-	-	
coronal	-	-	+	+	-	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	
alto	-	-	-	-	+	+	-	-	-	-	-	-	+	+	+	-	-	+	-	+	-	+	-	-	-	-	-	+	+	+	
recuado	-	-	-	-	+	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	
arredondado	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+
baixo	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	
vozeado	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
tenso	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-

Fonte: Silva (2014, p. 195).

ANEXO E - ALGORITMO POTIGRAFONE

```
variaveis
    elemento      : caractere;
    simbolos      : listaCaractere;
    silaba        : listaCaractere;
    palavra       : listaCaractere;
    transcricao   : listaCaractere;

função existeElemento(elemento: caractere): inteiro;
função verificaSimbolos(): inteiro;
função transcrevePalavra(palavra :listaCaractere ): listaCaractere;
função regraTranscricao(sil :listaCaractere ): listaCaractere;
função principal;

/*----- Verifica se os símbolos pertencem ao conjunto de símbolos
de entrada -----*/
função existeElemento(elemento: caractere): inteiro
var
    aux      :inteiro;
    posicao  :inteiro;
inicio
    aux = 1;
    posicao = 0;
    enquanto posicao < tamanho(simbolos) faça
    inicio
        if( elemento==simbolos[posicao] ) então
            inicio
                aux = 1;
            fim if
        posicao++;
    fim enquanto
retorna aux;
fimfunção
```

```

/*----- Verifica se todos os caracteres da palavra são válidos
-----*/

função verificaSimbolos(): inteiro
var
    aux    :inteiro;
    posicao :inteiro;
inicio
    aux = 1;
    posicao = 0;
    enquanto posicao < tamanho(palavra) faça
    inicio
        if( existeElemento(palavra[i])=0 ) então
            inicio
                aux = 0;
                retorna aux;
            fim if
        fim enquanto
    retorna aux;
fimfunção

/*----- Implementa os passos da transcricao -----*/
função transcrevePalavra(palavra :listaCaractere ): listaCaractere
var
    auxTranscricao    : listaCaractere;
    silaba             : conjunto de listaCaractere;
    posicao :inteiro;
inicio
    silaba = divisaoSilabica (palavra);
    posicao = 0;
    enquanto posicao < numeroElementos(silaba) faça
    inicio

```

```

                                auxTranscricao      =      auxTranscricao      +
regraTranscricao(silaba[posicao]);
                                fim enquanto
                                retorna auxTranscricao;
fimfunção

                                /*----- Implementa os passos da transcricao -----*/
função regraTranscricao(sil :listaCaractere ): listaCaractere
var
                                auxTranscricao      : listaCaractere;
início
                                /* implementacoes das regras de transcricao p/diversos
casos */
                                retorna auxTranscricao;
fimfunção

/*----- Função Principal -----*/
principal()
início
                                escreva("Digite a palavra");
                                leia (palavra);
                                if (verificaSimbolos(palavra) ) então
                                início
                                        transcricao = transcrevePalavra();
                                fim
                                senão
                                início
                                        imprime "palavras com caracteres inválidos"
                                fim
                                imprime transcricao;
fim
fim

```