

# Geração Automática de Variantes de Léxicos do Português Brasileiro para Sistemas de Reconhecimento de Fala

Izabel C. Seara, Fernando S. Pacheco, Rui Seara Jr., Sandra G. Kafka, Simone Klein e Rui Seara

**Resumo**—Neste trabalho, propomos uma estratégia para geração automática de variantes de pronúncia do léxico de um sistema de reconhecimento de fala para o português brasileiro, independente de locutor. As variações descritas referem-se a processos fonológicos próprios das línguas naturais, tais como inserção ou apagamento de segmentos, assimilação segmental e estruturação silábica. Foram consideradas as variações intra- e inter-palavras. Comparamos o número e os tipos de variantes apresentadas nas transcrições obtidas automaticamente com aquelas realizadas por um *expert* linguísta e verificamos que a concordância entre elas ficou muito próxima de 100%.

**Palavras-Chave**—Processos fonológicos, variantes de pronúncia, geração automática de transcrições, sistema de reconhecimento de fala.

**Abstract**—In this paper, we propose a strategy for automatically generating a lexicon of pronunciation variants, independent of speaker, to be used in a speech recognition system for Brazilian Portuguese. The variants described are referred to the phonological processes inherent in natural languages, such as insertion or deletion of segments, segmental assimilation, and syllabic structure. Intra- and inter-word variants have been taken into account. The number and the types of variants obtained by automatically generated transcriptions are compared with the ones performed by an expert linguist. We have verified a very good agreement between both results (almost 100%).

**Index Terms**—Phonological processes, pronunciation variants, automatically generated transcriptions, speech recognition system.

## I. INTRODUÇÃO

Os sistemas de reconhecimento de fala que buscam independência de tarefa e, principalmente, de locutor devem identificar as fontes de variação de pronúncia que reduzem o desempenho do reconhecimento [1]-[5]. Essas perdas se devem às alterações que sofrem os segmentos justapostos,

Izabel C. Seara, Fernando S. Pacheco, Rui Seara Jr., Sandra G. Kafka, Simone Klein e Rui Seara, LINSE – Laboratório de Circuitos e Processamento de Sinais, Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, SC, E-mails: {izabels, fernando, ruijr, kafka, klein, seara}@linse.ufsc.br.

Este trabalho foi parcialmente financiado pela Dígito Tecnologia Ltda. e pelo CNPq.

constituindo os chamados processos fonológicos, que devem ser considerados por tais sistemas. As variações acontecem entre falantes de diferentes regiões geográficas, classes sociais ou mesmo faixas etárias [6]. Sistemas que apresentam um único modelo de pronúncia não são muito eficazes para a tarefa de reconhecimento de fala. Assim, a geração de variantes de pronúncia para as palavras que compõem o léxico melhora significativamente o desempenho do sistema. Além do mais, a identificação dos processos que envolvem tais variações também possibilita a geração automática, permitindo seu emprego com maior independência do *corpus* original e da aplicação envolvida, levando em conta o maior número possível de variações que, no entanto, não acarretem um aumento da perplexidade de busca do sistema [7]. Ou seja, deve haver um compromisso entre número de variantes transcritas e taxa de desempenho.

Para apresentarmos de forma mais clara a integração dos processos de variação de pronúncia na geração automática de variantes, dividimos este estudo em 5 seções, como seguem. A Seção II apresenta as estratégias empregadas para se obter as transcrições elaboradas pela máquina. Na Seção III, apresentamos alguns exemplos dos processos fonológicos em estudo e como são elaboradas as regras que realizam a transcrição automática. Na Seção IV, aplicamos essas regras para a obtenção automática das variantes de pronúncia de 16.000 vocábulos e verificamos sua concordância com as transcrições já realizadas por um *expert* humano. Finalmente, na Seção V, apresentamos as conclusões relacionadas à obtenção automática das variantes, sugerindo os próximos passos para maximizar o desempenho do reconhecedor.

## II. ESTRATÉGIAS PARA A AVALIAÇÃO E AUTOMATIZAÇÃO DAS VARIANTES

Sistemas de reconhecimento de fala que se baseiam em uma única transcrição apresentam um desempenho muito pobre, necessitando-se, dessa forma, múltiplas transcrições das palavras contidas no léxico do sistema.

Algumas estruturas utilizam esquemas *data-driven* para derivar suas variantes de pronúncia de palavras [3]. Nesses esquemas, as regras não são modeladas explicitamente, mas sim, a partir de um teste de treinamento. Por exemplo, o apagamento de um fonema pode ser modelado através da

eliminação de quadros relacionados ao fonema não realizado. A vantagem das transcrições baseadas em regras, que apresentam processos lingüísticos menos dependentes de contexto ou tarefas, é que elas podem ser aplicadas a um novo léxico contendo palavras que ainda não ocorreram nos dados, possibilitando então uma maior adequação a tarefas diferentes.

Para que se possa estabelecer quais variações de pronúncia são pertinentes a uma dada língua, usando a transcrição por regras, deve-se observar os processos fonológicos presentes naqueles contextos e já esperados nas línguas naturais<sup>1</sup> para criar regras de reescrita desses segmentos fônicos.

Processos fonológicos advêm da justaposição de diferentes segmentos fonéticos, podendo ocorrer dentro de uma mesma palavra ou entre palavras subseqüentes. Na fala espontânea, os processos fonológicos, catalogados até então para o português brasileiro, muitas vezes não conseguem explicitar todas as variações que correspondem à realidade deste estilo de fala, já que a velocidade de elocução modifica vários parâmetros, inclusive apagando por completo sílabas inteiras. Essas sílabas são virtualmente recriadas pelo ouvinte, pois, em sua representação cognitiva de fala (sistema fonológico), tais segmentos ou sílabas devem estar presentes para que se possa, através do significante, associar-lhe o significado. Esse é um problema bastante complexo para os sistemas de reconhecimento de fala, podendo levar a uma considerável redução de seu desempenho [1]-[5].

Para se estabelecer o que seria variação, deve-se primeiramente considerar uma transcrição de referência sobre a qual serão introduzidas as variantes. Neste estudo, estabelecemos, como transcrições de referência, aquelas feitas segundo regras de correspondência grafema-fonema do português brasileiro, já empregadas, por exemplo, em sistemas de conversão texto-fala [9] e aquelas inseridas manualmente em um dicionário de exceções, em geral referentes a nomes próprios e estrangeiros. Essa transcrição de referência apresenta o modo de falar característico do nível culto da língua, no qual se teria uma elocução em velocidade mais lenta, isto é, não apareceriam apagamentos ou inserções de segmentos, característicos da maior velocidade de fala e das situações mais informais [3]. Na Tabela 1, mostramos alguns exemplos<sup>2</sup>.

TABELA 1  
TRANSCRIÇÃO DE REFERÊNCIA E RESPECTIVAS VARIANTES)

Transcrição de referência	Possíveis variações
Pouco [ 'powku ]	[ 'poku ]
Advogado [ advo 'gadu ]	[ adivo 'gadu ] [ adzivo 'gadu ]

<sup>1</sup> Línguas naturais são específicas da espécie humana em sua totalidade; instrumentos de comunicação e expressão que repousam sobre propriedades universais próprias de toda linguagem humana [8].

<sup>2</sup> Neste artigo, todas as transcrições foram feitas com símbolos fonéticos do *International Phonetic Alphabet* (IPA).

Para avaliar o desempenho de nosso transcritor automático, comparamos as múltiplas transcrições por ele efetuadas com aquelas elaboradas por um lingüista *expert* para 16.000 vocábulos, pertencentes a um léxico de um sistema de reconhecimento de fala. Assim, consideraremos que nosso sistema gerará automaticamente um conjunto de variantes de novas palavras inseridas no léxico, não necessitando recorrer a lingüistas, quando o percentual de concordância entre as transcrições do *expert* e as geradas automaticamente estiver muito próximo de 100%. Em nosso caso, essa concordância é verificada quanto ao número de variantes e na consideração ou não do processo fonológico em questão.

Apesar de termos levado em conta todos os processos fonológicos envolvidos na elocução de palavras do português brasileiro, isto é, os relacionados à estruturação silábica, assimilação, e enfraquecimento ou fortalecimento; para esta discussão, enfocaremos somente alguns exemplos desses processos. A partir das transcrições de referência, sobreposamos as regras relacionadas aos processos fonológicos do português brasileiro, obtendo assim as variações de pronúncia. Técnica semelhante foi usada em [3].

### III. PROCESSOS FONOLÓGICOS DO PORTUGUÊS BRASILEIRO

Os sistemas lingüísticos são formalizados através da organização dos sons, segundo sua estrutura e funcionamento, sendo essa sistematização objeto de estudo da fonologia [10]-[12]. Para podermos entender melhor esses conceitos fonológicos, temos que inicialmente entender de que forma os sons são classificados, isto é, como são definidas as classes fonéticas. Essa definição não é aleatória e vê o fonema como um conjunto de propriedades, através das quais se pode classificar os sons segundo as relações explicitadas por essas mesmas propriedades ou traços. Essa teoria, chamada de fonologia de traços, considera o fonema formado por um conjunto de traços binários de valores distintivos. Por exemplo, o que distingue as consoantes /k/ e /g/ é o traço binário [±vozeado], pois esses dois fonemas consonantais apresentam propriedades semelhantes, tais como: [-contínuo], [-soante], [-anterior] e [-coronal], distinguindo-se apenas por /k/ ser [-vozeado] e /g/, [+vozeado].

Os processos fonológicos consistem então em uma seqüência de regras que se aplicam a cadeias de segmentos em ordem linear. Cada regra age sobre uma cadeia modificada pelas regras anteriores. Essas regras operam em termos de traços binários (citados anteriormente) e adicionam, anulam ou permutam os segmentos, ou mudam os valores de seus traços distintivos. Elas explicitam processos dinâmicos próprios das línguas naturais [10]-[12], que podem ser agrupados em assimilatórios, estruturação silábica, enfraquecimento ou fortalecimento.

Para a nossa pesquisa, consideramos os processos fonológicos aparentes em estruturas internas às palavras e entre palavras. Esses processos vão compor as variantes de

pronúncia acessadas através de regras de reescrita dependentes de contexto, baseando-se nos processos fonológicos apresentados da seguinte forma:

$$A \boxed{B} C \rightarrow D, \quad (1)$$

onde A e C podem representar fonemas ou conjunto de fonemas<sup>3</sup>, dependendo do alcance da regra, ou limites de palavras (# representa início e final de palavra); B e D representam fonemas. Interpreta-se essa regra da seguinte forma: se um fonema B tem, como contexto à esquerda, o fonema ou o conjunto de fonemas A e, à direita, o fonema ou o conjunto de fonemas C, ele pode ser pronunciado como D.

Descrevemos também fenômenos que ocorrem na coarticulação entre palavras, uma vez que consideramos os limites de palavras como um dos contextos, podendo-se estabelecer as variantes que estariam presentes devido à palavra antecedente ou subsequente.

As regras criadas são aplicadas e reaplicadas a diferentes partes de uma pronúncia de referência e, ao final, são eliminadas as transcrições idênticas.

#### A. Processo de Assimilação

Um dos processos mais comuns e naturais de mudança é a assimilação. O processo assimilatório faz com que um segmento assumira os traços distintivos de um segmento adjacente. Assim, uma consoante pode tomar os traços distintivos de uma vogal; uma vogal, os traços distintivos de uma consoante ou ainda uma vogal pode afetar outra. Esse último tipo de assimilação é o que chamamos de harmonização vocálica. Nesse processo, a vogal de uma sílaba torna-se mais semelhante à vogal de uma outra.

No português brasileiro, tem-se harmonização vocálica em vogais médias átonas, [e] e [o], que tornam-se [i] e [u], respectivamente, quando, na sílaba seguinte, houver uma vogal alta ([i] ou [u]).

#### Exemplo 1:

“menino” /me'nino/ → [me'ninu]  
→ [mi'ninu]

#### Regra de reescrita do Exemplo 1<sup>4</sup>:

$$\$CA \boxed{e} \$C \$VA \$C \rightarrow i$$

<sup>3</sup> Quando o contexto se refere à classe de fonemas, ele é iniciado pelo símbolo \$.

<sup>4</sup> \$CA corresponde ao conjunto de fonemas diante dos quais comumente a vogal média se harmoniza com a vogal alta [mŋkɔpfs]; \$C corresponde ao conjunto de todas as consoantes do português brasileiro; \$VA, ao conjunto das vogais altas em posição tônica [i u].

Além do Exemplo 1, temos ainda como processos assimilatórios:

- vozeamento de fricativas em coda silábica seguida por consoante vozeada: “mesmo” [ 'mezmu ];
- palatização das consoantes oclusivas alveolares, transformando-se em africadas: “sete” [ 'setʃɪ ];
- vocalização ou velarização da lateral em final de sílaba ou palavra: “sal” [ 'saw ] [ 'saɫ ];
- nasalização assimilatória, principalmente em sílabas tônicas: “cama” [ 'kẽmẽ ], dentre outros.

#### B. Processo de Fortalecimento

Um processo que envolve o fortalecimento de segmentos fônicos é o fenômeno da ditongação que combina segmentos, aumentando o contraste entre vogais acentuadas e não acentuadas [11]. No português brasileiro, as sílabas tônicas de vocábulos oxítonos e monossílabos tônicos terminados em “e” ou “a” seguidas por “s” ditongam-se:

#### Exemplo 2:

“Através” /atra'ves/ → [atra'ves]  
→ [atra'veys]  
“Rapaz” /ra'pas/ → [ra'pas]  
→ [ra'pays]

#### Regras de reescrita do Exemplo 2:

$$\boxed{e} s \# \rightarrow ' \epsilon y \\ \rightarrow ' \epsilon \text{ɪ} \\ \boxed{a} s \# \rightarrow ' a y \\ \rightarrow ' a \text{ɪ}$$

#### C. Processo de Enfraquecimento

O processo de enfraquecimento ocorre com vogais finais de palavras ou semivogais, que se apresentam como fracas e sofrem apócope, isto é, perda da vogal final ou os grupos vocálicos (ditongos) reduzem-se a uma única vogal, sendo a nova estrutura silábica mais simples. Por exemplo, sílabas do tipo C<sub>1</sub>VC<sub>2</sub>C<sub>3</sub> e C<sub>1</sub>VC<sub>2</sub>, tônicas ou átonas, nas quais, na coda silábica<sup>5</sup>, C<sub>2</sub> constitui-se de semivogais e C<sub>3</sub> de consoantes fricativas<sup>6</sup>, podem passar para C<sub>1</sub>VC<sub>3</sub> e C<sub>1</sub>V, respectivamente, conforme o Exemplo 3. É o que chamamos de monotongação.

#### Exemplo 3:

“dois” / 'doys/ → [ 'doys ]  
→ [ 'dos ]  
“abaixar” /abay'ʃar/ → [abay'ʃar]  
→ [aba'ʃar]

<sup>5</sup> Coda silábica é a última posição na sílaba, geralmente ocupada por consoantes e semivogais.

<sup>6</sup> As fricativas, neste caso, são [s z ʃ ʒ].

**Regras de reescrita do Exemplo 3:**

$$\begin{array}{l}
 \$C \boxed{\text{'oy}} s \# \rightarrow \text{'oI} \\
 \rightarrow \text{'o} \\
 \\
 \$C \boxed{\text{ay}} \rightarrow aI \\
 \rightarrow a
 \end{array}$$

**D. Estruturação Silábica**

Os processos referentes à estruturação silábica são enfatizados por [13], apresentando a organização silábica da língua falada como ponto primordial para a consideração da variação de pronúncia, uma vez que freqüentemente a fala difere da representação fonológica canônica (de referência). Segundo [13], segmentos fonéticos são constantemente omitidos (28% deles são codas consonantais) ou transformados em outros (35% dessas transformações ocorrem com segmentos vocálicos).

Estes apagamentos e substituições, se observados pelo ponto de vista da sílaba, são mais sistemáticos do que se forem analisados do ponto de vista do segmento fonético ou fonológico. Vê-se também que a entidade silábica não é sempre claramente demarcada em termos de constituintes. Geralmente um constituinte pode se estender sobre outros dois segmentos fonéticos, resultando em um compartilhamento consonantal, com conseqüente ressilabificação em que, por exemplo, a coda da sílaba inicial se torna o *onset*<sup>7</sup> da sílaba seguinte.

**Exemplo 4:**

“Ginástica”/ʒi'nastika/ → [ʒi'naʃikɐ].

**Regra de reescrita do Exemplo 4<sup>8</sup>:**

$$\begin{array}{l}
 \$V \boxed{\text{st}} i \rightarrow st\text{ʃ} \\
 \rightarrow \text{ʃ}t\text{ʃ} \\
 \rightarrow \text{ʃ}t \\
 \rightarrow \text{ʃ}
 \end{array}$$

Em fala espontânea, a palavra “ginástica”, muitas vezes, é pronunciada como [ʒi'naʃikɐ], pronúncia na qual o “s” da coda da sílaba “nas”, ou seja, o segmento [ʃ], transforma-se também no *onset* da sílaba seguinte “ti”, passando de [nasti] para [naʃi]. Esse processo só ocorre quando a coda é constituída pela fricativa palatal [ʃ].

São ainda exemplos de estruturação silábica:

<sup>7</sup> *Onset* silábico é a primeira posição na sílaba, em geral, ocupada por consoantes. Núcleo silábico é a posição relacionada à vogal. Em português, apenas as vogais podem formar o núcleo de uma sílaba.

<sup>8</sup> \$V corresponde ao conjunto de todas as vogais do português brasileiro; [tʃ] corresponde à consoante africada comum no dialeto carioca quando pronuncia a palavra “tia” e [ʃ] corresponde ao som de “ch” em português.

i) *Onset* silábicos que são preservados, aproximando-se da forma canônica (CV), principalmente *onset* complexos com ausência de coda silábica (como CCV).

Exemplo: “quatro” [ 'kwatru ][ 'kwatu ].

ii) Os elementos de coda silábica são com freqüência apagados ou transformados em um segmento homorgânico ao do *onset* da sílaba seguinte. Em geral, o percentual de sílabas CV é bastante alto, visto que se somam a ele o percentual de transformações das sílabas CVC em CV, através do processo de apagamento da coda silábica. Em [14], constata-se que a freqüência relativa do tipo silábico CV em um *corpus* de fala semi-espontânea é de 58,54%.

Exemplo: “Os moços” → [ uz 'mosu ].

iii) O núcleo silábico, que seria composto, em português brasileiro, por vogal, também pode se desviar da forma canônica, mas não com a freqüência da coda silábica. A probabilidade do núcleo ser pronunciado de forma canônica é mais alta quando o *onset* é também articulado no modo padrão.

Exemplo: “Banho” [ 'bɛ̃n ].

#### IV. AVALIAÇÃO DA TRANSCRIÇÃO OBTIDA AUTOMATICAMENTE VERSUS A DE UM EXPERT

Após o levantamento de todas as possibilidades de variação de pronúncia com os processos mais recorrentes apresentados na literatura para o português brasileiro, elaboramos 320 regras que permitiram a construção de grupos de variantes para as palavras do léxico.

Inicialmente, geramos automaticamente as variantes de 500 palavras, retiradas aleatoriamente do léxico, para ajustar as regras implementadas. As variantes obtidas foram avaliadas, levando-se em conta as transcrições que um *expert* linguísta consideraria como possibilidades de variação. Depois, a fim de validarmos as regras de geração de variantes, foram analisadas 16.000 palavras que se mostraram bastante consistentes com as variantes que seriam estabelecidas manualmente (*expert*).

Assim, foram gerados autômatos que criavam as diferentes variantes de pronúncia de uma mesma palavra. O grafo apresentado na Fig. 1 mostra um exemplo do processo de geração automática de variantes. As Tabelas 2 e 3 apresentam exemplos de variantes e dos processos envolvidos para a obtenção dessas variantes de pronúncia.

Das 16.000 palavras testadas, foram obtidas 152.000 variantes, aproximadamente 10 variantes por vocábulo. A análise das transcrições constatou que 150 variantes não reproduziram variações pertinentes ao português brasileiro, ou seja, não correspondiam a uma das possibilidades que seriam transcritas por um *expert* linguísta. Isso equivale a um erro de apenas 0,1%. A concordância entre as transcrições de um *expert* e a obtida automaticamente pôde ser observada com respeito ao número e tipos de variantes de pronúncia selecionadas.

No entanto, nas situações discordantes, observamos que o sistema de geração automática é mais robusto do que o sistema manual, uma vez que as combinações de variantes determinadas pelos processos fonológicos são incluídas integralmente. No caso de um *expert*, poderia haver falhas, pois as transcrições sofrem variações intra-subjetivas. Assim, o método de criação de variantes a partir de regras se mostrou tão eficiente quanto o do transcritor humano na captura das variações de pronúncia do português brasileiro.

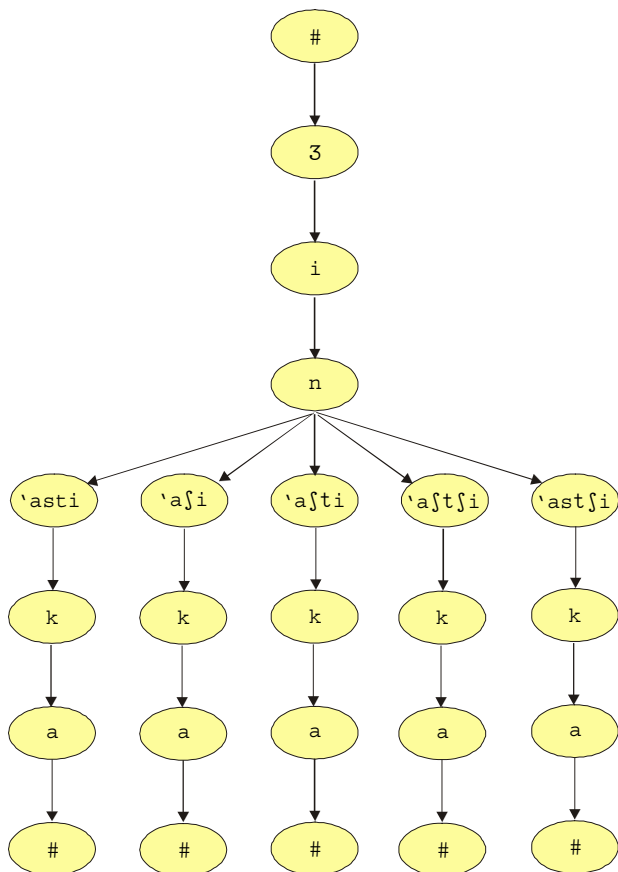


Fig.1. Grafo orientado da transcrição da palavra “ginástica”

TABELA 2

TRANSCRIÇÕES DO *EXPERT* E AS GERADAS AUTOMATICAMENTE DA PALAVRA “ABAIXAR” COM OS PROCESSOS UTILIZADOS PARA ESTABELECIMENTO DAS VARIANTES DE PRONÚNCIA

PROCESSOS	
[abay'ʃar]→[abay'ʃar]	Transcrição de Referência
[abay'ʃar]→[abay'ʃar]	Estruturação Silábica
[abay'ʃar]→[abai'ʃar]	Substituição
[abay'ʃar]→[abai'ʃar]	Estruturação Silábica
[abay'ʃar]→[aba'ʃar]	Enfraquecimento
[abay'ʃar]→[abay'ʃa]	Enfraquecimento
[abay'ʃar]→[aba'ʃa]	Enfraquecimento
[abay'ʃar]→[aba'ʃar]	Enfraquecimento
[abay'ʃar]→[abai'ʃa]	Substituição/Enfraquecimento

TABELA 3

TRANSCRIÇÕES DO *EXPERT* E AS GERADAS AUTOMATICAMENTE DA PALAVRA “GINÁSTICA” COM OS PROCESSOS UTILIZADOS PARA ESTABELECIMENTO DAS VARIANTES DE PRONÚNCIA

PROCESSOS	
[zi'nastike]→[zi'nastike]	Transcrição de Referência
[zi'nastike]→[zi'naʃtike]	Assimilação
[zi'nastike]→[zi'nastʃike]	Assimilação
[zi'nastike]→[zi'naʃtʃike]	Assimilação
[zi'nastike]→[zi'naʃike]	Estruturação Silábica

V. CONCLUSÕES

Nesta pesquisa, apresentamos uma estratégia para a obtenção automática de variantes de pronúncia para um léxico de um sistema de reconhecimento de fala do português brasileiro. Esse método baseia-se em um conjunto de regras de restrições que determinam, a partir de uma transcrição de referência, as possíveis variantes de pronúncia de uma palavra, focalizando, para isso, os processos fonológicos correntes encontrados no português brasileiro. As múltiplas pronúncias obtidas automaticamente substituem com eficiência aquelas que seriam inseridas por um *expert* linguísta.

O próximo passo deste estudo está em andamento e constitui-se no método de seleção automática de variantes (ou subconjunto de variantes), já que o custo computacional em sistemas de reconhecimento de fala é muito alto quando utilizamos todo o conjunto de variantes geradas. Assim, queremos maximizar o desempenho do reconhecedor, através da otimização do tamanho do subconjunto de variantes, levando-se em consideração a similaridade de contextos, mantendo ainda um bom nível de acertos.

AGRADECIMENTOS

Agradecemos à empresa de telecomunicações DÍGITRO Tecnologia Ltda. e ao CNPq pelo suporte financeiro dado a esta pesquisa.

REFERÊNCIAS

- [1] J. Kessens, M. Wester, and H. Strik. “Using Dutch Phonological Rules to Model Pronunciation Variation in ASR.” *Phonus 5: Proceedings of the Workshop on Phonetics and Phonology in ASR*. Saarbrücken: Institute of Phonetics, University of the Saarland, 2000, pp. 105-116.
- [2] O. Scharenborg and L. Boves. “Pronunciation Variation Modelling in a Model of Human Word Recognition” *Proceedings of Workshop on Pronunciation Modeling and Lexicon Adaptation*, Estes Park, USA, pp. 65-70.
- [3] N. Cremelie and J. P. Martens. “In Search of better pronunciation models for speech recognition”. *Speech Communication*, 29, pp. 115-136, 1999.
- [4] M. Wester, J. M. Kessens, and H. Strik. “Improving the Performance of a Dutch CSR by Modeling Pronunciation Variation.” *Proceedings of the workshop Modeling Pronunciation Variation for Automatic Speech Recognition*. Kerberade, 1998, pp. 145-150.
- [5] H. Strik and C. Cucchiari. “Modeling Pronunciation Variation for ASR: A Survey of the literature.” *Speech Communication*, 29, 1999, pp. 225-246.

- [6] M. Bagno. *A língua de Eulália: novela sociolinguística*. São Paulo: Contexto, 1997.
- [7] C. A. Ynoguti e F. Violaro. "Sobre a importância da transcrição fonética em sistemas de reconhecimento de fala". *Revista da Sociedade Brasileira de Telecomunicações*, vol. 15, nº 1, Junho 2000.
- [8] J. Dubois *et al.* *Dicionário de lingüística*. São Paulo: Cultrix, 1973.
- [9] I. C. Seara, S. G. Kafka, S. Klein e R. Seara. "Alternância vocálica das formas verbais e nominais do português brasileiro para aplicação em conversão texto-fala". *Revista da Sociedade Brasileira de Telecomunicações*, Rio de Janeiro, vol. 17, nº 1, pp. 79-85, 2002.
- [10] T. C. Silva. *Fonética e Fonologia do Português: roteiro de estudos e guia de exercício*. São Paulo: Contexto, 1999.
- [11] G. L. Istre. *Fonologia transformacional e natural: uma introdução crítica*. Ensaios de Lingüística da UFSC, Núcleos de Estudos Lingüísticos, Florianópolis, SC, 1980.
- [12] S. A. Shane. *Fonologia gerativa*. Tradução de Alzira S. da Rocha *et al.* Rio de Janeiro: Zahar Editores, 1975.
- [13] S. Greenberg. "Speaking in shorthand – A syllable-centric perspective for understanding pronunciation variation" *Speech Communication*, 29, pp. 159-176.
- [14] I. C. Seara. "Estudo estatístico dos fonemas do português brasileiro falado na capital de Santa Catarina para elaboração de frases foneticamente balanceadas". *Dissertação de Mestrado*. UFSC, Março 1994.