

CONSIDERAÇÕES SOBRE OS PROBLEMAS DE ALTERNÂNCIA VOCÁLICA DAS FORMAS VERBAIS DO PORTUGUÊS FALADO NO BRASIL PARA APLICAÇÃO EM UM SISTEMA DE CONVERSÃO TEXTO-FALA

IZABEL C. SEARA, SANDRA G. KAFKA, SIMONE KLEIN E RUI SEARA

LINSE: Circuitos e Processamento de Sinais
Departamento de Engenharia Elétrica
Universidade Federal de Santa Catarina
Campus Universitário, 88040-900 – Florianópolis – SC – Brasil
Tel: (0xx48) 331-9504, Fax: (0xx48) 331-9091, seara@linse.ufsc.br

RESUMO

Neste trabalho, propomos a criação de mecanismos para a consideração da alternância vocálica em verbos e substantivos para a aplicação em sistemas de conversão texto-fala. A resolução deste problema se baseia, fundamentalmente, em uma identificação adequada das classes gramaticais verbo e substantivo. Assim, apresentaremos uma visão geral das diferentes fases elaboradas para um *parser* morfo-sintático, bem como as principais técnicas de geração de alternância vocálica para situações gerais e para outras bastante restritas.

1. INTRODUÇÃO

Atualmente já existem diversas abordagens de sistemas de conversão texto-fala para o português falado no Brasil (doravante PB) [1-6]. As técnicas utilizadas nos diferentes estágios desses sistemas são muitas vezes distintas, mas seu objetivo final é a pronúncia correta da representação do texto de entrada com a máxima inteligibilidade e naturalidade na fala sintetizada. No entanto, um problema ainda não totalmente resolvido para o PB [1-4], no que concerne à pronúncia correta, é a alternância vocálica das formas verbais e nominais.

Para o PB, existem duas transcrições para as letras “e” e “o” não acentuadas graficamente no interior de palavras, podendo ser abertas (é, ó) ou fechadas (ê, ô). As regras ortográfico-fonéticas deram conta da maior parte delas. Porém, restaram as formas homógrafas não homófonas (mesma grafia com pronúncia diferenciada) cuja abertura e fechamento vocálicos distinguem verbos e substantivos. Por exemplo, em: *Eu olho no olho da menina* (formas homógrafas não homófonas), a primeira vogal “o” da forma verbal “olho” é aberta [ɔ], enquanto, na forma nominal, essa mesma vogal é fechada [o].

Além das homófonas, ainda temos problemas de abertura vocálica na 1ª, 2ª e 3ª pessoas do singular e 3ª do plural em

algumas conjugações verbais, para as quais as regras ortográfico-fonéticas não são satisfatórias, gerando vogais abertas quando deveriam ser fechadas e *vice-versa*. A frase: *Meta os pés na porta e acerte a meta*, ilustra esses problemas.

Assim, neste trabalho, propõe-se a elaboração de uma etapa dentro do processamento lingüístico que focaliza a resolução dos problemas de alternância vocálica.

A Seção 2 apresenta a descrição do sistema vocálico segundo seus traços fonológicos, juntamente com as regras fonológicas necessárias à alternância vocálica. A Seção 3 mostra as diferentes etapas do processamento lingüístico em sistemas de conversão texto-fala. Na Seção 4, apresentam-se os diferentes estágios do *parser* morfo-sintático elaborado para a realização de uma classificação gramatical consistente, principalmente de verbos e substantivos. A Seção 5 descreve as regras para a geração da alternância vocálica usando a classificação dos radicais verbais rotulados segundo seus diferentes esquemas casuais. Na Seção 6, apresentam-se alguns dos resultados obtidos com a implementação das regras expostas nas Seções 4 e 5. Finalmente, na Seção 7, são apresentadas as conclusões deste trabalho.

2. FORMALIZAÇÃO LINGÜÍSTICA DO SISTEMA VOCÁLICO DO PORTUGUÊS FALADO NO BRASIL

Os sistemas lingüísticos são formalizados através da organização dos sons segundo sua estrutura e funcionamento, sendo essa sistematização objeto de estudo da fonologia. Esses sistemas lingüísticos também podem ser formalizados pelo estudo da realidade física de um som, ou seja, do ponto de vista articulatório (articulação e produção dos sons da fala pelo aparelho fonador), acústico (propriedades físicas da produção e propagação dos sons) e perceptual (percepção dos sons), objeto de estudo da fonética.

Com o objetivo de tentar estabelecer relações entre o nível fonético e o fonológico, descrevendo e analisando como as línguas organizam os sons da fala, surgem as diferentes teorias fonológicas [7]. Conforme essas teorias, os sistemas vocálicos podem ser descritos segundo uma combinação de traços distintivos. Assim, as vogais em posição tônica no PB podem ser descritas pelos seguintes traços [8]:

Vogais	Não-Arredondada		Arredondada
	Anterior	Central	Posterior
Alta	/i/		/u/
Média Alta	/e/		/o/
Média Baixa	/ɛ/		/ɔ/
Baixa		/a/	

Figura 1. Quadro dos traços fonológicos das vogais orais do PB.

A alternância vocálica afeta as formas verbais do português brasileiro, atingindo somente verbos cujos radicais possuem vogal média alta /e, o/ ou média baixa /ɛ, ɔ/¹. Segundo seus traços fonológicos, as vogais médias /e, o/ são caracterizadas pelos traços [-alto] e [-baixo] e as médias /ɛ, ɔ/ pelos traços [-alto] e [+baixo].

De acordo com [7,9], um caso restrito de alternância vocálica é a harmonia verbal que se processa em verbos de vogal temática² “-e” e “-i” (2^a e 3^a conjugações, respectivamente) e estabelece que a última vogal do radical (exceto /a/) se harmoniza em altura com a vogal temática do verbo em questão. Dessa forma, seriam harmônicas a 1^a pessoa do singular do Presente do Indicativo e todas as pessoas do Presente do Subjuntivo.

[sent-i]+ o] → [sint+o] → [sinto] (1)

Em (1), há uma harmonização dos traços da vogal do radical do verbo orientada pela vogal temática, pois a vogal temática de *sentir* [i]³ tem o traço fonológico [+alto] e a vogal do radical [e] em [sent-] tem o traço [-alto]. Assim, a vogal do radical passa de [-alta] a [+alta]⁴.

O problema para a conversão texto-fala ocorre, no entanto, em formas verbais para as quais a Regra de Harmonia Vocálica não se aplica, isto é, em todas as pessoas do Presente do Indicativo e do Subjuntivo dos

¹ Os símbolos /e, o/ correspondem a (ê, ô) e /ɛ, ɔ/, a (é, ó).

² Vogal que determina a que conjugação pertence o verbo.

³ Vogal temática [i] corresponde à 3^a conjugação.

⁴ Para uma melhor observação da mudança de traços fonológicos da vogal do radical, veja a descrição das vogais do português brasileiro apresentada na Fig. 1.

verbos de 1^a conjugação e na 2^a e 3^a pessoas do singular e 3^a do plural desses mesmos tempos verbais para os verbos da 2^a e 3^a conjugações. Para essas formas, é aplicada a Regra de Abaixamento Vocálico, que só se processa em vogais com o traço [+acento] se a Regra de Harmonização Vocálica não for aplicada. Por exemplo, em:

(Tu) **[beb+e]+s]**, (2)

a vogal do radical tem a mesma altura da vogal temática, ou seja, é [-alta], não se aplicando a Regra de Harmonização Vocálica. Por conseguinte, as regras de Acento e de Abaixamento se aplicam.

A Regra de Acento preceitua que a vogal de um radical verbal recebe o traço [+acento] quando está diante de consoante seguida de vogal e de outra consoante (CVC) ou consoante seguida de vogal em final de palavra (CV#). Pela Regra de Acento, em:

(Tu) **[beb+e]+s]**_{verbo}, (3)

temos [¹beb+e]+s]⁵, isto é, a primeira vogal do radical verbal transforma-se na tônica, ou seja, apresenta o traço [+acento].

A Regra de Abaixamento estabelece que uma vogal com os traços [-alto] e [+acento] se torna [+baixa]. Essa regra, aplicada depois da Regra de Acento, leva [¹beb+e]+s] a [¹bɛb+e]+s], ou seja, a vogal acentuada média alta [e] passa a vogal média baixa [ɛ]⁶.

A resolução deste problema depende de uma classificação morfo-sintática bem elaborada. Assim, neste artigo, apresentaremos regras de alternância vocálica que estão inseridas no módulo de classificação gramatical (*parser* morfo-sintático). A correta representação da alternância vocálica dependerá de mecanismos de classificação gramatical suficientemente robustos para uma boa identificação de duas classes gramaticais, verbo e substantivo, necessária às regras de alternância vocálica.

3. PROCESSAMENTO LINGÜÍSTICO EM UM SISTEMA DE CONVERSÃO TEXTO-FALA

Em sistemas de conversão texto-fala [1-6,10-13], as etapas concernentes ao processamento lingüístico são bastante importantes para a etapa de síntese da fala. Para se ter uma idéia das etapas constituintes de um sistema de conversão texto-fala, elaboramos o diagrama da Fig. 2.

⁵ O diacrítico ¹ é uma notação fonológica e indica que a sílaba posterior ao diacrítico é a tônica.

⁶ Essa problematização foi um pouco sintetizada sendo colocada em uma linguagem mais acessível a não lingüistas, como é o caso da maior parte dos interessados em conversão texto-fala.

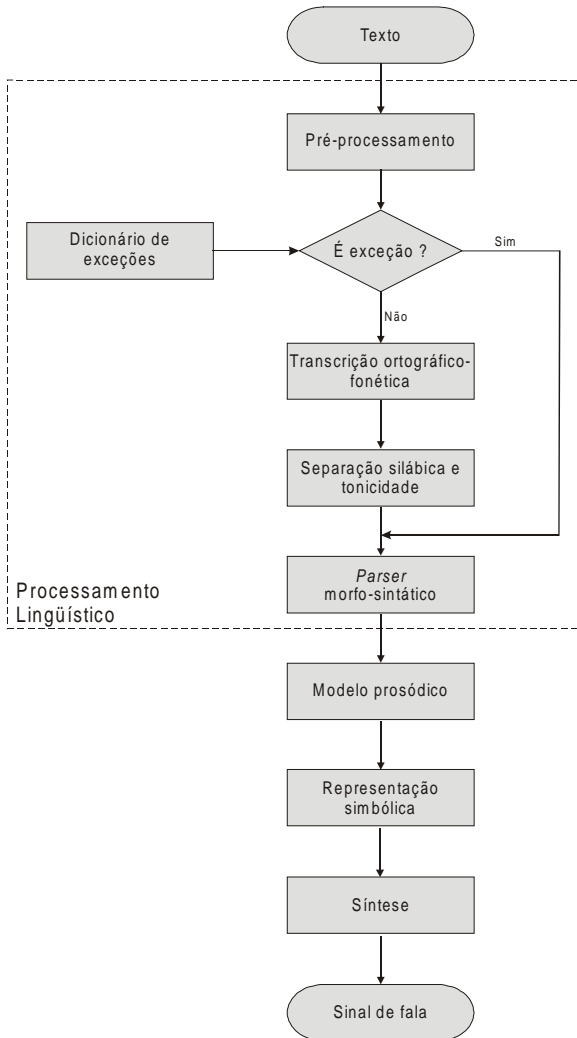


Figura 2. Diagrama das etapas do processo de conversão texto-fala.

Neste diagrama, percebe-se que a parte lingüística tem um peso bastante grande em todo o sistema, orientando inclusive a etapa de modelagem prosódica. Esta dependerá basicamente das informações lingüísticas das etapas anteriores para modificação de seus parâmetros prosódicos, ou seja, duração, intensidade e frequência fundamental.

Os estágios necessários à geração de regras para a alternância vocálica estão inseridos na etapa do *parser* morfo-sintático.

4. CLASSIFICAÇÃO GRAMATICAL

O *parser* morfo-sintático elaborado para o conversor texto-fala opera em quatro estágios:

1. inserção de palavras pertencentes a classes gramaticais fechadas em uma biblioteca com um trabalho manual de etiquetagem de classes gramaticais, radicais e desinências verbais;

2. geração de regras para classificação gramatical dos itens lexicais ainda não classificados;
3. geração de regras para desambigüização dos itens lexicais que possuem mais de uma classe gramatical;
4. geração de regras de alternância vocálica.

No primeiro estágio, são colocados em uma biblioteca os itens lexicais pertencentes às classes gramaticais fechadas, isto é, aquelas que apresentam um número bastante limitado de elementos, tais como artigos, pronomes, preposições, conjunções e os advérbios mais frequentes, com suas respectivas classificações. Em uma outra biblioteca, são etiquetados as desinências e os radicais verbais mais frequentes.

No segundo estágio, é elaborado um conjunto de regras que inicia sua classificação, baseando-se naquela inserida no primeiro estágio. Essa classificação anterior orienta a inserção de regras para a classificação dos demais itens lexicais contidos em uma determinada frase. Por exemplo, em:

O acordo foi fechado, (4)

ocorre a pré-classificação de “o” como artigo - [ART], demonstrativo - [DEM] e pronome oblíquo átono - [OBA]. Conseqüentemente, para esta frase, tem-se o seguinte diagrama de estados:

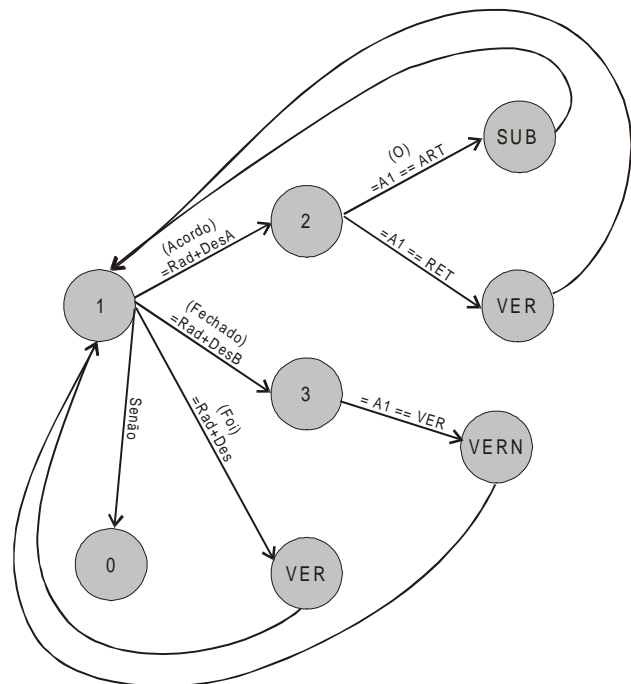


Figura 3 Diagrama de estados para classificação morfo-sintática dos itens lexicais do exemplo (4)⁷.

⁷ DesA: Formas desinenciais utilizadas na alternância vocálica; DesB: Desinências para formas verbo-nominais; Des: Demais desinências verbais.

No terceiro estágio, é feita a “desambigüização” de classes gramaticais, pois alguns dos itens lexicais inseridos na biblioteca apresentam mais de uma classificação gramatical. Um exemplo disso ocorre com o “o” apresentado em (4), que encerra as seguintes classes: artigo (ART), pronome pessoal oblíquo átono (OBA) e pronome demonstrativo (DEM). Por exemplo, em:

- i) *O menino saiu de casa.* (o → ART)
- ii) *Ele matou-o com duas facadas.* (o → OBA)
- iii) *O moço de verde foi o que me bateu.* (o → DEM)

Neste terceiro estágio, então, formulam-se regras que possibilitam uma única classificação. Para o caso dos exemplos (ii) e (iii) apresentados anteriormente, têm-se:

Exemplo (ii)⁸

SE (([i]∈C1) **E** ([i-1]=VER) **E** ([i+1]=PREP)) ⇒ [i]=OBA

Exemplo (iii)⁹

SE (([i]∈C1) **E** ([i+1]∈D2)) ⇒ [i]=DEM

O quarto estágio é o foco deste trabalho e depende basicamente de uma classificação consistente de verbos e nomes no terceiro estágio. Para esse estágio, foi elaborada uma divisão dos radicais necessários à identificação das formas verbais em 6 casos.

5. GERAÇÃO DE REGRAS PARA A ALTERNÂNCIA VOCÁLICA

Somente passarão pelo quarto estágio, o de geração de regras para a alternância vocálica, e receberão a nova classificação os radicais cujas vogais são afetadas pela alternância vocálica. Esses radicais foram etiquetados manualmente em 6 casos, uma vez que a alternância vocálica se dá, como veremos abaixo, de maneira diferenciada.

As regras de transcrição ortográfico-fonética resolveram grande parte dos problemas de abertura e fechamento vocálico, para a classe de substantivos. Porém, trouxeram alguns problemas para as formas verbais e suas desinências. Por exemplo, abriram vogais de desinências verbais que deveriam estar fechadas. Na forma verbal

⁸ [i]=Item Lexical-Alvo; [i-1]=Item Lexical Anterior; [i+1]=Item Lexical Posterior; C1=Grupo constituído pelos artigos definidos; VER=Grupo constituído pelos itens lexicais classificados como verbo; PREP=Grupo constituído pelos itens lexicais classificados como preposição; OBA=Grupo constituído pelos itens lexicais classificados como pronome oblíquo átono.

⁹ D1=Grupo constituído pelo pronome relativo *que*; DEM=Grupo constituído pelos itens lexicais classificados como pronome demonstrativo.

termos (1ª pessoa do plural do Infinitivo Pessoal), a fala sintetizada especificava a primeira vogal da desinência verbal como [+baixa] [ε], ao invés de especificá-la como [-baixa] [e]. Assim, esses casos também seriam candidatos potenciais para o estágio de alternância vocálica.

Esquemas de classificação prévia dos radicais:

CASO 1 – radicais de verbos da 1ª conjugação que podem ser classificados ora como nome ora como verbo e cuja vogal do radical é fechada, ou seja, tem o traço [-baixo] para os nomes e é aberta, traço [+baixo], para a 1ª, 2ª e 3ª pessoas do singular e 3ª do plural do Presente do Indicativo e do Presente do Subjuntivo dos verbos.

Ex.: *O almoço.* (5a)

Eu almoço. (5b)

Que ele almoce. (5c)

CASO 2 – radicais de verbos da 2ª conjugação cuja vogal do radical é fechada, ou seja, tem o traço [-baixo] na 1ª pessoa do Presente do Indicativo e em todas as pessoas do Presente do Subjuntivo e do Imperativo e é aberta, traço [+baixo], para a 2ª e 3ª pessoas do singular e 3ª do plural dos verbos do Presente do Indicativo e para os substantivos.

Ex.: *Meta-se com seus pares.* (6a)

Ele mete os pés na porta. (6b)

A meta é escrever. (6c)

CASO 3 – radicais de verbos da 3ª conjugação cuja vogal do radical é aberta, ou seja, tem o traço [+baixo], somente para a 2ª e 3ª pessoas do singular e 3ª do plural dos verbos do Presente do Indicativo.

Ex.: *Tu dormes.* (7a)

Eles dormem. (7b)

CASO 4 – radicais de verbos cuja alteração vocálica ocorre na desinência verbal, ou seja, a primeira vogal da desinência é aberta, traço [+baixo], e as regras de transcrição ortográfico-fonética não deram conta dessa alteração.

Ex.: *Se eu tiver um livro.* (8a)

Eles tiveram de sair cedo. (8b)

CASO 5 – a alteração vocálica ocorre na desinência verbal, sendo a vogal da desinência fechada, ou seja, tem o traço [-baixo], para o verbo e é aberta, traço [+baixo], para os nomes.

Ex.: *Eu vou colher uma flor.* (9a)

A colher de pau. (9b)

CASO 6 – radicais de certos verbos irregulares cuja vogal do radical é fechada, ou seja, tem o traço [-baixo] para os verbos e é aberta, traço [+baixo], para as outras classes gramaticais.

Ex.: *Ele fora aposentado.* (10a)

Ele está fora de casa. (10b)

A partir da classificação de formas nominais e verbais nos 6 casos anteriormente elencados, criou-se um bloco para as regras de alteração vocálica de radicais e desinências. Esse bloco levará em consideração, para a elaboração de suas regras, o código de alternância (CASOS) e a desinência verbal. Nos casos 3 e 4, as regras só foram válidas para as palavras classificadas como verbos (VER), nos demais casos, foram feitas regras dependentes da classificação verbo (VER) ou substantivo (SUB).

A seguir, apresentaremos algumas das regras elaboradas para a alteração vocálica¹⁰:

Para o CASO 1:

1) **SE** (([i]=VER) **E** (CASO=1) **E** ([DESA]=o)) ⇒ [RAD-1]=[+baixo]. Veja em (5b).

2) **SE** (([i]=SUB) **E** (CASO=1) **E** ([DESA]=o)) ⇒ [RAD-1]=[-baixo]. Veja em (5a).

Para o CASO 2:

1) **SE** (([i]=VER) **E** (CASO=2) **E** ([DESA]=a)) ⇒ [RAD-1]=[-baixo]. Veja em (6a).

2) **SE** (([i]=SUB) **E** (CASO=2) **E** ([DESA]=a)) ⇒ [RAD-1]=[+baixo]. Veja em (6c).

Para o CASO 3:

1) **SE** (([i]=VER) **E** (CASO=3) **E** ([DESA]=es)) ⇒ [RAD-1]=[+baixo]. Veja em (7).

Para o CASO 4:

1) **SE** (([i]=VER) **E** (CASO=4) **E** ([DESA]=er)) ⇒ [DES-1]=[+baixo]. Veja em (8).

Para o CASO 5:

1) **SE** (([i]=VER) **E** (CASO=5) **E** ([DESA]=er)) ⇒ [DES-1]=[-baixo]. Veja em (9a).

2) **SE** (([i]=SUB) **E** (CASO=5) **E** ([DESA]=er)) ⇒ [DES-1]=[+baixo]. Veja em (9b).

Para o CASO 6:

1) **SE** (([i]=VER) **E** (CASO=6) **E** ([DESA]=a)) ⇒ [RAD-1]=[-baixo]. Veja em (10a).

2) **SE** (([i]≠VER) **E** (CASO=6) **E** ([DESA]=a)) ⇒ [RAD-1]=[+baixo]. Veja em (10b).

6. RESULTADOS

A partir da implementação das regras apresentadas na Seção 5, foram feitos testes exaustivos, buscando observar o maior número possível de palavras que pudessem apresentar problemas de alternância vocálica. Nesses testes, obteve-se êxito na pronúncia da fala sintetizada para praticamente todos os itens testados. Problemas não resolvidos nas regras apresentadas em diversos sistemas de conversão texto-fala [1-4] foram solucionados. Por exemplo, as regras apresentadas em [3], não gerariam a forma correta para a vogal “e” do radical da forma correspondente à 1ª pessoa do singular do presente do indicativo “meto”. Suas regras estabelecem que a letra “e” das palavras que estão no presente do indicativo singular será aberta “é” (ou seja, terá o traço [+baixo]).

As regras aqui propostas resolveram inclusive problemas do tipo: *Quero que ele corte a carne com o corte que a corte recomendar*, no qual o verbo (*ele corte*) e o substantivo masculino (*o corte*) possuem a vogal de seus radicais aberta (traço [+baixo]) e o substantivo feminino (*a corte*) possui essa mesma vogal fechada (traço [-baixo]).

7. CONCLUSÕES

As regras de alternância vocálica discutidas neste trabalho resolveram a maior parte dos problemas concernentes ao fechamento e abertura vocálicos em nomes e verbos. Algumas dessas regras funcionam para casos bastante restritos, mas foram necessárias para uma correta pronúncia da representação do texto de entrada.

Os problemas que ainda não foram resolvidos dizem respeito às regras de classificação gramatical, ou seja, alguns itens lexicais são classificados como verbos quando deveriam ter sido classificados como substantivo e *vice-versa*. Sendo assim, ocorre uma inversão de alternância nas vogais dos radicais, abrindo quando deveriam fechar e fechando quando deveriam abrir. No entanto, se a classificação gramatical for correta, as regras de alternância vocálica apresentadas funcionam satisfatoriamente.

Os pequenos problemas ainda não resolvidos não são desencorajadores, já que o processo de ajuste fino em sistemas de conversão texto-fala é bastante delicado e evolutivo.

AGRADECIMENTOS

Agradecemos à Empresa de Telecomunicações DÍGITRO Tecnologia Ltda. pelo financiamento dado a esta pesquisa.

¹⁰ DESA=Formas desinenciais utilizadas na alternância vocálica; SUB=Grupo constituído pelos itens lexicais classificados como substantivo; [RAD-1]=Última vogal do radical; [DES-1]=Primeira vogal da desinência; [+baixo]=Vogal será aberta, ou seja, terá o traço [+baixo]; [-baixo]=Vogal será fechada, ou seja, terá o traço [-baixo].

REFERÊNCIAS

- [1] F. Egashira, “Síntese de voz a partir de texto para a língua portuguesa,” Dissertação de Mestrado, Unicamp, Campinas, Jul. 1992.
- [2] C. H. Silva, “Modelamento prosódico para conversão texto-fala do português falado no Brasil,” Dissertação de Mestrado, Unicamp, Campinas, Dez. 1995.
- [3] L. C. T. Gomes, “Sistema de conversão texto-fala para a língua portuguesa utilizando a abordagem de síntese por regras,” Dissertação de Mestrado, Unicamp, Campinas, Jul. 1998.
- [4] P. A. Aquino, “O papel das vogais reduzidas pós-tônicas na construção de um sistema de síntese concatenativa para o português do Brasil,” Dissertação de Mestrado, Unicamp, Campinas, Nov. 1997.
- [5] M. N. Souza, L. P. Calôba, J. M. Seixas, C. G. Machado e M. V. Ludolf, “Sintetizador de voz do projeto Processador Automático de Português,” *Proceedings of XII Brazilian Automatic Control Conference – XII CBA*, Uberlândia, MG, vol. VI, pp. 2093-2096, Set. 1998.
- [6] F. A. Figueiredo, L. A. B. Naviner e B. G. Aguiar Neto, “Uma nova abordagem para o sistema de conversão texto-fala para a língua portuguesa,” *XV Simpósio Brasileiro de Telecomunicações*, Recife, PE, pp. 328-331, Jan. 1997.
- [7] L. Bisol, *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 1996.
- [8] J. M. Câmara Jr., *Estrutura da língua portuguesa*. Petrópolis: Vozes, 1970.
- [9] M. H. M. Mateus, *Aspectos de fonologia do português*. Lisboa: Centro de Estudos Filológicos, 1975.
- [10] D. H. Klatt, “Review of text-to-speech conversion for English,” *Journal of the Acoustical Society of America*, vol. 82, no. 3, pp. 737-793, Sept. 1987.
- [11] M. Edgington, A. Lowry, P. Jackson, A. P. Breen e S. Minnis, “Overview of current text-to-speech techniques: Part I – text and linguistic analysis,” *BT Technol J*, vol. 14, no. 1, pp. 68-83, Jan. 1996.
- [12] T. Dutoit, *An introduction to text-to-speech synthesis*. Dordrecht: Kluwer, 1997.
- [13] R. V. Cox, C. A. Kamm, L. R. Rabiner, J. Schroeter e J. G. Wilpon, “Speech and language processing for next-millennium communication services,” *Proceedings of the IEEE*, vol. 88, no. 8, pp. 1314-1337, Aug. 2000.