

Caracterização semiautomática da tipologia rítmica do português brasileiro

Wellington da Silva
Grupo de Estudos de Prosódia da Fala
Universidade Estadual de Campinas
Campinas, Brasil
dablio_w@yahoo.com.br

Plínio A. Barbosa
Grupo de Estudos de Prosódia da Fala
Universidade Estadual de Campinas
Campinas, Brasil
pabarbosa.unicampbr@gmail.com

Resumo— No presente trabalho, investigamos a relação entre as medidas de produção do ritmo do português brasileiro apontadas por três métodos de caracterização do ritmo da fala e a percepção dele por uma bateria de ouvintes. O principal método utilizado foi uma técnica semiautomática fundamentada no modelo dinâmico do ritmo de Barbosa [3]. Utilizamos como corpora três gravações no estilo leitura e três no estilo semiespontâneo. Os outros métodos utilizados foram o índice de variabilidade pareada e o comprimento resultante em estatística circular. Aplicamos, então, um teste de discriminação com fala deslexicalizada, a fim de comparar os valores obtidos com esses métodos e a resposta dos sujeitos. Não foi possível chegar a resultados conclusivos, mas a análise se mostrou um passo crucial para o estudo da percepção do ritmo da fala.

Palavras-chave: Fonética; Prosódia; Ritmo da Fala; Teste de Discriminação

I. INTRODUÇÃO

Os esforços para caracterizar o ritmo das línguas naturais têm resultado em diferentes técnicas e abordagens. Houve um período no qual o ritmo era entendido como um fenômeno de regularidade temporal estrita e as línguas classificadas tipologicamente em ritmo acentual ou em ritmo silábico. Uma língua teria ritmo acentual se a duração dos grupos acentuais (formado por uma sílaba proeminente no enunciado e todas as átonas que a seguem, em línguas com cabeça à esquerda, ou que a precedem, em línguas com cabeça à direita) fosse a mesma (isocronismo acentual) e ritmo silábico se a duração das sílabas fosse a mesma (isocronismo silábico), vide [2] para uma revisão. Porém, tal isocronia absoluta nunca foi atestada empiricamente e o ritmo passou a ser visto como um ponto entre os dois extremos ideais: acentual e silábico, com a coexistência de ambos os tipos [4], sendo a isocronia entendida como um fenômeno da percepção da fala [9].

No presente trabalho, foi utilizado um procedimento semiautomático para caracterizar a tipologia rítmica do português brasileiro, à luz da teoria de osciladores acoplados de Barbosa [3], desenvolvida no seio de um programa de pesquisa que integra pesquisadores como Cummins, Port, e O'Dell. De acordo com essa teoria, o ritmo das línguas naturais é determinado pelo acoplamento de dois osciladores: acentual e silábico, existentes no sistema cognitivo do falante.

A força de acoplamento entre esses dois osciladores é o que indica se a língua tende ao ritmo acentual ou ao silábico. Os pulsos do oscilador silábico determinam os *onsets* vocálicos do enunciado e os pulsos do oscilador acentual, os locais de acento frasal.

Empiricamente atesta-se que a duração do grupo acentual tem uma componente linear que é função do número de sílabas nele contidas e, por isso, é possível obter uma regressão linear com essas duas variáveis. O'Dell e Nieminen [11] mostraram, através da técnica matemática *Averaged Phase Difference*, que a força de acoplamento (c) entre os dois osciladores pode ser estimada através da razão entre o coeficiente de interceptação da reta de regressão com o eixo das ordenadas (a) e o coeficiente de inclinação da reta (b) da equação (1), em que I é a duração do grupo acentual e n o número de sílabas fonéticas ou fonológicas do grupo acentual correspondente ($c = a/b$). Se a força de acoplamento (c) for igual a 1, significa que ambos os osciladores se influenciam de maneira equivalente. Se $0 < c < 1$, tem-se uma dominância do oscilador silábico sobre o acentual (tendência ao ritmo silábico), e se $c > 1$, o oscilador acentual domina o silábico (tendência ao ritmo acentual).

$$I = a + b \cdot n \quad (1)$$

Também foram utilizados, para comparação, dois outros métodos para caracterizar o ritmo da fala: o índice de variabilidade pareada (*normalized Pairwise Variability Index* de [10]) e uma técnica fundamentada em estatística circular, usada anteriormente por Brady e Port [8].

O índice de variabilidade pareada (PVI) mostra, em termos de duração, o quanto uma unidade linguística silábica ou subsilábica difere da unidade imediatamente seguinte. É computado de acordo com a fórmula (2), em que d_k corresponde à duração bruta da unidade V-V (unidade delimitada por dois *onsets* de vogais consecutivos, incluindo a vogal à esquerda) e m ao número total de unidades V-V. As unidades V-V em posição de fronteira de grupo acentual não foram consideradas para o cálculo. Um PVI alto reflete grandes irregularidades de unidade a unidade (ritmo acentual) e um PVI baixo, regularidade (ritmo silábico).

$$nPVI_{vv} = 100 \times \left[\frac{\sum_{k=1}^{m-1} |d_k - d_{k+1}|}{\sum_{k=1}^{m-1} (d_k + d_{k+1})/2} \right] / (m-1) \quad (2)$$

Este trabalho é resultado da Iniciação Científica de Wellington da Silva, intitulada "Caracterização automática da tipologia rítmica do português brasileiro", que foi orientada pelo prof. Dr. Plínio A. Barbosa e realizada com bolsa concedida pela Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, Processo 2010/01273-2, no período de 01/04/2010 a 31/03/2011.

A técnica fundamentada em estatística circular consiste em transformar as durações das unidades V-V em ângulos, calculando o quociente entre a duração de cada unidade V-V (d_{vv}) e a média das durações das unidades V-V ($meand_{vv}$), multiplicando-o por 2π para obter o valor em radianos. A seguir, calcula-se o comprimento resultante (R), que corresponde à soma dos vetores correspondentes aos pontos do círculo, de acordo com a fórmula (3). O comprimento resultante médio é então obtido através da divisão de R pelo número total de unidades V-V da amostra. Se os pontos do círculo forem mais agrupados, seus vetores correspondentes terão aproximadamente a mesma direção e o comprimento resultante médio será próximo de 1, que é o valor normalizado da principal direção resultante. Esse fato indica uma predominância de **ritmo silábico**, já que a característica dessa classe rítmica é de sílabas idealmente com a mesma duração. Se, por outro lado, os pontos do círculo estiverem muito dispersos, teremos vetores correspondentes em várias direções, e, portanto, o comprimento resultante médio será relativamente pequeno, o que indica predominância de **ritmo acentual**.

$$R^2 = \left[\sum \sin \left(\frac{d_{vv}}{meand_{vv}} \times 2\pi \right) \right]^2 + \left[\sum \cos \left(\frac{d_{vv}}{meand_{vv}} \times 2\pi \right) \right]^2 \quad (3)$$

II. METODOLOGIA

Para esse estudo, três locutores paulistas, duas mulheres e um homem, pós-graduandos na faixa de 30 a 45 anos de idade sem problemas auditivos ou fonoarticulatórios, leram o texto “O monge desastrado”, com cerca de 1.500 palavras, (estilo leitura) e logo em seguida contaram sobre o que ele se tratava (estilo narração; semiespontâneo). As seis gravações foram segmentadas de forma automática em unidades V-V com um *script* para o programa PRAAT [7], o *Beatextractor* [3], seguido de correção manual. A escolha da unidade V-V se justifica pelas pesquisas do *perceptual-center* como ponto de ancoragem universal para a percepção de isocronismo, o qual se dá no *onset* de uma vogal [2,5]. Além disso, Barbosa [1] mostrou que, no português brasileiro, a evolução das durações das unidades V-V sucessivas em frases isoladas indica pontos de culminância de crescendos duracionais que delimitam os grupos acentuais, fato não observável sistematicamente com as sílabas.

As unidades foram, então, etiquetadas manualmente. Em seguida, os grupos acentuais foram delimitados de forma automática, com o auxílio do *script* SGdetector [3], que detecta as fronteiras dos grupos acentuais a partir das durações normalizadas e suavizadas das unidades V-V e fornece uma tabela com a duração desse grupo e o número de unidades V-V nele contidas, além de uma estimativa da força da fronteira que delimita o grupo à direita. O número de sílabas fonológicas contidas em cada grupo acentual foi informado após contagem manual, pois estávamos interessados em comparar os resultados obtidos com as unidades V-V e com as sílabas, para esse método.

Assim como em [6], optou-se por considerar o nível de proeminência de cada grupo acentual para explicar a variância

da duração deste. Este nível de proeminência corresponde a uma medida normalizada (*z-score*) da duração da última unidade V-V do grupo acentual, a que carrega o acento frasal. A adição dessa variável resulta na regressão linear múltipla representada pela equação (4), em que z_p representa esse nível de proeminência, sendo os demais coeficientes os mesmos da equação (1). A força de acoplamento (c) continua sendo determinada por alb . Essa variável, juntamente com o número de sílabas (fonológicas ou unidades V-V), explica a maior porcentagem da variância da duração do grupo acentual. Aplicando esse método, obtemos um coeficiente de determinação (r^2) em torno de 70 %, isto é, 70 % da variância da duração dos grupos acentuais provêm da variância do número de unidades (V-V ou sílaba) no mesmo e da proeminência expressa pela duração normalizada.

$$I = a + b \cdot n + p \cdot z_p \quad (4)$$

Análises Estatísticas

A fim de avaliar possíveis mudanças no ritmo de fala dos locutores ao longo das leituras e narrações, dividimos as gravações em trechos e para cada um deles construímos uma regressão linear. Em seguida, aplicamos análises de covariância (ANCOVA) para avaliar a significância estatística da diferença entre as retas de regressão linear.

Para realizar a ANCOVA, primeiramente consideramos todos os coeficientes de inclinação (b e p) dos trechos. Quando os resultados para a influência do trecho considerando as duas inclinações não eram significativos, desconsideramos o efeito de trecho (inclinação indistinta do trecho 1) e refizemos a ANCOVA. No caso de intercepção de trecho não significativa, também foi considerada indistinta do trecho 1. No caso de o coeficiente de intercepção (a), ter valor não significativo, aceitamos a hipótese nula e consideramos o mesmo como nulo (ritmo silábico).

III. RESULTADOS

As tabelas 1 e 2 trazem as equações das regressões lineares para as sílabas fonológicas e unidades V-V, respectivamente, bem como as respectivas forças de acoplamento. Os locutores estão identificados pelas duas primeiras letras, sendo que a terceira corresponde ao sexo (F: feminino e M: masculino). Os estilos são identificados pelas siglas “Le” (leitura) e “Na” (narração; semiespontâneo). O número após a barra indica o trecho correspondente.

Pode-se ver que o ritmo do locutor FA é mais acentual no estilo narração do que na leitura, tanto para as sílabas fonológicas quanto para as unidades V-V. Quanto à locutora AG, a parte 1 para as sílabas fonológicas da narração teve um ritmo muito mais acentual do que os trechos da sua leitura para essa mesma unidade. Esse também foi, entre os locutores, o trecho mais acentual para as sílabas fonológicas, considerando tanto as leituras quanto as narrações.

TABELA 1: Equações das regressões lineares obtidas para as **sílabas fonológicas**. I corresponde à duração do grupo acentual, n ao respectivo número de sílabas fonológicas e z ao valor de z -score suavizado da respectiva proeminência. p corresponde à significância do coeficiente de interceptação (a) e ns indica não significativo. Coeficientes c significativos em negrito.

Locutor/Trecho	Regressão Linear	c
FAM – Le / 1	$I = 189 + 126n + 71z$ ($p < 0,04$)	1,5
FAM – Le / 2	$I = 0 + 166n + 50z$ (ns)	0
FAM – Le / 3	$I = 0 + 150n + 57z$ (ns)	0
FAM – Na	$I = 237 + 143n + 45z$ ($p < 0,07$)	1,7
AGF – Le / 1	$I = 0 + 170n + 59z$ (ns)	0
AGF – Le / 2	$I = 90 + 170n + 46z$ ($p < 0,05$)	0,53
AGF – Le / 3	$I = 196 + 170n + 41z$ ($p < 0,0004$)	1,15
AGF – Le / 4	$I = 135 + 170n + 42z$ ($p < 0,02$)	0,79
AGF – Le / 5	$I = 151 + 170n + 32z$ ($p < 0,05$)	0,89
AGF – Na / 1	$I = 366 + 137n + 43z$ ($p < 0,01$)	2,7
AGF – Na / 2	$I = -15 + 197n + 43z$ ($p < 0,04$)	0
AGF – Na / 3	$I = 0 + 137n + 43z$ (ns)	0
LLF – Le / 1	$I = 212 + 126n + 68z$ ($p < 0,03$)	1,68
LLF – Le / 2	$I = 0 + 161n + 41z$ (ns)	0
LLF – Le / 3	$I = -31 + 167n + 56z$ ($p < 0,02$)	0
LLF – Na / 1	$I = 0 + 184n + 47z$ (ns)	0
LLF – Na / 2	$I = 0 + 135n + 47z$ (ns)	0

As partes 2 e 3 da sua narração (considerando ainda as sílabas fonológicas), porém, apresentaram um ritmo mais silábico do que os trechos 2, 3, 4 e 5 da sua leitura. Para as unidades V-V dessa locutora, as partes 2,3 e 4 da leitura são um pouco mais acentuais do que os trechos da sua narração.

A parte 1 da leitura da locutora LL para as sílabas fonológicas é mais acentual do que a leitura do locutor FA e da locutora AG para a mesma unidade. É também mais acentual do que sua narração. Considerando as unidades V-V, a parte 2 da narração da locutora LL é o trecho mais acentual entre os locutores e estilos.

IV. TESTE DE DISCRIMINAÇÃO

Com o propósito de investigar se as pessoas são capazes de identificar as diferenças de ritmo assinaladas por esses métodos, elaboramos e aplicamos um teste de discriminação, inédito na literatura. O teste consistiu em apresentar aos sujeitos pares de trechos de fala deslexicalizados, extraídos das gravações utilizadas nas análises anteriores. O método usado para retirar a informação lexical dos enunciados foi o proposto por [13]. O uso de fala deslexicalizada é importante por garantir que os ouvintes se concentrem apenas na prosódia dos enunciados.

Solicitou-se a eles que julgassem se o modo de falar nos dois trechos do par apresentado era semelhante ou diferente, selecionando um valor em uma escala graduada de 0 a 6. O valor 0 indicava modos de falar idênticos e o valor 6, modos de falar diferentes e era possível selecionar valores intermediários. Quando a pessoa clicava com o mouse no valor desejado, o próximo estímulo já era apresentado. Era permitido ao sujeito ouvir cada estímulo apenas mais uma vez, caso desejasse, clicando em um botão apresentado na parte inferior da tela.

TABELA 2: Equações das regressões lineares obtidas para as **unidades V-V**. I corresponde à duração do grupo acentual, n ao respectivo número de unidades V-V e z ao valor de z -score suavizado da respectiva proeminência. p corresponde à significância do coeficiente de interceptação (a) e ns indica não significativo. Coeficientes c significativos em negrito.

Locutor/Trecho	Regressão Linear	c
FAM – Le / 1	$I = 0 + 180n + 69z$ (ns)	0
FAM – Le / 2	$I = 224 + 180n + 50z$ ($p = 0,05$)	1,24
FAM – Le / 3	$I = 0 + 180n + 55z$ (ns)	0
FAM – Na	$I = 262 + 175n + 42z$ ($p = 0,05$)	1,5
AGF – Le / 1	$I = 0 + 223n + 42z$ (ns)	0
AGF – Le / 2	$I = 15 + 223n + 42z$ ($p < 0,1$)	0,07
AGF – Le / 3	$I = 79 + 223n + 42z$ ($p < 0,005$)	0,35
AGF – Le / 4	$I = 43 + 223n + 42z$ ($p < 0,05$)	0,19
AGF – Le / 5	$I = 0 + 223n + 42z$ (ns)	0
AGF – Na / 1	$I = 0 + 194n + 40z$ (ns)	0
AGF – Na / 2	$I = 0 + 266n + 40z$ (ns)	0
AGF – Na / 3	$I = 0 + 194n + 40z$ (ns)	0
LLF – Le / 1	$I = 0 + 190n + 59z$ (ns)	0
LLF – Le / 2	$I = 188 + 190n + 36z$ ($p < 0,05$)	0,99
LLF – Le / 3	$I = 0 + 190n + 59z$ (ns)	0
LLF – Na / 1	$I = -258 + 284n + 43z$ ($p < 0,1$)	0
LLF – Na / 2	$I = 305 + 185n + 43z$ ($p < 0,02$)	1,65

O experimento foi montado com a ferramenta *ExperimentMFC*, através de um *script* para o programa PRAAT. Foram selecionados 36 trechos de 1 e de 2 segundos, pois também era de nosso interesse investigar se a duração dos trechos influenciaria nas respostas dos ouvintes. Cada estímulo (par) foi usado duas vezes, mas em ordem inversa, por exemplo: "AGreadA1,FAreadA1" e "FAreadA1,AGreadA1". Isso foi feito para que se testasse a coerência dos julgamentos dos sujeitos. No total, 76 estímulos compuseram o experimento. As respostas dos sujeitos foram convertidas em uma escala de -1 (mesmo modo) a 1 (modos diferentes), com os valores 0, -0,67, -0,33, 0, 0,33, 0,67 e 1. O valor -1 foi posteriormente transformado em 0, por indicar "mesmo modo", aplicando-se a mesma equação para todas as respostas.

Análises

O teste foi aplicado em 21 sujeitos, sendo 6 homens e 15 mulheres, de idade entre 18 e 25 anos e sem problemas auditivos ou fonoarticulatórios, por meio de um computador e com uso de fones de ouvido.

Das respostas obtidas, foram selecionadas aquelas que tiveram desvio-padrão menor que a "média dos desvios-padrão menos a metade de um desvio-padrão dos desvios-padrão", pois são as respostas menos variáveis entre os sujeitos. Depois, verificamos a consistência das respostas dos ouvintes nesses estímulos. Esperava-se que as respostas não variassem muito quando a ordem dos pares de trechos era invertida. Tal verificação foi feita a partir da diferença entre a média das respostas para os dois estímulos (o original e o de ordem invertida). Foram descartados 2 estímulos que apresentaram essa diferença maior que o critério descrito acima. Ao todo, foram selecionados 20 estímulos para a análise posterior.

Em seguida, calculamos as diferenças dos valores obtidos com os métodos (c , PVI e R) entre os trechos de cada par, por exemplo: para o par "AGreadA1,FAreadA1", tomamos o valor

do c da parte 1 da leitura de AG e subtraímos o valor do c da parte 1 da leitura de FA, e assim por diante, para os três métodos (assim, se ambos tivessem um mesmo valor de c a diferença seria 0, indicando ritmos de fala idênticos, segundo essa técnica). Então construímos gráficos de dispersão entre essas diferenças e a média das respostas dos sujeitos.

Resultados

Para o método fundamentado na teoria de osciladores acoplados, quando consideramos as regressões lineares múltiplas, não foram alcançadas correlações satisfatórias para os gráficos ($r^2 = 0,0018$ para as unidades V-V e $r^2 = 0,0244$ para as sílabas fonológicas). A razão disso se deve ao grande número de coeficientes não significativos obtidos. Por isso, optamos por utilizar as regressões lineares simples (calculadas sem levar em consideração o z -score da proeminência dos grupos acentuais como variável), a fim de investigar se dessa forma esse método consegue pelo menos prever parte da percepção dos sujeitos. Foi obtido um resultado interessante: foi alcançado um coeficiente de determinação maior quando consideradas as unidades V-V ($r^2 = 0,1852$), em comparação às sílabas fonológicas ($r^2 = 0,0037$). O resultado obtido para as unidades V-V também foi estatisticamente significativo: $p < 10^{-8}$ para o coeficiente de interceptação e $p < 0,06$ para o de inclinação.

Para o índice de variabilidade pareada (PVI) foi obtido um coeficiente de determinação de 0,1169. No entanto, essa correlação foi significativa somente para o coeficiente de interceptação da reta: $p < 10^{-5}$.

Para o método fundamentado em estatística circular não foi obtido um coeficiente de determinação relevante ($r^2 = 0,0773$).

Um teste T mostrou que a duração dos trechos não influenciou na coerência das respostas dos sujeitos: $p \approx 0,59$. Ou seja, a diferença entre a média das respostas para os estímulos e seus equivalentes invertidos com a duração de 1 segundo foi estatisticamente a mesma com a duração de 2 segundos.

V. CONSIDERAÇÕES FINAIS

O uso das unidades V-V permite a automatização da técnica fundamentada na teoria de osciladores acoplados. No entanto, em alguns casos o uso das sílabas fonológicas se mostrou mais eficaz, resultando em um maior número de coeficientes significativos. Ao relacionar as medidas de produção com a percepção dos sujeitos, o uso das unidades V-V alcançou resultados mais expressivos, pelo menos quando consideradas as regressões lineares simples. O número de coeficientes não significativos foi grande, fato que pede uma investigação mais apurada, com maior número de dados. De fato, o número de dados necessários para se obter resultados significativos é uma das limitações desse modelo. Contudo, a possibilidade de ser submetido a uma análise estatística não deixa de ser uma vantagem.

O uso do nível de proeminência dos grupos acentuais para explicar a variância da duração destes juntamente com o

número de unidades linguísticas contidas neles melhorou consideravelmente a capacidade explicativa do modelo.

Como se pode ver nas tabelas 1 e 2, o método de osciladores acoplados foi capaz de captar diferenças no ritmo de fala entre os locutores, entre os dois estilos (leitura e narração) e também ao longo das gravações.

É possível que a escala de 7 pontos utilizada no teste de discriminação exija dos sujeitos um grau de percepção do ritmo da fala bastante apurado, o que possivelmente eles não têm. Por isso, seria interessante utilizar em futuros experimentos uma escala com apenas 5 pontos, com o propósito de testar essa hipótese.

Os métodos que melhor previram as respostas dos sujeitos no teste de discriminação foram o índice de variabilidade pareada (PVI) e o método fundamentado na teoria de osciladores acoplados (para o caso das unidades V-V com regressões lineares simples), mas ambos com uma correlação baixa.

VI. AGRADECIMENTOS

Agradecemos aos locutores e a todos os sujeitos que participaram do experimento aqui relatado, especialmente a Beatriz de O. Salgado e Ana Luiza B. C. Roxo por ajudarem no recrutamento de sujeitos para o experimento.

VII. REFERÊNCIAS

- [1] BARBOSA, P. A. *At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration: emphasis on segmental duration generation*. Cadernos de Estudos Linguísticos, 31: 33-53, 1996.
- [2] BARBOSA, P. A. *Syllable-Timing in Brazilian Portuguese: uma Crítica a Roy Major*. DELTA, 16 (2), p. 369-402, 2000.
- [3] BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas, Brasil: Pontes/Fapesp, 2006.
- [4] BARBOSA, P. A. *Measuring speech rhythm variation in a model-based framework*. In Proc. of Interspeech 2009 - Speech and Intelligence. Londres: Causal Productions. p. 1527 – 1530, 2009.
- [5] BARBOSA, P. A., ARANTES, P., MEIRELES, A. R., VIEIRA, J. M. *Abstractness in Speech-Metronome Synchronisation: P-Centres as Cyclic Attractors*. In Proceedings of the Ninth European Conference on Speech Communication and Technology (Interspeech 2005) Lisbon, Portugal, p. 1441-1444, 2005.
- [6] BARBOSA, P. A.; VIANA, M. C.; TRANCOSO, I. *Cross-variety Rhythm Typology in Portuguese*. In Proc. of Interspeech 2009 - Speech and Intelligence. Londres: Causal Productions. p. 1011 – 1014, 2009.
- [7] BOERSMA, P. e WEENINK, D. *Praat: doing phonetics by computer* (Versão 5.1.04). Online em <<http://www.praat.org>>, acessado em 2009.
- [8] BRADY, M. C. e PORT, R. F. *Quantifying Vowel Onset Periodicity in Japanese*. International Congress of Phonetic Sciences, Saarbrücken. p. 337-342, 2007.
- [9] LEHISTE, I. *Isosynchrony reconsidered*. Journal of Phonetics. 5: 253-263, 1977.
- [10] LOW, E. L., GRABE, E., NOLAN, F. *Quantitative characterisations of speech rhythm: Syllable-timing in Singapore English*. Language and Speech, 43:377–401, 2000.
- [11] O'DELL, M. e NIEMINEN, T. *Coupled Oscillator Model of Speech Rhythm*. Proc. XIVth ICPhS, San Francisco, 1075-1078, 1999.
- [12] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>. Acessado em 2010.

- [13] VAINIO, M. *et al.* *New Method for Delexicalization and its Application to Prosodic Tagging for Text-to-Speech Synthesis*. In Proc. of Interspeech 2009 - Speech and Intelligence. Londres: Causal Productions. p. 1703 – 1706, 2009.