

Tipologia rítmica de dialetos do português brasileiro

Alexsandro Rodrigues Meireles

Programa de Pós-Graduação em Linguística
Universidade Federal do Espírito Santo
Vitória, Brasil
meirelesalex@gmail.com

Viviany de Paula Gambarini

Departamento de Línguas e Letras
Universidade Federal do Espírito Santo
Vitória, Brasil
vivi-blink@hotmail.com

Resumo — Este artigo analisa o ritmo da fala de três dialetos do português brasileiro sob a luz de uma abordagem de osciladores acoplados à tipologia rítmica. Os dados mostram uma escala de um dialeto menos acentual a mais acentual da seguinte forma: MG (força de acoplamento = 1.83, taxa de elocução = 6.3) > SP (1.66, 6.6) > BA (1.26, 4.2) > ES (0.95, 4.0). Além disso, os resultados mostram que o ritmo acentual parece estar relacionado a taxas mais rápidas. Finalmente, coeficientes de intercepto significativos em uma equação de regressão ocorreram mais frequentemente na taxa rápida. Este fato é apoiado por resultados anteriores que mostram um decréscimo de desvio-padrão nas taxas rápidas. Conseqüentemente, equações de regressão significantes podem ser mais fáceis de serem encontradas nessas taxas.

Palavras-chave: ritmo de fala, taxa de elocução, ritmo de fala, ritmo acentual e silábico.

I. INTRODUÇÃO (HEADING 1)

A literatura apresenta duas abordagens principais na pesquisa do ritmo da fala: a) a abordagem dos osciladores acoplados [4, 22]; e b) a abordagem descritiva [16, 23].

A abordagem descritiva realiza uma tipologia rítmica de acordo com parâmetros como %V (durações de intervalos vocálicos, [23]), 8C (durações de intervalos consonantais, [23]), PVI (índice de variabilidade pareada, [20]) e CCI (índice controle/compensação, [8]). Embora os dois últimos parâmetros tenham melhorado consideravelmente as medidas de ritmo da fala, por incluírem informações de diferenças locais de duração, nenhum desses modelos considera as duas propriedades essenciais de produção do ritmo postuladas por [14]: periodicidade e estrutura. Da mesma forma, Kohler [15] comenta que nem as durações de intervalos consonantais e/ou vocálicos nem o PVI e um modelo explicativo do ritmo da fala; e apenas uma forma de “data sorting on the basis of consonantal and vocalic, i.e. local segmental, durations in their point-to-point variability, not with global rhythmical patterns”.

Por outro lado, a abordagem dos osciladores acoplados, a qual lida com a interação entre um oscilador silábico e um acentual, é capaz de explicar “both universal and language-specific properties of rhythm by means of general principles applicable to all languages” [7]. Além disso, uma comparação entre o índice PVI do tamanho da sílaba e a força de acoplamento entre os dois osciladores do modelo dinâmico do ritmo (de agora em diante, MDR) mostrou que o último parâmetro é mais afinado com o desempenho rítmico dos falantes, pois reflete melhor diferenças em estilos de fala [7].

Geralmente, as abordagens de osciladores acoplados classificam o ritmo da fala pela força de acoplamento entre osciladores silábicos e acentuais, a qual é medida pelo quociente (força relativa, de agora em diante, r) entre o coeficiente intercepto (a) e inclinação (b) da equação de regressão linear $I = a + bn$, em que I é a duração do grupo acentual e n é o número de unidades do tamanho da sílaba. Assim, línguas de ritmo acentual teriam uma maior influência do oscilador acentual ($r > 1$) e línguas de ritmo silábico teriam uma maior influência do oscilador silábico ($r \leq 1$). Foi demonstrado por [7], porém, que a inclusão do nível de proeminência (de agora em diante, p) do oscilador acentual reflete melhor a variância das durações dos grupos acentuais. Por isso, p foi incluído em nossa análise, a qual é baseada nos conceitos teórico-metodológicos do MDR [5]. A fórmula utilizada para análise do ritmo da fala foi, portanto:

$$I = a + bn + cp \quad (1)$$

Barbosa [2,7] e Meireles [19], entretanto, mostram que devemos ser cuidadosos na análise do ritmo da fala, pois r pode variar bastante em função do estilo de fala [1,6,7], da taxa de elocução [2,17,19,21] e dos dialetos [10,19].

Meireles [19], por exemplo, realizou um estudo interdialeto para estudar a variação rítmica nos dialetos mineiro e baiano com as seguintes premissas: a) línguas de ritmo silábico possuem menor desvio-padrão da duração da unidade vogal-avogal (de agora em diante, VV), bem como maior desvio-padrão da duração do grupo acentual (de agora em diante, GA); b) línguas de ritmo acentual possuem maior desvio-padrão da duração da unidade VV, bem como menor desvio-padrão da duração do grupo acentual (de agora em diante, GA).

Essa diferente abordagem para analisar tipologia rítmica, baseada em medidas descritivas de dados acústicos, embora plausível, perde a interação entre os osciladores silábico e acentual. Esta pesquisa, por outro lado, compara os resultados anteriores de Meireles [19] sobre esses dois dialetos usando a abordagem de classificar os ritmos da fala de acordo com o parâmetro força de acoplamento, bem como acrescenta o dialeto capixaba na análise.

II. METODOLOGIA

A. Corpora (Heading 2)

Para este estudo, um excerto de um livro de Monteiro Lobato [17] foi lido por 12 falantes nativos do português

brasileiro, na faixa de 15 a 25 anos, em três taxas de elocução (lenta, normal e rápida). Os falantes são mineiros de Belo Horizonte (2 homens e 2 mulheres), baianos de Conceição do Jacuípe (2 homens e 2 mulheres) e capixabas de Vitória (2 homens e 2 mulheres). São, pois, considerados representantes típicos de seus dialetos.

As distintas taxas de elocução foram obtidas com as seguintes instruções e ordem: (1) normal: fala de maneira confortável; (2) lenta: fale o mais lento que puder, sem introduzir pausas entre as sentenças; (3) rápida: fale o mais rápido que puder sem introduzir distorções na fala.

B. Procedimentos

Procedimentos semi-automáticos foram usados na observação de diferentes estruturas rítmicas com o aumento da taxa de elocução. Primeiramente, as unidades VV foram etiquetadas no Praat [9] com o script BeatExtractor [4], seguida de correção manual. O intervalo do começo de uma vogal até a próxima vogal define a unidade VV. Por exemplo, “em seguid(a)” foi segmentada como: /eNs/, /eg/ e /id/, e então foi extraída a duração de cada unidade VV. Por fim, o script Stress Group Detector do Praat [9] foi rodado, resultando em informações como: (i) duração do VV momento-a-momento; (ii) duração do grupo acentual; e (iii) unidades VV por GA.

Os grupos acentuais foram delimitados automaticamente pelos seguintes passos: a) normalização por z-score para eliminar os efeitos da duração intrínseca dos segmentos em cada duração do VV [4]; b) suavização da evolução dos valores de z-scores usando a fórmula:
$$z_{smoothed} = \frac{5z^i + 3z^{i-1} + 3z^{i+1} + z^{i-2} + z^{i+2}}{15}$$
. Esse procedimento minimiza os efeitos de oscilação local de duração (acentos lexicais não marcados frasalmente e efeitos remanescentes de duração intrínseca) e determina as fronteiras de grupos acentuais por pontos de máximo na curva de duração suavizada; c) computação da duração do GA e do número de unidades VV por GA para comparar as distribuições em três taxas de elocução rates — lenta, normal e rápida — para cada falante (a distribuição de duração e número de unidades VV. O número de sílabas fonológicas por grupo acentual foi contado manualmente.

C. Força de acoplamento entre os osciladores

Como visto anteriormente, a força de acoplamento entre os osciladores silábico e acentual é calculada pelo quociente entre o intercepto (a) e a inclinação (b) da equação de regressão ($r = a/b$). Além disso, como o acréscimo do nível de proeminência (p) reflete melhor a variância das durações dos grupos acentuais, essa medida foi acrescentada na regressão linear múltipla. Essa última medida é simplesmente o z-suavizado de cada acento frasal, que é retornado pelo script SGDetector. Como [6,7] aponta: “due to the techniques applied, this value is not a subpart of the stress group duration”, mas melhora os coeficientes de correlação dos dados.

III. RESULTADOS

Semelhantemente aos resultados de [6, 7], o número de unidades VV não produziram interceptos significativos para a maioria dos falantes e condições (6 de 32 coeficientes, 2 dialetos x 4 falantes x 3 taxas, mais o dialeto baiano x 4 falantes x 2 taxas (normal, rápida)). Por essa razão, incluímos o número de sílabas fonológicas (SF) como uma variável independente na análise. No entanto, diferentemente dos estudos de Barbosa, o uso da SF não resultou em um maior número de interceptos significativos (5 de 32 coeficientes). Como podemos notar, a taxa lenta foi excluída da análise do dialeto baiano, pois todos falantes introduziram pausas entre as palavras.

É importante ressaltar que dos 6 interceptos significativos com unidades VV, 5 foram encontrados na taxa rápida. Além disso, todos coeficientes significativos com SF foram achados na taxa rápida. Esse achado relaciona-se com os resultados de Meireles [17, 18, 19] sobre a diminuição do desvio-padrão da duração das unidades VV e do GA nas taxas rápidas. Considerando essa diminuição, a probabilidade de achar resultados significados em regressões lineares é maior nas taxas rápidas.

Como tínhamos um conjunto complexo de dados para correlação (3 dialetos x 3 taxas x 4 falantes), decidimos primeiramente usar a taxa normal para comparação dialeto, e, então, comparar individualmente os efeitos da taxa de elocução em cada dialeto. Assim, executamos ANOVAs unifatoriais para observar similaridades de taxa de elocução entre os dialetos e falantes. Em outras palavras, queríamos saber se as taxas qualitativas (lenta, normal, rápida) eram similares às correspondentes taxas quantitativas (sílabas/s). Com base nesses resultados, por não serem suas taxas qualitativas iguais a dos outros falantes, os seguintes falantes foram excluídos das seguintes taxas: a) lenta: BC e RC (MG); b) normal: VH (MG); e c) rápida: BM (ES), VH (MG), e BA (BA). Todas as análises subsequentes, portanto, não incluíram esses falantes.

Uma ANOVA unifatorial com taxa de elocução em função dos dialetos, todos falantes incluídos, mostrou uma diferença significativa entre os dialetos apenas para a taxa normal ($F(2, 2068)=8,6228$, $p < 0.00019$). Esse resultado pode ser explicado pelas instruções para aquisição da taxa. Os sujeitos falaram o mais lentamente possível na taxa lenta e o mais rapidamente possível na taxa rápida, mas, confortavelmente, na taxa normal, i.e., essa última taxa reflete melhor uma forma natural de pronunciar sentenças nos dialetos. Em acréscimo, essa taxa reflete resultados anteriores comparativos entre os dialetos mineiro e baiano. Meireles e colegas [19] mostraram que os falantes do sudeste falam mais rápido que os do nordeste. Assim, nossos dados corroboram esse resultado, pois os dois dialetos do sudeste (MG, ES) foram significativamente diferentes do do nordeste (BA) na taxa normal (MG média (x) = 4.6, e desvio-padrão (dp) = 0.5; ES x = 4.4, e sd = 0.4; BA x = 3.9 média, e dp = 0.4).

No entanto, apenas um exemplo de intercepto significativo foi achado na taxa normal. Por outro lado, todos os outros interceptos significativos foram encontrados na taxa rápida.

Assim, todas comparações rítmicas entre os dialetos foram baseadas nessa taxa. Além disso, os textos sem significância para este parâmetro foram excluídos da comparação dialetal. Por fim, tivemos 2 falantes representando o dialeto baiano, 2 representando o mineiro e 1 falante o capixaba. Como podemos ver, esses falantes foram agrupados na tabela 1.

A tabela 1 mostra as equações de regressão linear na taxa rápida de acordo com o dialeto (MG, ES, BA) e o tipo de unidade silábica (VV ou SF). A força de acoplamento c é o quociente entre o intercepto (a) e a inclinação das equações. Todos coeficientes de correlação foram altamente significantes ($R^2 > 0.82$, $p < 10^{-4}$). As significâncias desses interceptos estão indicada entre parênteses. c é considerado indefinido para um valor não-significativo de intercepto ou para um a negativo. A taxa de elocução (te) é dada em VV/s ou SF/s. Esses resultados estão plotados em função da taxa de elocução na figura 1.

Tabela 1. Equações de regressão na taxa rápida para 3 dialetos do PB (MG, ES, BA) considerando o número de unidades VV (nVV) e o número de sílabas fonológicas ($nSil$). Força de acoplamento (c), speech rate (te), e a magnitude do acento frasal (p) são dadas.

Dialect	Regression equation	r	te
MG	$I = 222 + 132nVV + 106p$ ($6.4e^{-5}$)	1.68	6.3
MG	$I = 187 + 102nSil + 86p$ ($7.1e^{-5}$)	1.83	6.3
ES	$I = -232 + 272nVV + 60p$ (0.01)	u	4.0
ES	$I = 121 + 128nSil + 49p$ (0.02)	0.95	4.0
BA	$I = 125 + 153nVV + 63p$ (0.002)	0.81	4.2
BA	$I = 143 + 113nSil + 63p$ ($6.4e^{-3}$)	1.26	4.2

Considerando o número de SF, os resultados mostram que o dialeto mais acentual é o mineiro (1,83), seguido pelo baiano (1,26) e capixaba (0,95), o que está de acordo com os resultados de Meireles [19], o qual seguiu uma diferente abordagem metodológica. Pelo menos para dois dialetos, esse padrão é o mesmo se consideramos o número de VVs (MG (1.68) > BA (0.81)). Em acréscimo, embora não foi encontrada significância para taxa em função dos dialetos, a maior taxa foi encontrada para o dialeto mineiro, que também está de acordo com resultados prévios de Meireles.

Apesar de nossos dados sugerirem que o dialeto baiano é mais acentual do que o capixaba, mais dados precisam ser coletados para clarear esse ponto, pois apenas um a significativo foi encontrado para este dialeto. Pode até ser que o falante capixaba estudado seja um falante naturalmente devagar que não segue as tendências rítmicas de sua área dialeto. Além disso, nenhuma análise ANCOVA foi rodada para observar se as equações de regressão desses dialetos foram estatisticamente diferentes.

IV. DISCUSSÃO

Os resultados dão suporte ao uso da abordagem de osciladores acoplados para explicar a variabilidade rítmica nas línguas do mundo. Não apenas o MDR é capaz de representar os ritmos linguísticos levando em consideração todos os aspectos importantes para produção do ritmo (em essência a

periodicidade e a estruturação), mas também é sensível a variações linguísticas em função do estilo de fala, da taxa de elocução, das estruturas sintáticas e do conhecimento prévio do texto lido.

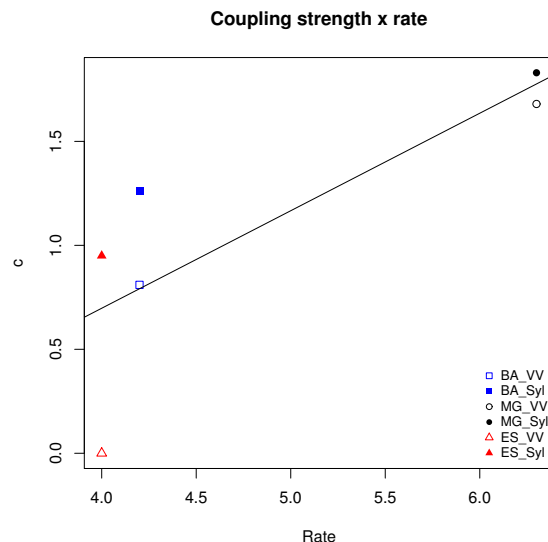


Figure 1. Relação entre taxa de elocução (unidades VV ou SF) e a força de acoplamento para 3 dialetos do PB.

Um ponto a ser considerado na abordagem, contudo, e o fato da análise ser bastante demorada. Muitos falantes precisam ser usados para serem selecionados poucos dados com o coeficiente intercepto estatisticamente significativo. Em nosso caso, apenas 5 regressões lineares (de 36) puderam ser usadas na análise. Também precisamos investigar porque a sílaba fonológica pode ser um parâmetro mais confiável do que a unidade VV (vide [6,7]).

Nossos resultados mostram uma aplicação da abordagem de osciladores acoplados para tipologia rítmica. De acordo com os dados, devidamente tratados estatisticamente, encontramos uma escala de ritmo menos acentual para mais acentual da seguinte forma: MG ($r = 1,83$; taxa de elocução = 6,3) > SP (1,66; 6,6) > BA (1,26; 4,2) > ES (0,95; 4,0).

Barbosa [6] argumenta que o aumento da taxa de elocução favorece o ritmo silábico. Esse argumento parece estar refletido em seus dados (vide tabela 2 do artigo de Barbosa [6]). No entanto, não temos certeza se isso é verdadeiro para o português brasileiro em geral, pois, em nossos dados, os dialetos mais acentuais foram os mais rápidos. Meireles e colegas [19] discutiram essa questão e concluíram que as taxas rápidas podem, ao contrário de Barbosa, estar associada com ritmo acentual. Assim, estudos posteriores com maior número de falantes e dialetos precisarão ser realizados para investigar melhor os efeitos da taxa de elocução no ritmo da fala.

Outra questão relevante apontada por nossos resultados e a ocorrência de coeficientes significativos do intercepto principalmente na taxa rápida. Esse fato encontra respaldo nos estudos de Meireles [17,18,19], os quais mostram um decréscimo do desvio-padrão em taxas rápidas, i.e., uma

menor variabilidade duracional e esperada em taxas extremas. Como consequência, pode ser mais fácil encontrar equações de regressão linear com parâmetros significativos nas taxas rápidas.

ACKNOWLEDGMENTS (HEADING 5)

Os autores agradecem Plínio Barbosa, Karen Currie e Pablo Arantes por sugestões e o CNPq por financiar a pesquisa.

REFERENCES

[1] Abaurre-Gnerre, M. B. Processos fonológicos segmentais como índices de padrões prosódicos diversos nos estilos formal e casual do português do Brasil. *Caderno de Estudos Lingüísticos*, 2: 23-34, 1981.

[2] Barbosa, P. A. Syllable-timing in Brazilian Portuguese”: uma crítica a Roy Major. *D.E.L.T.A.*, 16 (2), 369-402, 2000.

[3] Barbosa, P. A., “Explaining Cross-Linguistic Rhythmic Variability via a Coupled-Oscillator Model of Rhythm Production”, Proc. Speech Prosody 2002 Conf. [CD], Aixen- Provence, 163–166, 2002.

[4] Barbosa, P. A., *Incursões em torno do ritmo da fala*, Campinas, Brazil: Pontes/Fapesp, 2006.

[5] Barbosa, P. A., “From syntax to acoustic duration: a dynamical model of speech rhythm production”, *Speech Communication*, 49:725–742, 2007.

[6] Barbosa, P. A. “Measuring speech rhythm variation in a modelbased framework”. In.: Proc. Interspeech 2009, Brighton, UK, 1527-1530, 2009.

[7] Barbosa, P. A, Viana, M. C., Trancoso, I. “Cross-variety Rhythm Typology in Portuguese”. In.: Proc. Interspeech 2009, Brighton, UK, 1011-1014, 2009.

[8] Bertinetto, P. M., Bertini, C., “Towards a unified predictive model of Natural Language Rhythm”, *Quaderni del Laboratorio di Linguistica della SNS*, 7, 2007/08.

[9] Boersma, P., Weenink, D., “Praat: doing phonetics by computer” (Version 5.0.35) [Computer program], Online: <http://www.praat.org>, accessed in 2008.

[10] Cagliari, L. C. and Abaurre, M. B. M. Elementos para uma investigação instrumental das relações entre padrões rítmicos e processos fonológicos no português brasileiro. *Caderno de Estudos Lingüísticos*, 10: 39-57, 1986.

[11] Cummins, F., Port, R., “Rhythmic constraints on stress timing in English”, *J. Phon.*, 26:145–171, 1998.

[12] Deterding, D., “Letter to the Editor. The measurement of rhythm: a comparison of Singapore and British English”, *J. Phon.*, 29:217–230, 2001.

[13] Fraise, P. Les rythmes. *Journal Français d’Oto-Rhino-laryngologie*, n. Supplément numéro 7, p. 23–33, 1968.

[14] Fraise, P. *La Psychologie du Rythme*. Paris: Presses Universitaires de France, 1974.

[15] Kohler, K. J. “Rhythm in speech and language: a new research paradigm”, *Phonetica*, 66:29-45, 2009.

[16] Low, E. L., Grabe, E., Nolan, F., “Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English”, *Language and Speech*, 43:377–401, 2000.

[17] Meireles, A. R. *Self-organizing rhythms in Brazilian Portuguese: speech rate as a system perturbation*. Germany: VDM Verlag, 2009.

[18] Meireles, A. R. and Barbosa, P. A. Speech rate effects on speech rhythm. In: Speech Prosody 2008 Conference, 2008, Campinas. *Proceedings of the Speech Prosody 2008 Conference*. Campinas: RG. v.1. p.327 – 330, 2008.

[19] Meireles, A. R., Tozetti, J. P., Borges, R. R. Speech rate and rhythmic variation in Brazilian Portuguese. In: Speech Prosody 2010 Conference, 2010, Chicago. *Proceedings of the Speech Prosody 2010 Conference*. Chicago: RG. v.1. p.1 – 4, 2010.

[20] Mok, P. P. K., Dellwo, V., “Comparing native and nonnative speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English”, Proc. Speech Prosody 2008, Campinas, 423-426, 2008.

[21] Moraes, J. A. and Leite, Y. F. Ritmo e velocidade da fala na estratégia do discurso: uma proposta de trabalho. In: R. Ilari (org.) *Gramática do*

Português falado, v. II: níveis de análise lingüística. Campinas: Editora da Unicamp, 1992.

[22] O’Dell, M., Nieminen, T., “Coupled Oscillator Model of Speech Rhythm”, Proc. XIVth ICPhS, San Francisco, 1075-1078, 1999.

[23] Ramus, F., Nespors, M., Mehler, J., “Correlates of linguistic rhythm in the speech signal”, *Cognition*, 73:265–292, 1999.