

# PARA ALÉM DOS PACOTES ESTATÍSTICOS VARBRUL/ GOLDVARB E RBRUL: QUAL A CONCEPÇÃO DE GRAMÁTICA?<sup>1</sup>

## *BEYOND THE STATISTICAL PROGRAMS VARBRUL/ GOLDVARB AND RBRUL: WHAT IS THE CONCEPT OF GRAMMAR?*

Christina Abreu Gomes

Universidade Federal do Rio de Janeiro/CNPq/Faperj

### RESUMO

Esse artigo discute a questão da modelagem teórica da variação e a metodologia estatística utilizada na quantificação dos dados variáveis, centrada na comparação entre os Programas Varbrul/Goldvarb e Rbrul. Discute, ainda, o papel do indivíduo e do item lexical em relação à variação e à mudança e suas consequências para um modelo teórico que defende um status representacional para a variação sociolinguística.

**Palavras-chave:** indivíduo; item lexical; softwares Varbrul/Goldvarb e Rbrul; variação sociolinguística.

### ABSTRACT

This paper discusses the issue of theoretical modeling of variation and the statistical methods used in quantifying variable data, centered in a comparison between Varbrul/Goldvarb and Rbrul softwares. It also discusses the role of the individual and that of the lexical item in relation to variation and change and their consequences for the theoretical approach that defends a representational status for sociolinguistic variation.

**Keywords:** individual; lexical item; Varbrul/Goldvarb and Rbrul softwares; sociolinguistic variation.

<sup>1</sup> *Dedico esse artigo à minha querida companheira de projetos, reuniões, congressos e confraternizações, Cláudia Roncarati, a quem sempre admirei pela dedicação ao trabalho e pela coragem de enfrentar a vida e suas surpresas.*

## INTRODUÇÃO

Um estudo sociolinguístico, que utiliza dados de produção espontânea, tem o objetivo de capturar o caráter sistemático da variabilidade observada no uso e definir que aspectos da gramática – isto é, internos – e do uso – isto é, sociais – são responsáveis por essa sistematicidade. A variação é concebida como indexada, ou seja, veicula características sociais dos falantes como idade, sexo, classe social, escolaridade, e condições de uso, observadas através da variável estilo de fala. Sendo assim, uma análise desta natureza congrega fatores estruturais e sociais para identificar processos de mudança em curso ou variação estável em uma comunidade de fala. Uma questão central, então, é poder identificar o efeito dos diversos efeitos/condicionamentos em competição e determinar qual a contribuição de cada categoria (estrutural e não-estrutural) associada à ocorrência de determinada forma em oposição a outra ou outras. Em última instância, a sistematicidade da variação revelada no efeito observado das variáveis estruturais e sociais traz evidências a respeito do conhecimento linguístico internalizado pelo falante. Portanto, subjacente ao estudo de uma variável sociolinguística está uma determinada concepção de gramática.

Desde o início dos estudos variacionistas, foi proposto um modelo de análise multivariada associado à criação de um programa computacional especificamente desenhado para lidar com dados de variação linguística (Labov, 1969, Cedergreen e Sankoff, 1974, Rousseau e Sankoff, 1978). A questão relativa aos métodos estatísticos para a análise de dados de variação linguística tem se centrado principalmente na evolução dos modelos matemáticos subjacentes ao tratamento quantitativo dos dados variáveis. Naro (2003:15-225) apresenta uma breve cronologia desses modelos (aditivo, multiplicativo e logístico) subjacentes ao pacote computacional Varbrul e, mais recentemente, Goldvarb, em suas diversas versões. Scherre (1996) e Guy e Zilles (2007) trataram de aspectos teóricos e metodológicos relacionados à aplicação e uso do programa Varbrul/Goldvarb. Oliveira (2009) compara resultados de regressão logística para dados de variação sociolinguística obtidos com a utilização do Varbrul/Goldvarb e do Programa SPSS com o objetivo de identificar limitações e contribuições das duas ferramentas. Mais recentemente, foi elaborado por Daniel Ezra Johnson (Johnson, 2009) outro programa estatístico de regressão logística

para análise de dados de variação linguística, batizado de Rbrul, com o intuito de avançar e refinar alguns aspectos da análise estatística que não são possíveis de serem realizados pelo Programa Goldvarb. No entanto, esta nova ferramenta também abre a possibilidade de análises que remetem a questões fundamentais relativas ao status da variação linguística na gramática.

Neste artigo vamos apresentar as diferenças fundamentais entre os Programas Varbrul/Goldvarb e o Rbrul e analisar as consequências para a compreensão da variação linguística no que diz respeito à concepção de gramática que está em jogo quando são adotados determinados procedimentos metodológicos. Esperamos contribuir, com o presente artigo, para divulgar o tratamento estatístico dos dados através da ferramenta Rbrul e abrir a discussão sobre a relação entre procedimentos metodológicos e pressupostos teóricos<sup>2</sup>. O Programa Rbrul, em função da maneira como foi concebido, aumenta o alcance explanatório de algumas questões importantes como, por exemplo, o papel do indivíduo na variação, a compreensão de processos de mudança que envolvem o efeito do item lexical e a incorporação de variáveis contínuas como frequência.

## 1. A ferramenta Rbrul

O Programa Rbrul é um software gratuito, idealizado por Daniel Ezra Johnson. O programa e o manual de uso estão disponíveis na página <http://www.danielezrajohnson.com/rbrul.html>. A atual versão do Rbrul é a 2.01<sup>3</sup>, de 7 de setembro de 2011. O programa roda na plataforma R, que também é um software gratuito disponível para download em [www.r-project.org](http://www.r-project.org). O autor realiza atualizações periódicas no programa e mantém atualizadas as informações sobre o programa no site acima referido, assim como as modificações realizadas em cada versão e suas consequências, em um arquivo com o objetivo de facilitar o uso do Rbrul (ver <http://www.danielezrajohnson.com>).

<sup>2</sup> *Essa pesquisa é financiada pelo CNPq (Processo no. 304400/2010-6) e pela FAPERJ (Processo no. 26/102.405/2009).*

<sup>3</sup> *Consultar a página mencionada para verificar novas atualizações e versões do programa.*

O Programa foi idealizado para cumprir 3 grandes objetivos, a saber:

- realizar todas as funções que as versões Goldvarb e Varbrul realizam (regressão múltipla, tabulação cruzada, *step up/step down*);

- realizar funções que as versões Goldvarb e Varbrul não realizam (por exemplo: rodar variáveis contínuas como variáveis independentes; variáveis contínuas como variável dependente; dar conta de modelos mistos – by-speaker, by-item; estimar efeitos *between-groups* – gênero/sexo – e *within groups* – indivíduos);

- estabelecer interface com as capacidades gráficas do R.

Ainda, o programa apresenta resultados em log-odds, além de apresentar o peso relativo, dá conta dos knockouts sem precisar excluí-los, e não tem limites para número de fatores por grupo.

O programa reconhece arquivos de dados de natureza diferente. Tanto arquivos do tipo .tkn, usados no Goldvarb, quanto arquivos .txt, usados no Varbrul e convertidos para Excel, servem de arquivos de entrada para a rodada. Há uma formatação específica para cada um dos dois tipos para que o arquivo seja compatível com os requisitos do Rbrul. A conversão de um arquivo .tkn é feita no decorrer dos procedimentos da rodada e a conversão de um arquivo .txt deve ser realizada previamente para preparação deste arquivo até chegar à formatação como .csv, através do Programa Excel<sup>4</sup>.

Um exemplo de arquivo de dados que serve de entrada para uma rodada no Rbrul pode ser visto na Figura 1 abaixo. A codificação através de caracteres do teclado pode ser usada e/ou transformada em texto. O mecanismo de conversão de um arquivo do tipo .txt, que contém uma cadeia de codificação conforme a usada para rodar dados no Varbrul/Goldvarb, em um arquivo de Excel é muito simples e, nesse processo, caracteres podem ser transformados em texto, ou os contextos, que normalmente acompanham os dados codificados após a cadeia de codificação, podem ser transformados em uma variável independente. A figura 1 apresenta o resultado da transformação de um arquivo .txt em um arquivo .xls, inicialmente com uma cadeia de codificação do tipo exemplificado a seguir, seguida de contexto, de um estudo sobre a coda fricativa:

---

<sup>4</sup> Não é objetivo deste artigo abordar o passo a passo da conversão do arquivo .txt ou mesmo das rodadas.

(1) (0&fmacnMi mas de vez em quando

Após eliminar o contexto mais amplo que a palavra, restando somente o item com a variável em questão como em

(2) (0&fmacnMi mas

o arquivo é transformado em arquivo de Excel. Neste novo arquivo as colunas do .txt são interpretadas como colunas do arquivo Excel e uma última coluna é criada com todos os itens lexicais, que passarão a constar como mais um grupo de fatores ou variável independente.

FIGURA 1: resultado de conversão de arquivo .txt para excel

126	a	internal	tri	am&fmacnMi;non-morg;informal	Bod&fmacnMi	prop&fmacnMi	prop&fmacnMi			prop&fmacnMi
127	a	internal	tri	am&fmacnMi;non-morg;informal	Bod&fmacnMi	prop&fmacnMi	prop&fmacnMi			prop&fmacnMi
128	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
129	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
130	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
131	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
132	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
133	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
134	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
135	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
136	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
137	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
138	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
139	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
140	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri
141	h	final	no&fmacnMi	am&fmacnMi;non-morg;informal	Bod&fmacnMi	tri	tri			tri

O arquivo do Excel precisa ser convertido para .csv (comma separate values) para servir de entrada de dados para o Rbrul.

Com relação às funções não realizadas pelo Varbrul/Goldvarb destacamos três. A primeira diz respeito à possibilidade de se utilizar variável contínua como variável dependente ou independente sem ser necessário discretizá-la, como por exemplo, formantes de vogais ou faixa etária. Outro aspecto bastante importante é a capacidade de o programa reconhecer vários tipos de arquivos. Essa ferramenta, associada ao fato de que o programa reconhece texto como fator de um grupo e de que não

há limite de fatores por grupo, cria a possibilidade de se observar o papel do item lexical na variação sonora. Além disso, o programa faz diferença entre variáveis independentes de efeito fixo (*fixed effect*) e de efeito aleatório (*random effect*). Iremos nos ater a essas duas últimas e suas consequências para a discussão sobre o status da variação linguística na gramática na próxima seção.

No entanto, cabe ressaltar o ganho metodológico e teórico de ser possível fazer a diferença entre variáveis fixas e aleatórias numa rodada. Se dois efeitos são aninhados, isto é, as características de um são mapeadas em outra variável, então o efeito aninhado será o efeito aleatório e o outro será o efeito fixo. Se, por exemplo, as seguintes variáveis estiverem sob consideração, falante e sexo/gênero, falante está aninhado em sexo/gênero e será o efeito aleatório. Com relação a variáveis de efeito fixo e aleatório, é importante lembrar que no Varbrul/Goldvarb, as variáveis independentes (ou grupos de fatores) são concebidas como independentes entre si, mesmo que não sejam. Por isso, numa rodada em que haja interesse de verificar o indivíduo, e outras características estão sendo mapeadas, como, por exemplo, sexo/gênero, faixa etária e escolaridade, é sempre necessário realizar rodadas diferentes, a que levará em conta as características sociais de estratificação da amostra e a que observará somente o indivíduo. No Rbrul basta identificar qual a(s) variável(is) aleatória(s) e qual(is) a(s) fixa(s) para que, naquela rodada, o programa identifique a relação entre os diferentes tipos. Por exemplo, é possível realizar uma rodada com a variável independente indivíduo (*random effect*), e as variáveis faixa etária, escolaridade e sexo/gênero (*fixed effect*).

## 2. O papel do indivíduo e o papel do item lexical

Por trás da questão metodológica envolvendo o modelo estatístico utilizado no Varbrul/Goldvarb, existe o pressuposto de que indivíduo e item lexical não têm efeito sobre a variação. Esse tratamento se baseia em posições teóricas assumidas tanto em relação ao indivíduo quanto ao item lexical, respectivamente, em função da relação entre indivíduo e comunidade de fala, e da concepção de gramática fonológica subjacente à análise da variação fonológica (ou sociofonética).

Na abordagem dos Modelos baseados no Uso, a relação entre gramática e uso é capturada através da postulação de que gramática é a organização cognitiva da experiência do indivíduo com a língua (Bybee, 2006). A gramática é abstraída a partir de mecanismos cognitivos de domínio geral, que permitem, categorizar e estabelecer relações de semelhança e diferença e de certos aspectos da experiência do indivíduo que participam da modelagem das representações. A abordagem da sociolinguística situa a experiência do indivíduo na sua dimensão social.

Assim, com relação ao indivíduo, o esperado é que a variação observada para cada falante reflita os valores observados para a comunidade de fala. Guy (1980) mostra, em estudo sobre o -t,-d *deletion* no inglês da Filadélfia, que os efeitos das variáveis independentes analisadas na comunidade de fala atuam da mesma maneira nos indivíduos, apesar de haver diferença nas médias individuais. Por outro lado, alguns estudos têm mostrado que indivíduos podem não apresentar a regularidade esperada em virtude dos resultados obtidos para a comunidade de fala. Oliveira (1992), em estudo sobre acentuação de vogais, defende que o indivíduo é mais regular que a comunidade, se se levar em consideração o item e não o segmento, uma vez que alguns falantes demonstraram um comportamento categórico para alguns itens. Na introdução de Paiva e Duarte (2003:28), acerca de trabalhos desenvolvidos pelo grupo de pesquisas PEUL/UFRJ como estudo de painel (dados do mesmo falante em momentos diferentes), há referência ao comportamento irregular de alguns indivíduos, uma vez que alguns consistem evidência para reforçar a hipótese clássica de estabilidade da gramática, enquanto outros apresentam “flutuações quantitativas e qualitativas (...) demonstrando, assim, algumas descontinuidades de comportamento”. Martins (2007), observando diferentes variáveis fonológicas, aponta que a variação intra-individual também é parte fundamental para a compreensão da variação linguística.

O fato é que diversos estudos têm demonstrado que nem sempre os indivíduos apresentam um comportamento que reflita o da comunidade de fala. Daí a importância de ser possível capturar o papel do indivíduo numa metodologia estatística que permita avaliar a relação entre o comportamento do falante e as demais variáveis sociais em questão. É importante que se diga que Labov não descarta a variabilidade do indivíduo. Qualquer trabalho de

variação que utiliza o construto do tempo aparente precisa identificar, em relação à distribuição das variantes por faixa etária, se há uma mudança em progresso (mudança na comunidade de fala) ou se a distribuição constitui um caso de gradação etária (mudança no indivíduo – *age grading*) (cf. Bailey, 2002). Labov localiza a gramática do indivíduo de maneira diferente, no sentido de que os idioletos (a gramática de cada indivíduo) não têm importância per se, uma vez que a abordagem baseada exclusivamente nos idioletos não é suficiente para capturar a estrutura sociolinguística de uma comunidade de fala. Assim, o indivíduo precisa ser situado em relação à estrutura social em que se insere, identificado através das variáveis idade, sexo/gênero, classe social, etc.

O ponto que quero defender aqui é que aspectos específicos do comportamento do indivíduo associados aos comportamentos mais gerais observados nos grupamentos sociais devem ser levados em conta no entendimento da variação e, mais especificamente, em relação à mudança linguística. Em última análise, ainda é importante definir e avaliar a relação existente entre o indivíduo e a comunidade de fala. De fato, todos os valores presentes numa comunidade de fala são compartilhados igualmente por todos os indivíduos? Se não, qual a consequência para o entendimento da gramática da comunidade de fala? Sobre essa questão, Clark (no prelo), a partir de uma abordagem da variação do ponto de vista da Linguística Cognitiva e dos Modelos Baseados no Uso, que estabelece o pressuposto de que a gramática do indivíduo é abstraída principalmente da sua experiência linguística, propõe que nenhum indivíduo vai abstrair a mesma gramática, ou em suas palavras “no two speakers will share the same grammar”. No entanto, embora haja questões importantes levantadas nos Modelos baseados no Uso quanto à arquitetura da gramática e a acomodação dos fatos variáveis, considero que a anulação ou exclusão da estrutura da sociedade no modelo é problemática. De fato, as experiências individuais podem ser diferentes numa sociedade, mas elas não são apenas experiências individuais. São, sobretudo, experiências dos indivíduos que se dão em um determinado contexto histórico e social, que levará à convergências, mas não necessariamente a homogeneidades. Assim, é possível esperar mais de uma gramática numa mesma comunidade de fala, se há, nessa comunidade de fala, discrepâncias profundas entre os diversos grupamentos sociais



que a compõem. Experiências individuais podem ser distintas, mas são experiências sociais e não estritamente individuais. Assim, sempre haverá alguma convergência, uma vez que o indivíduo não é autônomo em relação à estrutura social em qualquer aspecto, não só o linguístico. É exatamente a proposta inaugurada por Labov, nos estudos de variação sociolinguística, que procura capturar essa dimensão da gramática do indivíduo através da relação entre sociedade e conhecimento linguístico.

Com relação ao item lexical, assume-se, na maioria dos trabalhos de variação sociolinguística, uma abordagem neogramática em relação à variação fonológica, a de que o segmento é a unidade da mudança e que a motivação fonética exclui a motivação lexical, muito embora haja evidências de efeito lexical em diversos processos de mudança (Bybee, 2002). Labov (1981) procura equacionar a tensão segmento/palavra levantando como hipótese que determinados tipos de processos tendem a ser do tipo neogramático e outros do tipo difusão lexical. Já Oliveira (1991) defende que todas as mudanças se propagam através das palavras.

Do ponto de vista metodológico essa discussão esbarra no fato de que não é possível avaliar o efeito da palavra usando o Varbrul/Goldvarb por duas razões. Uma razão diz respeito à configuração do arquivo de entrada para a análise estatística. Embora, no Goldvarb não haja limite de fator por grupo, há uma limitação estabelecida em função da configuração do arquivo de entrada, que utiliza somente um caracter por coluna (código) e a quantidade de caracteres disponíveis no teclado para dar conta de uma variável independente como item lexical. A outra razão é que, no Varbrul/Goldvarb, os grupos de fatores são concebidos como independentes (*fixed effect*) sempre. Portanto, um estudo de uma variável fonológica (ou sociofonética) que leve em conta o efeito do item lexical, mesmo que pudesse contornar o problema do número de caracteres para codificações de itens, esbarraria no fato de que a rodada trataria da mesma maneira grupos de fatores que se relacionam com características do segmento no item, como por exemplo – posição do segmento no item, se medial e final, no caso de uma coda, se em sílaba tônica ou átona, ou até mesmo o tamanho do item, e o item propriamente dito. Além disso, a observação do efeito da palavra também fica prejudicada com um procedimento metodológico bastante utilizado nos trabalhos, o da exclusão, nos dados a serem submetidos à

análise estatística, dos itens que tendem a apresentar um percentual alto de uma das variantes. Enfim, uma conjunção de fatores de diversas naturezas – como a adoção implícita ou explícita de que a variação fonológica apresenta características neogramáticas e um instrumental estatístico que corrobora essa posição – não tem permitido um olhar mais aprofundado a uma questão central para a compreensão da mudança sonora, com consequências para o entendimento da implementação da mudança sonora, para a teoria fonológica, e, em última instância, para a arquitetura da gramática. Essa questão é o entendimento das motivações fonéticas e o papel do item lexical na variação e mudança sonora.

Gomes e Melo (2009), utilizando o Rbrul como ferramenta estatística, analisaram o efeito do item no uso variável da coda fricativa na comunidade de fala do Rio de Janeiro, tendo como aplicação a ocorrência da variante velar/glotal, e encontraram, para um determinado grupo social, uma direcionalidade diferente da observada na comunidade de fala. Segundo os trabalhos já realizados para essa comunidade de fala (Guy, 1981, Scherre e Macedo, 2000, Callou e Brandão, 2009), a variante glotal é de baixa ocorrência (abaixo de 7% em todas as amostras estudadas) e, segundo projeta Guy, tenderia a desaparecer devido ao seu caráter estigmatizado e o fato de que as camadas populares passaram a ter mais acesso à educação, e, portanto, aos valores relacionados a essa variante e difundidos pelas diversas instituições sociais. No entanto, no grupo social do estudo de Gomes e Melo (2009), menores infratores cumprindo medida socioeducativa, o percentual médio de ocorrência da glotal foi de 30%, um valor bastante superior ao observado nas amostras dos estudos mencionados. Além do percentual global significativamente superior, os resultados obtidos replicam os condicionamentos encontrados nos trabalhos anteriores. A contribuição do trabalho foi a de poder controlar o papel do item lexical na propagação do que parece ser o desenvolvimento de um padrão sociolinguístico numa direção diferente da que se observa para o restante da comunidade de fala, isto é, na direção da glotal. Ao lado de uma motivação fonética forte (efeito do contexto seguinte) e de outros condicionamentos relacionados à posição da coda na palavra, e ao tamanho do item, os autores puderam identificar que o item propriamente dito aliado à sua alta frequência de ocorrência no corpus também atuam na propagação da velar/glotal.

Esses resultados constituem evidência para a coexistência tanto de motivação fonética quanto propagação da mudança através do léxico e dialogam com modelos de gramática que propõem que a variação tem status representacional. Conforme defende Bybee (2002), efeitos de frequência e de contexto são indicativos de que o uso linguístico tem impacto na representação mental. De acordo com Bybee, mudanças que afetam primeiramente as palavras de alta frequência são o resultado da automação da produção, isto é, resultam da sobreposição entre redução e gestos articulatórios que vêm com a fluência.

A autora propõe um modelo de variação e mudança, baseado em Pierrehumbert (1994), segundo o qual variação e mudança não são externas ao léxico e à gramática, mas inerente a eles. A mudança sonora é capturada na representação do item desde o começo de sua implementação, uma vez que as diversas instâncias de produção da palavra são armazenadas, resultando numa representação detalhada da experiência do falante em produzir e percebê-las, e que vai sendo atualizada em função dessa experiência.

## **CONSIDERAÇÕES FINAIS**

Além de sedimentar um conceito de gramática, o que incorpora a heterogeneidade estruturada, conforme postulada por Weinreich, Labov e Herzog (1968) com o objetivo de explicar a mudança linguística (Labov, 2006), os estudos variacionistas devem contribuir para a discussão sobre a organização do conhecimento linguístico, principalmente no que diz respeito ao status da variação socialmente indexada na gramática, e de que maneira a modelagem teórica captura a variação na arquitetura da gramática proposta (Gomes e Silva, 2004). Na verdade, já há inúmeras evidências do papel da variação na produção e na percepção. Assim, fora da literatura exclusivamente sociolinguística, Munson et al. (2005) defendem a necessidade de a modelagem fonológica capturar quatro tipos de conhecimento fonológico: o conhecimento relativo às características perceptuais e acústicas dos sons da fala (conhecimento perceptual), o relativo às características articulatórias dos sons da fala (conhecimento articulatório), o conhecimento das restrições fonotáticas e de como os sons podem ser combinados em palavras (conhecimento de nível mais alto

fonológico) e o conhecimento de como a variação na fala pode ser utilizada para veicular identidade social (conhecimento socialmente-indexado). Em outras palavras, defende-se, para além da proposta inaugural de Weinreich, Labov e Herzog, centrada na explicação da mudança linguística, uma modelagem teórica que incorpora a variação como central à gramática em termos representacionais, com consequências não só para o entendimento do comportamento do falante típico como também para o de falantes de população atípica.

## REFERÊNCIAS BIBLIOGRÁFICAS

BAILEY, G. *Real and apparent time*. In: J. K. Chambers, P. Trudgill, & N. Schilling-Estes (org). *The handbook of Language Variation and Change*. Oxford: Blackwell, 2003. p. 312-332.

BYBEE, J. *Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change*. *Language variation and change* 14, 2002. p.261-290.

\_\_\_\_\_. *From usage to grammar: the mind's response to repetition*. *Language* 82(4), 2006, p. 711-733.

CALLOU, D. M. I.; BRANDÃO, S. F. *Sobre o /S/ em coda silábica no Rio de Janeiro: falas culta e popular*. In: Salgado, Ana Claudia Peters; Barretto, Mônica M. Guimarães Savedra. (Org). *Sociolinguística no Brasil: uma contribuição dos estudos sobre línguas em/de contato: homenagem ao Prof. Jürgen Heye*. Rio de Janeiro: 7 Letras, 2009. p. 27-34.

CEDERGREEN, H. & SANKOFF, D. *Variable rules: performance as a statistical reflexion of competence*. *Language* 50. 1974. p. 333-335.

CLARK, Lynn. *Dialect data, lexical frequency and usage-based approach*. In de Vogelaar, Gunther & Sieler, Guido (eds.) *Dialects as a Testing Ground for Theories of Language Change*. *Studies in Language Variation*. Amsterdam/Philadelphia: John Benjamins, (no prelo)

GOMES, C.A. & MELO, M. A. L. S. *Developing new patterns in the speech community: a case study about fricative lenition in Brazilian Portuguese*. *NWAV 38 New Ways of Analyzing variation*, University of Ottawa, 2009.

GOMES, C. A.; SILVA, T. C. *Variação linguística: questão antiga e novas perspectivas*. Língua(gem), ILAPEC/Macapá, v. 1, n. 2, 2004. p. 31-41.

GUY, G. R. *Variation in the group and the individual: the case of final stop deletion*. In: Labov, W. (org) *Locating Language in Time and Space*. New York, Academic Press, 1980. p. 1-36.

\_\_\_\_\_. *Linguistic Variation in Brazilian Portuguese: aspects of the phonology, syntax, and language history*. Ph. Dissertation, University of Pennsylvania, 1981. p. 391.

GUY, G. R. & ZILLES, A. *Sociolinguística Quantitativa*. São Paulo: Parábola Editorial, 2007.

JOHNSON, D. E. *Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis*. *Language and Linguistics Compass* 3/1, 2009. p. 359–383.

LABOV, W. *Resolving the neogrammarian controversy*. *Language* 57, 1981. p.267-309.

\_\_\_\_\_. *Contraction, deletion and inherent variability of the English vernacular*. *Language* 45, 1996. p. 715-62.

\_\_\_\_\_. *A sociolinguistic perspective on sociophonetic research*. *Journal of Phonetics* 34(6), 2006. p. 500-515.

MARTINS, R. M. F. *A organização do Componente Fonológico e o Papel do Indivíduo*. Tese de Doutorado, Faculdade de Letras, UFMG, 2007. 275 p.

MUNSON, B., EDWARDS, J., & BECKMAN, M. E. *Phonological knowledge in typical and atypical speech sound development*. *Topics in Language Disorders*, 25, 2005. p. 190-206.

NARO, A. J. *Modelos quantitativos e tratamento estatístico*. In: Mollica, M. C. e Braga, M. L. (org) *Introdução à Sociolinguística*. Rio de Janeiro: Contexto, 2003. p. 15-25.

OLIVEIRA, A. J. *Análise quantitativa no estudo da variação linguística: noções de estatística e análise comparativa ente Varbrul e SPSS*. *Revista de Estudos da Linguagem*. v. 17, n. 2, UFMG, Belo Horizonte. 2009. p. 93-119.

OLIVEIRA, Marco Antonio de. *The Neogrammarian Controversy Revisited*. International Journal of The Sociology of Language, v. 89, n. 1, 1991. p. 93-105.

\_\_\_\_\_. *Aspectos da difusão lexical*. Revista de Estudos da Linguagem. Belo Horizonte, Faculdade de Letras da UFMG, v. 1, 1992. p. 31-41.

PAIVA, Maria da Conceição; DUARTE, Maria Eugênia Lamoglia (Orgs.). *Mudança linguística em tempo real*. Rio de Janeiro: Contra-Capa, 2003.

PIERREHUMBERT, J. *Knowledge of Variation*. Papers from the Parasession on Variation. 30<sup>th</sup> Meeting of the Chicago Linguistic society. Chicago: Chicago Linguistic Society, 1994. p.232-256.

ROUSSEAU, P. & SANKOFF, D. *Advances in variable rule methodology*. In: Rousseau, P. & Sankoff, D. (org) *Linguistic variation: models and methods*. New York: Academic Press. 1978. p. 57-69.

SCHERRE, M. M. P. *Pressupostos teóricos e suporte quantitativo*. In: Oliveira e Silva, G. M. & Scherre, M. M. P. (org) *Padrões Sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Tempo Brasileiro, Rio de Janeiro. 1996. p. 27-36.

SCHERRE, M. M. P.; MACEDO, A V. T. *Variação e mudança: o caso da pronúncia do s pós-vocálico*. In: MOLLICA, M. C.; MARTELOTTA, M. E. (Org). *Análises linguísticas: a contribuição de Alzira Macedo*. Rio de Janeiro: UFRJ, 2000. p. 52-64.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Empirical foundations for a theory of language change*. In: LEHMANN, W. P.; MALKIEL, Yakov (Eds.). *Directions in Historical Linguistics*. Austin, TX:University of Texas Press, 1968. p. 95-188.