

GREGORY R. GUY e ANA ZILLES



SOCIOLINGÜÍSTICA  
QUANTITATIVA



INSTRUMENTAL DE ANÁLISE

$\pi$   
Parábola

**EDITOR:** *Marcos Marcionilo*

**CAPA E PROJETO GRÁFICO:** *Andréia Custódio*

**CONSELHO EDITORIAL:** *Ana Stahl Zilles* [Unisinos]

*Carlos Alberto Faraco* [UFPR]

*Egon de Oliveira Rangel* [PUCSP]

*Gilvan Müller de Oliveira* [UFSC, Ipo]

*Henrique Monteagudo* [Univ. de Santiago de Compostela]

*Kanavillil Rajagopalan* [Unicamp]

*Marcos Bagno* [UnB]

*Maria Marta Pereira Scherre* [UFRJ, UnB]

*Rachel Gazolla de Andrade* [PUC-SP]

*Salma Tannus Muchail* [PUC-SP]

*Stella Maris Bortoni-Ricardo* [UnB]

**CIP-BRASIL. CATALOGAÇÃO NA FONTE**  
**SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ**

---

G998a

Guy, Gregory Riordan

Sociolinguística quantitativa - instrumental de análise / Gregory R. Guy e Ana Zilles. - São Paulo : Parábola Editorial, 2007.

(Lingua[gem]; 23)

Anexos

Inclui bibliografia

ISBN 978-85-88456-45-1

1. Linguagem e línguas - variação - Pesquisa. 2. Linguística - Pesquisa. 3. Linguística - Pesquisa - Metodologia. 4. Pesquisa quantitativa. 5. Sociolinguística - Pesquisa. I. Zilles, Ana Maria Stahl. II. Título. III. Série.

06-3564

CDD:401.9

CDU81:27

---

Direitos reservados à

**PARÁBOLA EDITORIAL**

Rua Clemente Pereira, 327 - Ipiranga

04216-060 São Paulo, SP

Fone: [11] 6914-4932 | Fax: [11] 6215-2636

home page: [www.parabolaeditorial.com.br](http://www.parabolaeditorial.com.br)

e-mail: [parabola@parabolaeditorial.com.br](mailto:parabola@parabolaeditorial.com.br)

---

Todos os direitos reservados. Nenhuma parte desta obra pode ser reproduzida ou transmitida por qualquer forma e/ou quaisquer meios (eletrônico ou mecânico, incluindo fotocópia e gravação) ou arquivada em qualquer sistema ou banco de dados sem permissão por escrito da Parábola Editorial Ltda.

---

ISBN: 978-85-88456-45-1

© do texto: Gregory R. Guy, Ana Maria Stahl Zilles

© desta edição: Parábola Editorial, São Paulo, setembro de 2007

# SUMÁRIO

Lista de tabelas, **7**

Lista de figuras, **8**

Siglas, **8**

**Apresentação, 9**

*Maria Marta Pereira Scherre, Anthony Julius Naro*

**Introdução, 13**

*Ana Zilles*

## **PARTE I. FUNDAMENTAÇÃO**

**1. Introdução à análise quantitativa da variação lingüística**

*Gregory R. Guy* [Tradução de Ana Zilles e Leonardo Z. Maya.], **19**

**2. Varbrul: análise avançada**

*Gregory R. Guy* [Tradução de Ana Zilles], **47**

## **PARTE II. GUIA PRÁTICO PARA O USO DA METODOLOGIA QUANTITATIVA EM SOCIOLINGÜÍSTICA**

1. Por que realizar análises quantitativas?, **73**

2. Quais os tipos de variáveis com que trabalhamos em sociolingüística?, **74**

3. Qual a utilidade de calcular proporções e percentagens?, **80**

4. O que é significância estatística?, **85**

5. Qual a diferença entre análises univariadas e multivariadas?, **98**

6. O que são modelos quantitativos?, **101**

7. O que é e por que usar o pacote de programas estatísticos Varbrul?, **105**

8. Como se preparar para uma análise com o Varbrul?, **108**

9. Que problemas da amostra devem receber atenção especial do pesquisador?, **126**

10. O que é um sistema analítico?, **135**

11. Como definimos a variável dependente?, **141**

12. Como começar a análise estatística com o Varbrul?, **149**

13. O que são nocautes?, **158**

14. Como eliminamos nocautes?, **159**

15. O que são rodadas *step-up/step-down* e *one level*?, **164**

16. Quando se usa um ou outro tipo de rodada (um nível ou *step-up/step-down*)?, **168**

17. Como proceder quando há muitos grupos no início?, **170**
18. Como é que se aperfeiçoa a análise quantitativa?, **174**
19. Como decidir manter ou retirar grupos de fatores da análise?, **182**
20. Que princípios embasam a amalgamação de fatores?, **188**
21. Como e quando se realiza teste de significância estatística dentro de grupos de fatores?, **190**
22. O que é convergência e o que significa?, **198**
23. Como fazer comparações entre rodadas usando diferentes grupos de fatores, ou em casos de recodificações complicadas (por ex., segmentos *versus* traços)?, **201**
24. Como apresentar os resultados?, **205**
25. Como interpretar os resultados numéricos?, **211**
26. Deve-se ou pode-se apresentar resultados sem significância estatística?, **214**
27. A análise com o Varbrul deve proceder do mesmo modo com fatores sociais e fatores lingüísticos?, **218**
28. Como construir e interpretar gráficos relacionados com os resultados?, **222**
29. Pode-se usar o Varbrul para analisar outros tipos de dados?, **224**
30. Quando se escolhe a variante designada como 'aplicação', qual o efeito sobre os pesos calculados pelo *Varbrul*?, **228**

#### **Referências bibliográficas, 231**

#### **Anexo I — Tabela de valores de qui-quadrado, 234**

#### **Anexo II — Exemplos de sistemas analíticos sob a forma de arquivos de especificação de fatores, 235**

#### **Glossário, 239**

# LISTA DE TABELAS

## PARTE I

### Capítulo 1

TABELA 1: Uso de pronúncias consonantais (r-1) de /r/ pós-vocálico na cidade de Nova York (exemplo hipotético)

TABELA 2: Efeito dos fatores morfológicos no apagamento de -t,d (Guy & Boyd, 1990:7)

### Capítulo 2

TABELA 1. Grupos de fatores não-ortogonais: “o pior caso”

TABELA 2. Espanhol de Porto Rico: apagamento do -s plural; efeito da posição no *SN* e da presença de marca precedente (cf. Poplack, 1980)

TABELA 3. Grupos de fatores não-ortogonais: subcategorização

TABELA 4. Assimetria na distribuição dos dados (número de ocorrências)

TABELA 5. Distribuição equilibrada dos dados (número de ocorrências)

TABELA 6.- Percentagens de aplicação da regra para os conjuntos de dados de teste

TABELA 7.- Resultados dos conjuntos de dados de teste, mostrando o efeito da crescente assimetria na distribuição dos dados

TABELA 8. Português brasileiro: concordância no *SN*; efeito do local de informação adicional de plural (cf. Guy, 1981)

TABELA 9. Português brasileiro: concordância sujeito-verbo; efeito da posição do sujeito (cf. Guy, 1981)

TABELA 10. Português brasileiro: apagamento de -s final; efeito da consoante seguinte (probabilidades de apagamento, cf. Guy, 1981)

TABELA 11. Português brasileiro: apagamento de -s final; número de ocorrências por consoante seguinte (cf. Guy, 1981)

## PARTE II

TABELA 1. Orações relativas em dois falantes de inglês australiano: *wh-* e *that* em posição de sujeito com antecedente [+ humano]

TABELA 2. Orações relativas em dois falantes de inglês australiano: *wh-*, *that* e  $\emptyset$  em posição de objeto com antecedente [+ humano]

TABELA 3. Complementizador *que* no francês de Montréal (de Cedergren & Sankoff, 1974)

TABELA 4. Tabela de contingência para o exemplo do francês de Montréal

TABELA 5. Exemplos de Tabelas de contingência mostrando diferentes graus de associação entre idade dos falantes e uma variável lingüística hipotética

TABELA 6. Valores esperados para o exemplo do francês de Montréal, de acordo com a hipótese nula

TABELA 7. Passos no cálculo do qui-quadrado

TABELA 8. Cruzamento de tipo de sujeito e contexto fonológico precedente em inglês afro-americano; número total de ocorrências (cf. Weldon 2003)

TABELA 9. Grupos de fatores sociais significativos para o uso de *a gente* em Porto Alegre na década de 1990: N, % e pesos (Banco de Dados Varsul)

TABELA 10. Freqüência e probabilidade de concordância verbal, segundo o grupo de fatores posição do sujeito em relação ao verbo (Monguilhott & Coelho, 2002:199)

TABELA 11. Ordem vs (% e peso) em relação à estrutura do sn em quatro cidades do RS, dados do Varsul (Zilles, 2000:87)

TABELA 12. Distribuição (N e %) de formas variáveis de concordância com a 1ª pessoa do plural (zero *versus* —mo e —mos) em POA, SBO e PAN, dados do Varsul/ RS (cf. Silva & Zilles, 2002)

## LISTA DE FIGURAS

### PARTE I

#### Capítulo 1

Figura 1. Estratificação do (r) por classe na cidade de Nova York (segundo Labov 1966, 160)

Figura 2. Probabilidade de apagamento de -t,d em verbos semifracos por idade (Guy & Boyd, 1990, 8)

Figura 3. Mapa dialetal quantitativo (Chambers & Trudgill 1980, 130)

Figura 4. Um mapa de componentes principais (Horvath & Sankoff 1987, 190)

#### Capítulo 2

Figura 1. Estratificação social de (r) na cidade de Nova York (cf. Labov, 1966, 240)

### PARTE II

Figura 1. Gênero *versus* nível de escolaridade no uso de *a gente* nos anos 1990; Porto Alegre, Banco de Dados Varsul (cf. Zilles, 2005).

Figura 2. Percentagem de formas variáveis de concordância com a 1ª pessoa do plural (zero *versus* -mo e -mos) em Porto Alegre, São Borja e Panambi, dados do Varsul/RS (cf. Silva e Zilles, 2002)

## SIGLAS

### IAQ

Sigla para o título do artigo “Introdução à análise quantitativa da variação lingüística”, de Gregory Guy (“The Quantitative Analysis of Linguistic Variation”, 1993), que integra este volume.

### VAA

Sigla para o título do artigo “Varbrul: análise avançada”, de autoria de Gregory Guy, (“Advanced Varbrul Analysis”, 1988). Este artigo foi traduzido para o português e publicado inicialmente nos *Cadernos de Tradução do Instituto de Letras da UFRGS*, nº 1, em 1998. Também integra este volume.

# APRESENTAÇÃO

---

Maria Marta Pereira Scherre  
Anthony Julius Naro

---

O grande prazer de apresentar o livro de Gregory Guy & Ana Zilles, *Sociolingüística quantitativa — Instrumental de análise*, para os pesquisadores e estudantes brasileiros tem duas razões principais: uma é de contextualização histórica do nosso contato com os dois autores; a outra tem a ver com o conteúdo das duas partes que compõem o livro.

Nosso primeiro contato com Greg se deu na década de 1970, quando a pesquisa variacionista de linha laboviana estava se iniciando no Brasil, com o grupo hoje denominado de Programa de Estudos sobre o Uso da Língua (PEUL), sediado no Departamento de Lingüística e Filologia da Faculdade de Letras da Universidade Federal do Rio de Janeiro. À época, Greg era aluno do curso de doutorado na Universidade da Pensilvânia, sob orientação de William Labov. A convite do projeto Competências Básicas do Português, coordenado por Míriam Lemle, e com a indicação pessoal de Labov, Greg foi o responsável pelo início das análises quantitativas desenvolvidas no Brasil. Foi ele quem nos trouxe as primeiras ferramentas computacionais de análise quantitativa e quem melhor representou a mola propulsora das pesquisas realizadas à época, quando ainda se fazia a entrada de dados por meio de cartões perfurados. Da participação no projeto Competências Básicas, resultou, em 1981, a tese de doutorado de Greg, intitulada *Linguistic Variation in Brazilian Portuguese: Aspects of the Phonology, Syntax, and Language History*, baseada nos dados coletados por aquele projeto e nas análises aqui realizadas. Em sua tese, Greg faz uma análise profunda de fenômenos variáveis do português brasileiro — concordância verbal, concordância nomi-

nal, queda da nasal e do -s em final de palavra. Além de importantes discussões com relação a aspectos da estrutura lingüística envolvendo os fenômenos analisados, Greg coloca em cena de forma instigante a discussão sobre a hipótese da crioulização e descrioulização do português brasileiro. Todos os pontos abordados na tese de doutorado renderam — e ainda rendem — boas e acaloradas discussões na literatura lingüística. Os debates fizeram crescer o nosso conhecimento sobre a estrutura e as origens do português brasileiro.

Nosso primeiro contato com Ana se deu bem mais tarde, no início da década de 1990. Ocorreu em função do crescimento do PEUL como grupo irradiador de experiência com a pesquisa sociolingüística de linha laboviana. Diversos membros do PEUL transferiram o conhecimento acumulado pelo grupo do Rio, colaborando assim para a criação e estabelecimento do Projeto de Variação Lingüística no Sul (VARSUL). O VARSUL utilizou com maestria a experiência recebida desta e de outras fontes e é hoje um grupo consolidado, produzindo pesquisa de ponta na área. Ana faz parte desse grupo, que se envolveu de forma séria e intensa no desenvolvimento de pesquisas variacionistas. O investimento sociolingüístico de Ana se consolidou no pós-doutorado nos Estados Unidos em 2001-2002. Greg, já professor titular na Universidade de Nova York (NYU), foi seu grande *anfitrião*, mas Ana recebeu também pleno apoio de Gillian Sankoff, da Universidade da Pensilvânia. Com seu retorno ao Brasil, Ana presenteia a comunidade acadêmica brasileira com o principal produto prometido ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), um livro em co-autoria com Gregory Guy a respeito da pesquisa sociolingüística quantitativa e, em especial, dos programas Varbrul. É precisamente este livro, *Sociolingüística quantitativa — Instrumental de análise*, que temos a grata satisfação de apresentar.

A comunidade acadêmica brasileira recebe o livro *Sociolingüística quantitativa — Instrumental de análise*, com um conteúdo desejado por muitos, particularmente pelos pesquisadores que desenvolvem pesquisas na linha da sociolingüística laboviana. O livro apresenta duas partes. A primeira parte, que consta da tradução de dois artigos de Greg, abre a discussão e dá respostas e esclarecimentos precisos às dúvidas do imaginá-

rio e da realidade das pesquisas com os dados de variação lingüística. A segunda parte é fruto do trabalho de pós-doutorado de Ana. Trata-se de um texto de Guy & Zilles, em que se busca trazer a público respostas a diversas perguntas que os pesquisadores variacionistas, usuários do aparato teórico-metodológico da sociolingüística laboviana, se fazem no dia-a-dia, no trabalho meticoloso que envolve a análise de grandes massas de dados produzidos em circunstâncias reais de uso. Esse tipo de análise envolve necessariamente a utilização de instrumentos estatísticos. No livro, os autores dialogam com os pesquisadores, apresentando respostas para as dúvidas que povoam a experiência com o tratamento dos dados. E fazem isso de forma interessante, em busca de um equilíbrio entre o teórico e o metodológico. Todas as respostas são formuladas de maneira clara e direta, com explicações completas, expressas de maneira a tornar a compreensão do texto fácil e agradável.

Em síntese, o conteúdo deste livro permitirá ao leitor curioso a resolução de grande parte de suas indagações, tanto em nível teórico/metodológico mais amplo quanto em nível das preocupações diárias da pesquisa real. Permitirá também o compartilhamento das incertezas e pontos problemáticos, que são de todos — de iniciantes e de iniciados — e que, mesmo ainda sem soluções definitivas, fazem crescer a busca de conhecimento da realidade que nos cerca. Vamos à sua leitura, sagaz, perspicaz, instigante e esclarecedora.