



**Vander Paula Viana**

**Verbos modais em contraste**  
**Análise de corpus da escrita de universitários em inglês**

**Dissertação de Mestrado**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Letras da PUC-Rio.

Orientador: Profa. Dra. Lúcia Pacheco de Oliveira

Rio de Janeiro

Abril de 2008



**Vander Paula Viana**

**Verbos modais em contraste**

**Análise de corpus da escrita de universitários em inglês**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Letras da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Profa. Dra. Lúcia Pacheco de Oliveira**

Orientadora  
PUC-Rio

**Profa. Dra. Sonia Zyngier**

UFRJ

**Profa. Dra. Violeta de San Tiago Dantas Barbosa Quental**

PUC-Rio

**Prof. Dr. Paulo Fernando Carneiro de Andrade**

Coordenador Setorial do Centro de Teologia e Ciências Humanas  
PUC-Rio

Rio de Janeiro, 09 de abril de 2008

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e da orientadora.

### Vander Paula Viana

Graduou-se em Letras (Inglês / Literaturas) pela Universidade do Estado do Rio de Janeiro em 2005. Integra o grupo de pesquisa Research and Development in Empirical Studies (REDES-BRA) desde 2003. Atua principalmente nas áreas de língua inglesa e lingüística de corpus, tendo ministrado cursos na confluência destas duas áreas. Faz parte do corpo editorial do *Asian EFL Journal*, *The Linguistics Journal* e *Chinese EFL Journal*, sendo editor associado do último. Publicou oito artigos em periódicos especializados, editou cinco livros e tem oito capítulos de livros publicados em volumes nacionais e internacionais. Participou de aproximadamente 50 eventos, tendo colaborado na organização de oito deles. Trabalha com o ensino de língua inglesa há quase dez anos e exerce o cargo de segundo secretário da Associação de professores de Língua Inglesa do Estado do Rio de Janeiro (APLIERJ) desde 2007.

#### Ficha Catalográfica

Viana, Vander Paula

Verbos modais em contraste: análise de corpus da escrita de universitários em inglês / Vander Paula Viana; orientadora: Lúcia Pacheco de Oliveira. – 2008.

230 f. : il. (color.) ; 30 cm

Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

Inclui bibliografia

1. Letras – Teses. 2. Língua inglesa. 3. Verbos modais. 4. Lingüística de corpus. 5. Redação universitária. 6. Colocação. 7. Padrões de uso. 8. Abordagem textual. I. Oliveira, Lúcia Pacheco de. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Letras. III. Título.

CDD: 800

À minha mãe por tudo e em todos os momentos; por ser, estar e realizar por mim.

## Agradecimentos

À Profa. Dra. Lúcia Pacheco de Oliveira pela orientação acadêmica, pela convivência afetuosa e gentil, pelo incentivo a uma postura autônoma e pela confiança em meu trabalho. Sou igualmente grato aos convites por ela realizados durante o curso de mestrado como, por exemplo, a progressão direta para o doutorado.

Aos alunos de graduação das quatro universidades nas quais a coleta de dados foi realizada (PUC-Rio, UERJ, UFF e UFRJ) por terem aceitado empregar parte de seu tempo para redigirem as redações que tornaram este estudo real.

Aos professores destes alunos nas referidas universidades, que permitiram que a coleta de dados fosse realizada em suas turmas, generosamente cedendo suas aulas para esta finalidade.

Ao Prof. Dr. Tony Berber Sardinha pela acolhida no projeto de compilação do BRICLE (The Brazilian Portuguese Sub-corpus of the International Corpus of Learner English) e à Denise Delegá pelas informações a respeito de coletas anteriormente realizadas.

À Profa. Dra. Tania Shepherd, pela orientação de iniciação científica durante meus estudos de graduação, introduzindo-me à pesquisa acadêmica.

À PUC-Rio pela bolsa de isenção concedida.

Às Profas. Dras. Márcia Martins, Sonia Zyngier e Violeta Quental por terem prontamente aceitado o convite para integrar a banca examinadora desta dissertação.

À Elisa Borges e à Olívia Fialho pelo auxílio bibliográfico em momentos distintos do mestrado.

Às Profas. Dras. Lúcia Pacheco de Oliveira, Margarida Basílio, Maria Carmelita

Dias, Maria do Carmo de Oliveira e Violeta Quental por suas aulas durante o curso de mestrado, fomentando a minha formação acadêmica.

À Danielle Menezes que tive a oportunidade de conhecer melhor e com quem pude estreitar laços de amizade durante o curso, pelo apoio contínuo, pelo imenso carinho e pela disponibilidade constante.

À Luciana Amaral pelo feliz reencontro após anos, pela amizade, pelas longas conversas telefônicas e pelo auxílio em diversos momentos.

À Ana Carolina Gomes e à Christiane Alcântara, amigas desde a graduação, pela companhia afetuosa em momentos imprescindíveis.

Ao grupo de pesquisa REDES-BRA, pelas redes de relacionamento e pela possibilidade de desenvolvimento interpessoal e acadêmico. Dedico um agradecimento especial à Fabiana Fausto, por auxiliar-me a entender a estatística; à Gabriela Marques, pelas constantes trocas de e-mail e informação; à Juliana Jandre, pelo grande apreço a mim dispensado; e à Milena Mendes, por sua companhia única.

À Sonia Zyngier pelo seu pronto auxílio e incentivo desde 2003. O meu obrigado não somente por ter ajudado em minha formação e crescimento acadêmicos de forma espontânea, mas também pelo apoio extra-acadêmico.

À Viviane Viana, minha irmã, pelo carinho constante, pelo incentivo contínuo, pelo auxílio na concretização de planos e pelo generoso desejo do melhor para mim, sempre indicando caminhos e possibilitando os meios para que eles fossem efetivamente seguidos.

À Sandra Viana, minha mãe, que é a mais importante de todas as professoras em minha vida, aquela que me fez optar por esta profissão através de seu exemplo de dedicação ao magistério. Por ter sido e continuar sendo a minha maior referência de afeto, carinho e amabilidade, por ter feito sempre tudo por mim mesmo em condições adversas e por ser a pessoa que amo nesta vida, deixo aqui registrado o meu singelo e afetuoso agradecimento por toda uma vida.

## Resumo

Viana, Vander Paula; Oliveira, Lúcia Pacheco de. **Verbos modais em contraste: análise de corpus da escrita de universitários em inglês**. Rio de Janeiro, 2008. 230p. Dissertação de Mestrado – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

O presente estudo busca investigar o uso de verbos modais na produção escrita de universitários brasileiros em inglês como língua estrangeira (ILE), contrastando-o com o emprego realizado por alunos universitários americanos e britânicos que têm o inglês como primeira língua. Para tanto, lança-se mão do arcabouço teórico da lingüística de corpus (Sinclair, 1991; McEnery & Wilson, 1996; Hunston, 2002; Berber Sardinha, 2004; McEnery, Xiao & Tono, 2006) e de recursos tecnológicos, especialmente o programa computacional *WordSmith Tools* (Scott, 1999), para analisar dois corpora. O corpus de estudo consiste em uma fração do Br-ICLE (*The Brazilian Portuguese Sub-corpus of the International Corpus of Learner English*), coletada em quatro universidades do Estado do Rio de Janeiro, totalizando 51.430 palavras, enquanto o de referência corresponde a uma parte do LOCNESS (*Louvain Corpus of Native English Essays*), somando 165.135 palavras. Adotando uma abordagem textual à escrita (Hyland, 2002) e uma abordagem estatística para quantificar os resultados, a investigação tem como foco os verbos modais centrais (Biber et al., 1999), a saber, ‘can’, ‘could’, ‘may’, ‘might’, ‘must’, ‘shall’, ‘should’, ‘will’ e ‘would’. Os resultados indicam que os falantes brasileiros de inglês empregam menos verbos modais do que seus pares americanos e britânicos. Mais especificamente, nota-se que ‘will’ e ‘would’ são usados com frequência significativamente menor na escrita em ILE de brasileiros. Diferenças entre os modais também são encontradas quando se analisa o padrão de uso destes verbos. Os resultados da pesquisa evidenciam, de certa forma, a diferença na produção escrita dos dois grupos estudados, o que pode contribuir para o ensino de língua inglesa no contexto brasileiro.

## Palavras-chave

língua inglesa; verbos modais; lingüística de corpus; redação universitária; colocação; padrões de uso; abordagem textual

## Abstract

Viana, Vander Paula; Oliveira, Lúcia Pacheco de. **Modal verbs in contrast: a corpus-based analysis of undergraduate writing in English.** Rio de Janeiro, 2008. 230p. M.A. Thesis – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

The present study aims at investigating and contrasting the use of modal verbs in the writing of Brazilian undergraduates majoring in English as a foreign language (EFL), contrasting it to that of British and American university students who speak English as a first language. To this end, it makes use of both the theoretical background of Corpus Linguistics (Sinclair, 1991; McEnery & Wilson, 1996; Hunston, 2002; Berber Sardinha, 2004; McEnery, Xiao & Tono, 2006) and of technological resources, especially the computer program *WordSmith Tools* (Scott, 1999), to analyze two corpora. The research corpus consists of a section of the Br-ICLE (*The Brazilian Portuguese Sub-corpus of the International Corpus of Learner English*) which has been collected in four universities in the city of Rio de Janeiro, totaling 51,430 words, while the reference one corresponds to a part of the LOCNESS (*Louvain Corpus of Native English Essays*), accounting for 165,135 words. Adopting a textual approach to writing (Hyland, 2002) and using statistical tests to quantify the results, the investigation focuses on central modal verbs (Biber et al., 1999), namely, ‘can’, ‘could’, ‘may’, ‘might’, ‘must’, ‘shall’, ‘should’, ‘will’ and ‘would’. Results indicate that the Brazilian speakers of English make less use of modal verbs than their American and British counterparts. More specifically, it is observed that both ‘will’ and ‘would’ present significantly lower frequencies in Brazilian EFL writing. Differences among modals are also found when the patterns of use of these verbs are taken into account. Research results show that there are certain differences in the written production of the two groups, which may contribute to the teaching of English in the Brazilian context.

## Keywords

English language; modal verbs; corpus linguistics; university-student essays; collocation; patterns of use; textual approach

## Sumário

1 Introdução	19
2 Lingüística de corpus	23
2.1. Corpus	23
2.1.1. Tamanho	31
2.1.2. Tipologia	34
2.2. Padrões de uso	35
2.3. Resumo	41
3 Escrita	42
3.1. Abordagens à escrita	43
3.2. Escrita na universidade	51
3.2.1. Corpora de ILE e abordagem textual	56
3.3. Resumo	58
4 Verbos modais	60
4.1. Descrições gramaticais	61
4.1.1. Perspectiva funcional	61
4.1.2. Perspectiva pedagógica	63
4.1.3. Perspectiva tradicional	65
4.1.4. Perspectiva baseada em corpus	67
4.2. Foco analítico	71
4.3. Estudos anteriores	72
4.3.1. Gêneros e variedades de IL1	73
4.3.2. Estudos contrastivos	79
4.3.3. Pesquisa e ensino de ILE	88
4.4. Resumo	92
5 Dados coletados	94
5.1. Corpora de pesquisa	94

5.1.1. Br-ICLE	94
5.1.1.1. Princípios de coleta de dados	96
5.1.1.2. Coletas de dados	98
5.1.1.3. Condições de realização das redações	101
5.1.1.4. Formatação do corpus	102
5.1.1.5. Caracterização do corpus	104
5.1.2. LOCNESS	108
5.1.2.1. Recorte do corpus	110
5.1.2.2. Princípios de coleta de dados	111
5.1.2.3. Coletas de dados	113
5.1.2.4. Condições de realização das redações	113
5.1.2.5. Formatação do corpus	114
5.1.2.6. Caracterização do corpus	115
5.2. Participantes	118
5.2.1. Br-ICLE	118
5.2.2. LOCNESS	123
5.3. Resumo	124
6 Instrumentos	126
6.1. <i>WordSmith Tools</i>	127
6.1.1. <i>WordList</i>	129
6.1.2. <i>KeyWords</i>	133
6.1.3. <i>Concord</i>	135
6.2. SPSS	138
6.2.1. ANOVA	141
6.3. <i>Excel</i>	141
6.3.1. Informação mútua	142
6.3.2. Escore T	143
6.4. Resumo	143
7 Análise de dados	145
7.1. Análise freqüencial no Br-ICLE	146
7.1.1. Verbos modais	147
7.1.2. Referência temporal	152

7.1.3. Grupos de significados	154
7.2. Análise freqüencial no LOCNESS	155
7.2.1. Verbos modais	156
7.2.2. Referência temporal	159
7.2.3. Grupos de significados	161
7.3. Análise freqüencial entre corpora	163
7.3.1. Grupos de significados	165
7.3.2. Referência temporal	167
7.3.3. Verbos modais	168
7.3.4. Análise estatística	175
7.4. Análise dos padrões de uso	176
7.4.1. Padrões de ' <i>can</i> '	177
7.4.2. Padrões de ' <i>may</i> '	180
7.4.3. Padrões de ' <i>must</i> '	182
7.4.4. Padrões de ' <i>should</i> '	183
7.4.5. Padrões de ' <i>will</i> '	184
7.4.6. Padrões de ' <i>would</i> '	185
7.5. Particularidades de uso	187
7.6. Resumo	189
8 Considerações finais	193
9 Referências bibliográficas	201
10 Anexos	212
10.1. Página de inserção de dados no Br-ICLE	213
10.2. Página de registro de dados no Br-ICLE	215
10.3. Questionário	216
10.4. Estatísticas do Br-ICLE	217
10.5. Estatísticas das redações da PUC-Rio	218
10.6. Estatísticas das redações da UERJ	219
10.7. Estatísticas das redações da UFF	220
10.8. Estatísticas das redações da UFRJ	221
10.9. Estatísticas das redações do LOCNESS	222

10.10. Estatísticas das redações de participantes americanos (LOCNESS)	223
10.11. Estatísticas das redações de participantes britânicos (LOCNESS)	224
10.12. Freqüência relativa de verbos modais por corpus por grupos de 100.000 itens	225
10.13. Freqüência relativa de verbos modais por redação em cada corpus por grupos de 1.000 palavras	226
10.14. Resultados do teste ANOVA para a relação contrastiva de verbos modais nos dois corpora	227
10.15. Resultados do teste ANOVA para o uso contrastivo de cada verbo modal nos corpora investigados	228
10.16. Freqüência relativa de verbos lexicais em sintagmas modalizados no LOCNESS	229

## Lista de figuras

Figura 1: Elementos definidores do termo 'corpus' encontrados na literatura	30
Figura 2: Proposta classificatória de corpora com base em tamanho (Berber Sardinha, 2002, p. 119)	33
Figura 3: Produção lingüística à luz da lingüística de corpus (com base em Sinclair, 1991)	36
Figura 4: Escrita conforme sua natureza (Grabe & Kaplan, 1996)	44
Figura 5: Abordagens à escrita e suas respectivas perspectivas (Hyland, 2002)	46
Figura 6: Tipos de investigação textual (Hyland, 2002)	50
Figura 7: Modalidade e suas ramificações na GSF (cf. Halliday, 1994 [1985])	62
Figura 8: Taxonomia de significados de verbos modais proposta por Swan (1995 [1980])	64
Figura 9: Taxonomia de verbos modais proposta por Biber et al. (1999)	67
Figura 10: Distribuição de verbos modais centrais por referência temporal (cf. Biber et al., 1999)	69
Figura 11: Distribuição de alunos por universidade (Br-ICLE)	99
Figura 12: Exemplo de redação em formato digital (091P4cSc)	103
Figura 13: Parágrafo original (041E4oSn)	103
Figura 14: Parágrafo após formatação (041E4oSn)	103
Figura 15: Distribuição do número de palavras no corpus por universidade	105
Figura 16: Média de palavras por redação em cada universidade	106
Figura 17: Distribuição de redações por tópicos	107
Figura 18: Distribuição de redações no Br-ICLE por época de coleta	108
Figura 19: Distribuição de alunos por universidade (LOCNESS / Dados americanos)	113

Figura 20: Distribuição do número de palavras no corpus por nacionalidade	116
Figura 21: Média de palavras por redação por nacionalidade	116
Figura 22: Distribuição de redações no LOCNESS por época de coleta	117
Figura 23: Distribuição dos participantes por faixa etária	119
Figura 24: Conhecimento de línguas estrangeiras diferente do inglês	120
Figura 25: Anos de estudo de inglês na universidade	122
Figura 26: Lista de palavras por ordem alfabética	130
Figura 27: Lista de palavras por ordem de freqüência	131
Figura 28: Tela de estatísticas do <i>WordList</i>	131
Figura 29: Palavras-chave positivas características do LOCNESS	134
Figura 30: Linhas de concordância para ' <i>should</i> ' no Br-ICLE	136
Figura 31: Lista de colocados para ' <i>should*</i> ' no Br-ICLE	137
Figura 32: Verbos modais no Br-ICLE	147
Figura 33: Referência temporal associada a verbos modais no Br-ICLE	152
Figura 34: Distribuição de verbos modais por significados no Br-ICLE	154
Figura 35: Freqüência normalizada de verbos modais no LOCNESS	156
Figura 36: Referência temporal associada a verbos modais no LOCNESS	159
Figura 37: Distribuição de verbos modais por significados no LOCNESS	162
Figura 38: Comparação freqüencial entre corpora por significados	166
Figura 39: Comparação freqüencial entre corpora por referência temporal	167
Figura 40: Comparação freqüencial entre corpora por verbos modais	168

## Lista de tabelas

Tabela 1: Distribuição de alunos por curso de graduação	121
Tabela 2: Distribuição de alunos por universidade (Br-ICLE)	121
Tabela 3: Quantidade de anos cursados na universidade	121
Tabela 4: Significância estatística por verbo modal nos corpora investigados	175

## Lista de abreviaturas e siglas

ANOVA	<i>Analysis of variance</i>
BNC	<i>British National Corpus</i>
BoE	<i>Bank of English</i>
BPI	Bacharelado em Português, Inglês e Literaturas correspondentes
Br-ICLE	<i>The Brazilian Portuguese Sub-corpus of ICLE</i>
BTI	Bacharelado em Tradução (Português / Inglês)
CAp-UFRJ	Colégio de Aplicação da UFRJ
CECL	<i>Centre for English Corpus Linguistics</i>
FROWN	<i>Freiburg-Brown Corpus of American English</i>
GSF	Gramática Sistemico-Funcional
ICAME	<i>International Computer Archive of Modern and Medieval English</i>
ICE	<i>International Corpus of English</i>
ICLE	<i>International Corpus of Learner English</i>
IL1	Inglês como primeira língua
IL2	Inglês como segunda língua
ILE	Inglês como língua estrangeira
IND	<i>Indiana University</i>
KWIC	<i>Key word in context</i>
L1	Primeira língua
L2	Segunda língua
LA	Linguística aplicada
LE	Língua estrangeira
LOB	Lancaster-Oslo / Bergen (Corpus)
LOCNESS	<i>Louvain Corpus of Native English Essays</i>
LPI	Licenciatura em Português, Inglês e Literaturas correspondentes
MIC	<i>University of Michigan</i>
MRQ	<i>Marquette University</i>
NURC-RJ	Projeto da Norma Urbana Oral Culta do Rio de Janeiro
PALC	<i>Practical Applications in Language and Computers</i>
PRB	<i>Presbyterian College</i>
PUC-Rio	Pontifícia Universidade Católica do Rio de Janeiro

PUC-SP	Pontifícia Universidade Católica de São Paulo
REDES	<i>Research and Development in Empirical Studies</i>
SCU	<i>University of South Carolina</i>
SPSS	<i>Statistical Package for the Social Sciences</i>
TALC	<i>Teaching and Language Corpora</i>
UERJ	Universidade do Estado do Rio de Janeiro
UFF	Universidade Federal Fluminense
UFRJ	Universidade Federal do Rio de Janeiro
UFSC	Universidade Federal de Santa Catarina
USP	Universidade de São Paulo

*The more we learn about the world and the deeper our learning, the more conscious, specific and articulate will be our knowledge of what we do not know, our knowledge of our ignorance. For this, indeed, is the main source of our ignorance – the fact that our knowledge can be only finite, while our ignorance must necessarily be infinite.<sup>1</sup>*

(Popper, 2002 [1963], p. 38)

---

<sup>1</sup> Tradução livre para o português: “Quanto mais aprendemos a respeito do mundo e mais profundo o nosso aprendizado, mais consciente, específico e articulado será o nosso conhecimento do que não sabemos, nosso conhecimento da nossa ignorância. Isto é, na verdade, a principal origem de nossa ignorância – o fato de que nosso conhecimento só pode ser finito enquanto nossa ignorância deve necessariamente ser infinita”.