

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**Faculdade de Letras**

**PRISCILLA TULIPA DA COSTA**

**OS *PHRASAL VERBS* NA PRODUÇÃO ESCRITA DE APRENDIZES  
BRASILEIROS DE INGLÊS COMO SEGUNDA LÍNGUA: UMA  
ANÁLISE BASEADA EM *CORPUS***

**Belo Horizonte**  
**2017**

**PRISCILLA TULIPA DA COSTA**

**OS *PHRASAL VERBS* NA PRODUÇÃO ESCRITA DE APRENDIZES  
BRASILEIROS DE INGLÊS COMO SEGUNDA LÍNGUA: UMA  
ANÁLISE BASEADA EM *CORPUS***

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Estudos Linguísticos.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Estudos da Língua em Uso.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Adriana Maria Tenuta de Azevedo.

C837p

Costa, Priscilla Tulipa da.

Os *phrasal verbs* na produção escrita de aprendizes brasileiros de inglês como segunda língua [manuscrito] : uma análise baseada em corpus / Priscilla Tulipa da Costa. – 2017.

121 f., enc. : il., grafs., color.

Orientadora: Adriana Maria Tenuta de Azevedo.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Estudos da Língua em Uso.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 91-95.

Apêndices: f. 96-112.

Anexos: f. 113-120.

1. Língua inglesa – Estudo e ensino – Falantes de português – Teses. 2. Linguística de corpus – Teses. 3. Língua inglesa – Verbos – Teses. 4. Aquisição da segunda linguagem – Teses. 5. Redação acadêmica – Teses. I. Azevedo, Adriana Maria Tenuta de. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD : 420.7



UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS



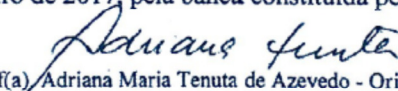
## FOLHA DE APROVAÇÃO

**OS PHRASAL VERBS NA PRODUÇÃO ESCRITA DE APRENDIZES  
BRASILEIROS DE INGLÊS COMO SEGUNDA LÍNGUA: UMA  
ANÁLISE BASEADA EM CORPUS**

**PRISCILLA TULIPA DA COSTA**

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGÜÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGÜÍSTICOS, área de concentração LINGÜÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Estudos da Língua em Uso.

Aprovada em 23 de fevereiro de 2017, pela banca constituída pelos membros:

  
Prof(a) Adriana Maria Tenuta de Azevedo - Orientador  
Universidade Federal de Minas Gerais

  
Prof(a) Bárbara Malveira Orfanó  
UFMG

  
Prof(a) Guilherme Fromm  
UFU

Belo Horizonte, 23 de fevereiro de 2017.

*Aos meus amados pais e irmãos, que sempre  
estiveram ao meu lado, me apoiando no  
desenvolvimento de meus projetos  
pessoais e profissionais.*

## AGRADECIMENTO

Por mais esta importante etapa concluída agradeço primeiramente a Deus, amigo inseparável que cuida de mim e que me proporciona, a cada dia, a oportunidade de vivenciar experiências tão maravilhosas. À minha família e ao Lucas, sempre presentes, e aos amigos, agradeço por me acompanharem e me apoiarem em minhas lutas, por compreenderem minhas correrias e, agora, por compartilharem de minha vitória.

À querida e competente professora Dr<sup>a</sup>. Adriana Maria Tenuta de Azevedo, agradeço por ter me orientado em mais esse projeto, por ter me ensinado bastante, me motivado a continuar desenvolvendo atividades de pesquisa, por me proporcionar espaço produtivo para realizar meus estudos e, principalmente, por acreditar e confiar em mim, por me incentivar e por ser tão compreensiva e paciente em relação às dificuldades que enfrentei ao longo do percurso acadêmico.

Agradeço também à Universidade Federal de Minas Gerais, que além de oferecer gratuitamente à população a oportunidade de realizar cursos de especialização, ainda coloca à disposição dos educandos uma equipe de profissionais competente, atualizada e focada na efetivação do conhecimento. Ao POSLIN/FALE - UFMG, pela presteza em sempre me ajudar a resolver minhas questões acadêmicas.

Aos colegas, agradeço pela companhia e agradáveis momentos de descontração e motivação na construção do conhecimento. A todos os professores, por dividirem comigo seu conhecimento e por me darem o suporte necessário em todas as etapas desta jornada.

Por fim, agradeço ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), por financiar essa pesquisa e permitir que eu tivesse recursos para conduzi-la.

*"It is often forgotten that (dictionaries) are artificial repositories, put together well after the languages they define. The roots of language are irrational and of a magical nature."*

*Jorge Luis Borges,  
Prólogo de "El otro, el mismo."*

## RESUMO

Esta pesquisa tem por objetivo investigar o uso dos *phrasal verbs* em produções escritas de aprendizes brasileiros da língua inglesa, estabelecendo uma comparação com aprendizes nativos. A base para o estudo está na importância dessas estruturas gramaticais para a proficiência na língua inglesa e nas características que as tornam difíceis de serem assimiladas por aprendizes não nativos do idioma. O trabalho, de caráter descritivo, tem fundamento na Linguística de *Corpus* e foi conduzido a partir de uma metodologia composta por exames quantitativos e qualitativos desses verbos. Como ferramentas, foram usados os *corpora* BR-ICLE, para estudo, e LOCNESS e BNC, para referência. Os dados foram tratados e analisados com o auxílio do *software* AntConc. Os resultados mostraram que os *phrasal verbs* são, em geral, pouco utilizados em textos acadêmicos quando comparados a outros tipos de *multi-word verbs*, entretanto, nota-se também que algumas das combinações produzidas tornaram-se típicas desse tipo de produção textual, o que sugere que os *phrasal verbs* têm ganhado espaço em textos de caráter mais formal. Além disso, o estudo também mostrou que há semelhanças e diferenças entre os grupos analisados no emprego dessas expressões, que os *phrasal verbs* mais utilizados pelos não nativos são também, em sua maioria, os mais utilizados pelos não nativos, mas que os aprendizes brasileiros, muitas vezes, aplicam esses verbos de forma diferente dos nativos, sugerindo desacordo em relação aos padrões da língua inglesa. A pesquisa indicou ainda a necessidade de estudos mais aprofundados, especialmente no campo da semântica, para melhor descrição do fenômeno.

**Palavras-chaves:** Linguística de *Corpus*. *Phrasal verbs*. *Corpus* de aprendizes. Escrita acadêmica.



## ABSTRACT

This research aims to investigate the use of phrasal verbs in academic essays written by Brazilian learners of English, in comparison with native speakers. The basis for the study is the importance of these grammatical structures to the proficiency in English, as well as the characteristics that make them difficult to be assimilated by non-native learners of the language. The work, which is descriptive, is based on Corpus Linguistics and was led by means of a methodology composed by quantitative and qualitative analysis of the phrasal verbs. As the learners' corpus, we used the Br-ICLE. LOCNESS and BNC were used as control corpora. The data were processed and analyzed with the support of AntConc software. The results showed that phrasal verbs are, in general, less used in academic texts when compared to other types of multi-word verbs; however, it was also noted that some of the combinations produced have become typical of this kind of textual genre, what suggests that the use of phrasal verbs are increasing in formal texts. In addition, the study also showed that there are similarities and differences between the phrasal verb production in the groups analyzed. It was observed that the most used phrasal verbs by non-native are the same most used by native students; however, Brazilian students often use phrasal verbs differently from natives, suggesting a disagreement to the English language patterns. The research also indicates the need for further studies, especially in the field of semantics, for a better description of the phenomenon.

**Keywords:** Corpus Linguistics. Phrasal verbs. Learner corpus. Academic writing.

## LISTA DE FIGURAS

<b>FIGURA 1</b> - <i>Interface</i> do programa AntConc 3.4.4w .....	41
<b>FIGURA 2</b> - Função “ <i>Sort</i> ” em linha de concordância do AntConc 3.4.4w.....	42
<b>FIGURA 3</b> - A ferramenta <i>File View Tool</i> no AntConc 3.4.4w.....	43
<b>FIGURA 4</b> - Lista de palavras do Br-ICLE gerada na função <i>Wordlist</i> do AntConc .....	45
<b>FIGURA 5</b> - Lista de palavras do LOCNESS gerada na função <i>Wordlist</i> do AntConc.....	48
<b>FIGURA 6</b> - Exemplo de análise da partícula <i>up</i> usando a ferramenta <i>Clusters</i> .....	52
<b>FIGURA 7</b> - Exemplo de análise da partícula <i>out</i> usando a ferramenta <i>Concordance</i> .....	53

## LISTA DE GRÁFICOS

- GRÁFICO 1** - Partículas adverbiais formando *phrasal verbs* no *corpus* de estudo ..... 60
- GRÁFICO 2** - Partículas adverbiais formando *phrasal verbs* no *corpus* de referência..... 61
- GRÁFICO 3** - As dez partículas adverbiais mais produtivas no CE e o número de verbos com os quais combinam para formar *phrasal verbs*..... 62
- GRÁFICO 4** - Número de verbos lexicais individuais formando PVs com partículas adverbiais..... 64
- GRÁFICO 5** - Os dez verbos mais produtivos no CE e o número de partículas com as quais combinam para formar *phrasal verbs* (ocorrência por 100 mil) ..... 65
- GRÁFICO 6** - Os dez verbos mais produtivos no CR e o número de partículas com as quais combinam para formar *phrasal verbs* (ocorrência por 100 mil) ..... 65
- GRÁFICO 7** - Frequência de *phrasal verbs* nos *corpora* (ocorrência por 100 mil)..... 67

## LISTA DE QUADROS

QUADRO 1 - Classificação dos <i>phrasal verbs</i> .....	24
QUADRO 2 - Testes e/ou critérios de distinção dos verbos frasais .....	30
QUADRO 3 - Tipologia empregada na Linguística de <i>Corpus</i> .....	36
QUADRO 4 - Títulos e/ou temas sugeridos pelo projeto ICLE para redações acadêmicas...	44
QUADRO 5 - Temas/assuntos sugeridos para as composições do LOCNESS.....	47
QUADRO 6 - Lista de partículas adverbiais .....	51
QUADRO 7 - Linhas de concordâncias para o lema “ <i>depend</i> ” + partícula “ <i>on</i> ” .....	54
QUADRO 8 - Os 25 verbos frasais mais frequentes nos <i>corpora</i> .....	68
QUADRO 9 - Comparação entre “ <i>go on</i> ” e seus verbos únicos equivalentes .....	72
QUADRO 10 - Comparação entre “ <i>sum up</i> ” e termos equivalentes.....	73
QUADRO 11 - Sobreuso e subuso dos verbos frasais no Br-ICLE .....	76

## LISTA DE TABELAS

<b>TABELA 1</b> - Dados estatísticos do <i>corpus</i> de estudo .....	46
<b>TABELA 2</b> - <i>Corpora</i> presentes no LOCNESS e quantidades de textos em cada <i>corpus</i> .....	46
<b>TABELA 3</b> - Dados estatísticos do <i>corpus</i> de referência .....	48
<b>TABELA 4</b> - Dados estatísticos dos <i>corpora</i> de estudo e de referência .....	57
<b>TABELA 5</b> - Frequência de partículas adverbiais nos <i>corpora</i> .....	59
<b>TABELA 6</b> - Teste <i>Log-likelihood</i> .....	67

## LISTA DE ABREVIATURAS E SIGLAS

**BNC** - *British National Corpus*

**Br-ICLE** - *Brazilian Subcorpus of the International Corpus of Learner English*

**CE** - *Corpus de estudo*

**CR** - *Corpus de referência*

**FC** - *Free combination*

**Freq.** - *Frequência*

**L1** - *Primeira língua*

**L2** - *Segunda língua*

**LL** - *Log-likelihood*

**LOCNESS** - *Louvain Corpus of Native English Essays*

**MWV** - *Multi-word verbs*

**Norm.** - *Normalizada*

**OPVD** - *Oxford Phrasal Verbs Dictionary for Learners of English*

**PAV** - *Partículas adverbiais*

**Prep.V** - *Prepositional verb*

**PV** - *Phrasal verbs*

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>16</b>
1.1 Justificativa .....	17
1.2 Objetivos.....	19
1.3 Perguntas de pesquisa .....	19
<b>2 FUNDAMENTAÇÃO TEÓRICA</b> .....	<b>21</b>
2.1 Os <i>multi-word verbs</i> - verbos constituídos por mais de uma palavra.....	21
2.2 Os <i>phrasal verbs</i> (verbos frasais) - Definição.....	23
2.2.1 A <i>semântica dos phrasal verbs</i> .....	24
2.2.2 <i>Critérios de distinção para phrasal verbs</i> .....	26
2.3 Phrasal verbs e o ensino/aprendizagem da língua inglesa .....	30
2.4 Linguística de <i>Corpus</i> .....	33
2.4.1 <i>Definição</i> .....	33
2.4.2 <i>Breve histórico</i> .....	34
2.4.3 <i>Tipos de corpus</i> .....	36
2.4.4 <i>Ferramentas de apoio nas pesquisas baseadas em corpus</i> .....	37
2.4.4.1 <i>Concordanciadores</i> .....	38
2.4.4.2 <i>Etiquetadores</i> .....	38
<b>3 METODOLOGIA</b> .....	<b>40</b>
3.1 Ferramentas utilizadas na pesquisa.....	40
3.1.1 <i>AntConc</i> .....	40
3.2 Os dados utilizados na pesquisa .....	43
3.2.1 <i>Corpus de estudo</i> .....	43
3.2.1.1 <i>Dados do corpus de estudo</i> .....	45
3.2.2 <i>Corpus de referência para a análise quantitativa</i> .....	46
3.2.2.1 <i>Dados do corpus de referência</i> .....	47
3.2.3 <i>Corpus de referência para a análise qualitativa</i> .....	48
3.3 Procedimentos gerais.....	49
3.4 Seleção, extração e validação dos <i>phrasal verbs</i> .....	50
3.5 Procedimentos para a análise qualitativa dos <i>phrasal verbs</i> .....	56
<b>4 ANÁLISE E DISCUSSÃO DOS RESULTADOS: ABORDAGEM QUANTITATIVA</b> .....	<b>57</b>
4.1 Dados estatísticos dos <i>corpora</i> e parâmetro de apresentação dos resultados.....	57
4.2 Frequência e produtividade das partículas adverbiais .....	58
4.3 Frequência e produtividade dos verbos lexicais .....	64
4.4 Os <i>phrasal verbs</i> nos <i>corpora</i> - visão geral .....	66
4.5 Frequência e produtividade dos <i>phrasal verbs</i> na escrita acadêmica .....	68
<b>5 ANÁLISE E DISCUSSÃO DOS RESULTADOS: ABORDAGEM QUALITATIVA</b> .....	<b>78</b>
5.1 Influências da primeira língua na aprendizagem do inglês.....	78
5.2 Criatividade na produção de verbos frasais .....	80
5.3 Naturalidade no idioma .....	82
5.3.1 <i>Desvios colocacionais</i> .....	85
5.4 Estilo e formalidade na escrita acadêmica .....	87

<b>6 CONSIDERAÇÕES FINAIS.....</b>	<b>89</b>
<b>REFERÊNCIAS.....</b>	<b>91</b>
<b>APÊNDICE A - AS 270 PALAVRAS MAIS FREQUENTES BR-ICLE (FREQ. BRUTA).....</b>	<b>96</b>
<b>APÊNDICE B - AS 270 PALAVRAS MAIS FREQUENTES LOCNESS (FREQ. BRUTA).....</b>	<b>98</b>
<b>APÊNDICE C - RELAÇÃO DOS 40 VERBOS LEXICAIS MAIS FREQUENTES NO CORPUS DE ESTUDO.....</b>	<b>100</b>
<b>APÊNDICE D - LISTA DE PARTÍCULAS ADVERBIAIS FORMANDO <i>PHRASAL VERBS</i>, CONFORME A FREQUÊNCIA (OCORRÊNCIAS BRUTAS E OCORRÊNCIAS POR 100 MIL).....</b>	<b>101</b>
<b>APÊNDICE E - PRODUTIVIDADE DAS PARTÍCULAS (COMBINAÇÃO COM VERBOS DIFERENTES).....</b>	<b>102</b>
<b>APÊNDICE F - LISTA DE <i>PHRASAL VERBS</i> ENCONTRADOS NOS <i>CORPORA</i>, CONFORME A FREQUÊNCIA.....</b>	<b>103</b>
<b>ANEXO A - <i>LEARNER PROFILE</i>.....</b>	<b>113</b>
<b>ANEXO B - CIRCUNSTÂNCIAS DE PRODUÇÃO DO LOCNESS .....</b>	<b>115</b>



## 1 INTRODUÇÃO

Esta dissertação apresenta a descrição de uma investigação acerca do uso dos chamados *phrasal verbs* (verbos frasais)<sup>1</sup> em produções escritas de aprendizes da língua inglesa. Na pesquisa, observou-se quais são as semelhanças, diferenças e preferências percebidas na escrita acadêmica em inglês de falantes nativos e não nativos (mais especificamente, dos aprendizes brasileiros) no que concerne à utilização desse item lexical, considerado por muitos (CELCE-MURCIA; LARSEN FREEMAN, 1999; WAIBEL, 2007) como uma das mais desafiadoras formações da língua inglesa, em função, especialmente, de fatores sintáticos e semânticos.

Para tanto, optou-se por buscar na Linguística de *Corpus* o suporte necessário para a realização da investigação. A pesquisa teve como base a utilização de dois *corpora* principais: um brasileiro, para estudo (Br-ICLE), e um compilado na Bélgica, para referência (LOCNESS), ambos contendo coleções de textos produzidos por aprendizes da língua inglesa. Um terceiro *corpus* (BNC) foi utilizado como *corpus* de referência na etapa de análise qualitativa. Para o tratamento e a análise dos dados foi aplicado o *software* AntConc, um programa de computador gratuito composto por um conjunto de recursos que permite a análise de textos e de concordâncias. A ferramenta, que será mais bem explicada ao longo da dissertação, foi utilizada para gerar as listas de frequência e concordância responsáveis por tornar a averiguação possível.

A investigação descrita vai ao encontro de diversos outros estudos baseados em *corpora* que também analisaram as características da escrita de aprendizes nativos e não nativos. Dentre os muitos trabalhos, pode-se citar Biber (1999, 2002, 2010), que estudou variados aspectos do discurso acadêmico, como padrões, variações de registro e fraseologias; e Granger (1997), que comparou as cláusulas de participio no inglês acadêmico em *corpora* de nativos e não nativos. No que concerne à conceituação dos *phrasal verbs*, o estudo buscou fundamentação nas perspectivas gramaticais de Downing e Locke (2006), Downing (2015), Greenbaum e Quirk (1990), Quirk *et al.* (1985) e Carter e McCarthy (2006), que apresentam

---

<sup>1</sup> Os termos *phrasal verbs* e verbos frasais (constantemente alternados ao longo do trabalho) serão utilizados aqui para se referir a um grupo mais reduzido de verbos, a uma abrangência bem específica estabelecida por um grupo de pesquisadores/gramáticos da língua inglesa, a saber: Downing e Locke (2006), Downing (2015), Greenbaum e Quirk (1990), Quirk *et al.* (1985) e Carter e McCarthy (2006). O que, mais amplamente, na literatura se chama *phrasal verbs*, nesta pesquisa, e para os gramáticos mencionados, é denominado *multi-word verbs* (verbos multipalavras).

conceitos similares. A obra de Carter e McCarthy, em especial, também contribui para o estudo porque evidencia exemplos da língua em uso, bem como revela traços das preferências de falantes nativos, o que pode colaborar bastante na resolução das questões indicadas neste trabalho, conforme será mostrado ao longo desta dissertação.

A pesquisa realizada busca contribuir para a descrição desse fenômeno linguístico por meio da identificação de padrões e das preferências dos estudantes, auxiliando não só para uma melhor compreensão da interlíngua, mas também para a conscientização acerca de possíveis adequações e inadequações de uso, bem como para o fomento da reflexão sobre processos de aquisição e uso dos *phrasal verbs*.

Este trabalho está organizado em cinco partes principais (além das considerações finais), que serão desenvolvidas a seguir. A **primeira** delas, de caráter introdutório, apresenta um panorama do trabalho, indicando a justificativa da investigação, os objetivos do estudo e as perguntas de pesquisa. A **segunda** parte, que constitui a fundamentação teórica, trata dos *phrasal verbs*, suas definições, classificações e importância na língua inglesa. O capítulo trata também dos fundamentos da Linguística de *Corpus*, apresentando um breve histórico e informações gerais dos processos nessa área de estudos. Na **terceira** parte, serão mostrados os procedimentos metodológicos usados no desenvolvimento da investigação, bem como os dados dos *corpora* e das ferramentas que foram utilizados nessa análise. Por fim, na **quarta** e na **quinta** partes, serão apresentados e discutidos os resultados encontrados nas análises quantitativa e qualitativa, respectivamente.

## 1.1 Justificativa

De grande importância para a proficiência na língua inglesa, os *phrasal verbs* constituem uma das marcas mais distintivas e criativas do inglês (GARDNER; DAVIS, 2007). Contudo, suas características os tornam difíceis de serem assimilados por estudantes não nativos, especialmente por aqueles cuja primeira língua não tem origem na família germânica (CELCE-MURCIA; LARSEN-FREEMAN, 1999; DARWIN; GRAY, 1999). Faria *et al.* (2007, p. 82) explicam que “o ensino/aprendizado de *phrasal verbs* constitui um desafio para estudantes de inglês graças a peculiares aspectos sintáticos e semânticos”, como a idiomaticidade, as classificações que os diversos autores conferem a essa categoria gramatical, as interferências e transferências linguísticas e a própria terminologia.

Nesse sentido, considerando, principalmente, o papel dos verbos frasais na aquisição da fluência em língua inglesa e, tangencialmente, as questões que permeiam seu

ensino/aprendizagem, este trabalho propõe um estudo descritivo (baseado na Linguística de *Corpus*) acerca do uso dos *phrasal verbs* em produções escritas de estudantes brasileiros. A escolha por pesquisar o fenômeno na língua escrita tem fundamento nos estudos de Biber *et al.* (1999), Fletcher (2005) e McCarthy e O'Dell (2004), que demonstram que, apesar de ser um traço mais frequente na língua oral e informal, essa unidade lexical também tem sido bastante usada na língua escrita formal.

Em estudo anteriormente realizado (COSTA, 2012), foram analisados os *multi-word verbs* (verbos compostos por mais de uma palavra) nas produções acadêmicas de aprendizes da língua inglesa, tendo como foco a identificação dos verbos frasais, preposicionados e frasais preposicionados mais comuns nos textos de estudantes nativos e não nativos, bem como a descrição das diferenças e semelhanças no emprego dessas unidades no gênero textual observado. Os resultados dessa primeira investigação, que também fora realizada com o suporte da Linguística de *Corpus* e que teve sua metodologia baseada na *Constrative Interlanguage Analysis* (GRANGER, 1996), sugeriram que, na comparação com os nativos, os aprendizes brasileiros empregam quantidades bem significativas de *multi-word verbs* em seus textos. Essa afirmação, porém, necessita de estudos mais aprofundados para ser confirmada, já que fora obtida por meio de uma investigação quantitativa limitada, em função do próprio objetivo da pesquisa.

O estudo também revelou a existência de erros na maneira como os verbos formados por mais de uma palavra são usados e que tais erros estão relacionados, principalmente, às produções textuais dos não nativos. Essa revelação também apontou para a necessidade de uma investigação mais específica, que possa gerar dados, por exemplo, qualitativos, para viabilizar a validade ou não dessa hipótese. Esses procedimentos tornariam mais seguras as afirmações acerca da ocorrência ou não de erros no uso dos *multi-word verbs*. Cabe aqui ressaltar que o termo “erro” não deve ser compreendido como julgamento de certo ou errado, de uso próprio ou impróprio, mas sim como maneira de identificar uma relação de sobreuso e subuso, de desacordo ou afastamento daquilo que é tido como usual nos padrões da língua inglesa.

Outra questão relevante levantada em Costa (2012) foi a tendência, por parte dos aprendizes brasileiros, em usar os *multi-word verbs* considerando o significado individual de cada palavra que os compõe, uma percepção que pode indicar, por exemplo, a influência (ou transferência) da língua portuguesa durante a aprendizagem do segundo idioma. Essa afirmação, contudo, também não é segura somente a partir do estudo realizado e suas

limitações. Ela poderia ser respaldada a partir de uma pesquisa mais abrangente, que incluísse avaliações não só quantitativas, mas também qualitativas.

Dessa forma, a investigação realizada por Costa (2012) serviu como estudo piloto para a observação aqui descrita, pois, ao analisar uma categoria maior (a dos *multi-word verbs*, que compreende os *phrasal*, os *prepositional* e os *phrasal-prepositional verbs*), o trabalho colaborou para situar as classes de verbos formados por mais de uma palavra na escrita acadêmica de aprendizes, além de apontar, por meio dos resultados, a representatividade dos *phrasal verbs* nos *corpora* estudados, o que motiva e justifica a escolha desse grupo específico como objeto de estudo da nova pesquisa que foi desenvolvida. Além de focar nesse fenômeno em particular, a nova observação ampliou a análise a partir de exames não só quantitativos, mas também qualitativos do uso de verbos frasais.

## 1.2 Objetivos

Como mencionado no tópico anterior, devido à importância dos *phrasal verbs* para a comunicação em inglês e à sua expressividade na escrita acadêmica, como percebido em Costa (2012), o estudo descrito neste documento teve como objetivo geral descrever e comparar o uso de tais estruturas em textos argumentativos de aprendizes nativos e não nativos (mais especificamente, aprendizes brasileiros) da língua inglesa, o que foi realizado por meio de análises quantitativas e qualitativas do fenômeno, conforme será mostrado ao longo do trabalho.

Como objetivos específicos, buscou-se: a) identificar e descrever padrões recorrentes e regulares nas produções textuais investigadas, que possam indicar escolhas e/ou preferências dos aprendizes nativos e não nativos em relação ao uso dos *phrasal verbs*; b) descrever as diferenças e similaridades entre as produções escritas dos dois grupos examinados, no que concerne ao uso dos verbos frasais; c) discutir as escolhas dos estudantes e suas motivações, considerando o contexto de uso das expressões e o significado atribuído a elas.

## 1.3 Perguntas de pesquisa

Para ajudar a organizar a pesquisa, bem como para dar suporte no alcance aos objetivos, algumas perguntas foram formuladas. Com a análise quantitativa pretendeu-se responder:

- Quais são os *phrasal verbs* mais frequentes em ambos os *corpora*?
- No emprego de *phrasal verbs*, quais são as similaridades e as diferenças percebidas na escrita de nativos e de aprendizes brasileiros?
- Quais são as tendências e/ou preferências de uso dos aprendizes em relação aos nativos? Há evidências que apontem que os aprendizes brasileiros prefiram verbos latinizados e/ou evitem o uso de *phrasal verbs*?

Essas respostas nos permitiram alcançar os dois primeiros objetivos específicos propostos (cf. item 1.3). Trabalhou-se aqui com a hipótese de que aprendizes nativos e não nativos utilizem quantidades significativas de *phrasal verbs* em seus ensaios, entretanto, é possível que aprendizes não nativos tenham preferência por palavras de origem latina.

Já a análise qualitativa nos permitiu responder às seguintes questões:

- No que diz respeito ao emprego dos *phrasal verbs* e seus significados, quais são os principais problemas encontrados nos textos de não nativos em comparação com os de nativos?
- Em relação aos problemas, quais são os possíveis motivos? Pode-se considerar que há ocorrências de novos significados atribuídos a verbos frasais já existentes e/ou formações de novas combinações verbo-partícula?

Por fim, essa etapa alcançou o terceiro objetivo de pesquisa (cf. item 1.3) e a hipótese a ser confirmada nessa análise é a de que os problemas dos aprendizes estão relacionados à interferência e transferência da língua materna, bem como à criatividade e ao uso indevido de colocações.

## 2 FUNDAMENTAÇÃO TEÓRICA

Essa pesquisa busca embasamento teórico em dois campos de estudo específicos que serão abordados a seguir: os verbos frasais (*phrasal verbs*), classificação dada às locuções verbais da língua inglesa formadas por um verbo lexical seguido de uma segunda palavra – um advérbio –, uma formação gramatical abarcada no grupo dos verbos formados por mais de uma palavra (*multi-word verbs*) e a Linguística de *Corpus*, área de pesquisa que se ocupa da coleta de *corpora* textuais para explorar a linguagem por meio de evidências empíricas.

### 2.1 Os *multi-word verbs* - verbos constituídos por mais de uma palavra

Os verbos na língua inglesa podem formar combinações com advérbios (ex. *find out, hang around, break down* etc.), com preposições (ex. *listen to, rely on, call for* etc.) e também com ambos, advérbio e preposição ao mesmo tempo (ex. *get out of, hang up with, settle down with* etc.), formando estruturas que funcionam como verbos únicos, ou seja, estruturas que, apesar de serem compostas por mais de uma palavra, apresentam apenas um significado. Na gramática tradicional do inglês, esse tipo de combinação lexical é denominada *multi-word verb* (verbo multipalavras ou constituído por mais de uma palavra) e compõe um grupo de verbos bastante comuns e produtivos na língua inglesa, especialmente em discursos informais.

Carter e McCarthy (2006) dividem os *multi-word verbs* (MWV) em três categorias principais: verbos frasais (*phrasal verbs*), verbos preposicionados (*prepositional verbs*) e verbos frasais preposicionados (*phrasal-prepositional verbs*). Para os autores, essas estruturas são combinações entre verbos lexicais e partículas - que podem ser advérbio ou preposição – formando uma única unidade de sentido, como em “*I think I’m going to **drop off** soon*” (*drop off* = *fall asleep*) e “*I **came across** some old pictures*” (*came across* = *find*). De acordo com eles, as palavras que formam os *multi-word verbs* são escritas separadamente, nunca como uma única palavra ou com a inserção de hífenes.

Assim como a definição de Carter e McCarthy (2006), a definição de Palmer (1988) também mostra uma classificação que distingue, basicamente, três formas de MWV (*phrasal, prepositional* e *phrasal-prepositional verbs*), que são diferenciadas pela função que a partícula exerce em cada uma delas (ora pode ser identificada como preposição, ora como advérbio, ora como advérbio seguido de preposição, como ocorre com os verbos frasais preposicionados).

Greenbaum e Quirk (1990) também apresentam classificação parecida para os MWV. Para eles, “as duas principais categorias de *multi-word verbs* [*phrasal verbs e prepositional verbs*] consistem de um verbo lexical acrescido de uma partícula, designação neutra para as categorias de advérbio e preposição que são usadas em tais combinações. [...] Além disso, há os *phrasal-prepositional verbs* com verbos com duas partículas, um advérbio seguido de uma preposição [...]” (GREENBAUM; QUIRK, 1990, p. 336, tradução nossa<sup>2</sup>). Os autores, entretanto, mencionam uma quarta categoria, que eles chamam de outras construções de *multi-word verb (other multi-word verb construction)* e que corresponde a combinações idiomáticas como verbo + adjetivo (ex.: *Meg put the cloth straight.*), verbo + verbo não finito (ex.: *Let it go.*) e verbo + duas preposições (ex.: *It developed from a small club into a mass organization in three years.*). Quirk *et al.* explicam ainda que os MWV são formados por “partículas morfologicamente invariáveis que, juntamente com o verbo, funcionam como uma única unidade gramatical”<sup>3</sup> (QUIRK *et al.*, 1985, p. 1150), ou seja, essas estruturas se comportam como um único verbo, seja lexical ou sintaticamente.

Sobre o fato de verbo e partícula terem valor de um único verbo, Downing e Locke afirmam que “a função de muitas das partículas é modificar a natureza da atividade expressa pelo verbo. O resultado é uma extensão de sentido que é frequentemente diferente do (s) sentido (s) dos verbos sozinhos”<sup>4</sup> (DOWNING; LOCKE, 2006, p. 303). Biber *et al.* (1999) confirmam essa característica ao relatarem que essa combinação tem, normalmente, sentido idiomático e que seu significado não pode ser deduzido a partir do significado individual de cada palavra.

Da mesma forma que Quirk *et al.* (1990), Biber, Conrad e Leech (1999) classificam os MWV em quatro categorias: os *phrasal verbs*, os *prepositional verbs*, os *phrasal-prepositional verbs* e outras construções de *multi-word verbs*. Em concordância com os demais autores aqui citados, Biber, Conrad e Leech (1999) também afirmam que a característica dessas unidades é que elas funcionam como um único verbo e, geralmente, possuem significado idiomático. Para os autores, nesse grupo de palavras, a condição da

---

<sup>2</sup> Tradução de minha autoria, bem como em todas as demais citações a partir de textos em outras línguas encontradas ao longo deste trabalho. No original: “The two main categories of multi-word verbs consist of a lexical verb plus a particle, a neutral designation for the overlapping categories of adverb and preposition that are used in such combinations. [...] In addition, there are PHRASAL-PREPOSITIONAL VERBS with verbs with two particles, an adverb followed by a preposition [...]”.

<sup>3</sup> No original: “[...] followed by a morphologically invariable particle, which functions with the verb as a single grammatical unit [...] behaves to some extent either lexically or syntactically as a single verb”.

<sup>4</sup> No original: “The function of many of the particles is to modify the nature of the activity expressed by the verb. The result is an extended meaning which is often different from the meaning(s) of the verb when it functions alone.”

idiomaticidade torna possível, muitas vezes, a substituição do termo formado por mais de uma palavra por palavras únicas, como, por exemplo, a possibilidade de se usar “*undertake*” em vez de “*carry out*”, ou “*discover*” em vez de “*find out*”.

Em suma, pode-se compreender que a conceituação dos *multi-word verbs* está relacionada à existência de pelo menos três tipos principais de combinação entre verbo e partícula, sendo um deles o dos verbos frasais, que é o objeto de estudo deste trabalho e que será mais bem definido no tópico a seguir.

## 2.2 Os *phrasal verbs* (verbos frasais) - Definição

Para Downing (2015), os verbos frasais consistem de um verbo lexical associado a uma partícula adverbial. Eles podem ser intransitivos, ou seja, quando não requerem objeto direto (ex.: *She gets up early.*) ou transitivos, quando requerem objeto direto (ex.: *She switched it off*). Para a autora, no plano sintático, quando o objeto é um substantivo, ele pode tanto preceder a partícula, quanto sucedê-la, como ocorre em “*They called off the wedding*” e “*They called the wedding off*”. Se o objeto for um pronome, por outro lado, a partícula deve ser posicionada sempre depois desse objeto, como em “*They called it off*”. A motivação para esta movimentação da partícula está ligada à atribuição de foco relacionada à distribuição da informação em uma sentença. Os pronomes geralmente não são usados para introduzir novos elementos no discurso, por isso não ocupam, tipicamente, posição final na frase, que é uma posição de foco (HALLIDAY, 2004).

Também Greenbaum e Quirk (1990) e Downing e Locke (2006) classificam os *phrasal verbs* em transitivos e intransitivos, mas ressaltam que algumas combinações, como *give in*, *blow up* e *break down*, por exemplo, podem ser tanto transitivas quanto intransitivas, resultando, em alguns casos, em diferenças no significado da combinação (ex. *They broke down the door to rescue the child.* => transitivo. Sentido = remover um obstáculo; *The car has broken down.* => intransitivo. Sentido = parar de funcionar).

De maneira geral, podemos considerar a classificação dos verbos frasais do inglês conforme o quadro abaixo:



**QUADRO 1** - Classificação dos *phrasal verbs*

	Verbo Lexical	Objeto direto	Partículas		Objeto direto
			Advérbio	Preposição	
<i>Phrasal verbs transitive</i>	<i>Drink</i>	<i>It</i>	<i>up</i>	-	-
	<i>Switched</i>	-	<i>off</i>	-	<i>the light</i>
<i>Phrasal verbs intransitive</i>	<i>Crop</i>	-	<i>up</i>	-	-

Fonte: Tabela adaptada de Greenbaum e Quirk (1990, p. 1181).

### 2.2.1 A semântica dos *phrasal verbs*

Como já visto anteriormente, uma das características que tornam os verbos frasais difíceis de serem aprendidos por estudantes não nativos é a questão da idiomaticidade. Conforme aponta Walková (2013), essas estruturas não são um grupo homogêneo e seu comportamento sintático está ligado às propriedades semânticas que, por serem muito variadas, acabam recebendo classificações diferentes na literatura.

Autores como Quirk *et al.* (1985), Celce-Murcia e Larsen Freeman (1999), Downing e Locke (2006) e Downing (2015) propõem categorizações para o nível de transparência/opacidade dos verbos frasais, ou seja, para a característica que permite a dedução ou não do significado do termo a partir de cada uma das palavras que os formam. Em verbos considerados mais transparentes, como o “*put on*” em “*Now, you can put on your helmet*”, é possível verificar com clareza que, na combinação das palavras, o verbo *put* mantém seu significado de “colocar” e a partícula *on*, da mesma forma, mantém seu sentido direcional, indicando que o capacete foi colocado sobre uma superfície. Downing e Locke (2006) e Downing (2015) chamam a esse tipo de construção de não idiomáticas. Isso não acontece, por outro lado, no verbo frasal “*eat up*” em “*They eat up the fruits*”, já que, nesse caso, o verbo *eat* mantém seu sentido de “comer”, mas a partícula *up* indica, em vez de direção ou movimento de uma posição mais baixa para outra mais alta, que as frutas foram consumidas integralmente, por completo. Ou seja, nesse segundo caso, considerado mais opaco que o primeiro, a partícula *up* funciona como um marcador aspectual, exprimindo valor completivo e sua função é reforçar a intensidade do verbo. Na conceituação de Downing e Locke (2006) e Downing (2015), essas são construções semi-idiomáticas. Em um terceiro exemplo, “*put down*” em “*The army put down the rebellions*” apresenta um significado totalmente idiomático, ou seja, o sentido da combinação dificilmente pode ser compreendido a partir do significado individual das palavras que compõem o verbo frasal. Assim, percebe-se

um sentido metafórico, que torna a expressão ainda mais opaca do que a mostrada no exemplo anterior. Na classificação de Downing e Locke (2006) e Downing (2015), essas são construções totalmente idiomáticas. E é esse nível de transparência/opacidade, esse grau de idiomaticidade, o distanciamento do significado original das palavras que nem sempre é assimilado por aprendizes não nativos, tornando o aprendizado dos verbos frasais complicados e difíceis.

No que tange à dificuldade dos aprendizes brasileiros, especificamente, na compreensão dos *phrasal verbs* podemos ilustrar da seguinte forma: o verbo frasal “*fix up*” em “*They fixed up the toys*” para os nativos da língua inglesa pode ser compreendido claramente com seu significado de totalidade, uma vez que a partícula *up* indica e reforça a completude da ação expressa pelo verbo (ou seja, *up* denota que os brinquedos foram consertados por completo). Já para os aprendizes brasileiros, que não têm referência de nenhuma construção similar aos verbos frasais na sua língua nativa (português), essa construção poderia ser naturalmente formulada por meio da frase “*They fixed the toys*”, sem a combinação com a partícula *up*. Dessa forma, pode-se perceber que, para aprendizes não nativos, o valor aspectual dos verbos frasais muitas vezes não é percebido/compreendido conforme consta nos padrões da língua inglesa e, por isso, a existência ou não da partícula adverbial não faz diferença na construção do significado.

Outro exemplo pode ser visto na frase “*She passed out*”, em que o verbo frasal *pass out* tem sentido de se tornar inconsciente repentinamente, desmaiar. Na escrita de aprendizes brasileiros, contudo, essa sentença poderia facilmente ser representada pelo correspondente *faint* (desmaiar), como em “*She almost fainted*”, demonstrando não só a influência da língua nativa, mas também a tendência pela preferência do uso de verbos de origem latina em detrimento das combinações de verbo lexical + partícula, já que aprender o verbo latinizado é mais fácil do que aprender *phrasal verbs* (SIDE, 1990). O problema disso, é que nem sempre essas escolhas propõem substituição com a mesma conotação dos verbos frasais, conforme aponta Cornell ao afirmar que essas estruturas “são muitas vezes mais específicas em significado do que suas equivalentes [verbos de uma só palavra] e geralmente carregam conotações para as quais seus potenciais usuários devem estar atentos”<sup>5</sup> (CORNELL, 1985, p. 275). Sendo assim, essa diferença pode contribuir para afastar os aprendizes da aquisição de uma *performance* mais natural na língua inglesa.

---

<sup>5</sup> No original: “*phrasal verbs are often more specific in meaning than their ‘equivalents’ and often carry connotations which their potential users must be aware of*”.

### 2.2.2 Critérios de distinção para *phrasal verbs*

Apesar de a maioria das atuais gramáticas que discorrem sobre os verbos frasais do inglês consentirem que essas estruturas são combinações entre um verbo e uma partícula adverbial (DOWNING, 2015; DOWNING; LOCKE, 2006; CARTER; MCCARTHY, 2006; BIBER *et. al.*, 1999; BIBER; CONRAD; LEECH, 2002), elas também reconhecem a dificuldade de se realizar a distinção clara destas unidades em relação a outros tipos de MWV, especialmente com os *prepositional verbs*. Isso ocorre, entre outros motivos, porque as partículas podem pertencer a mais de uma categoria (advérbio ou preposição) e, conseqüentemente, podem apresentar valores diferentes em determinados contextos. Um exemplo disso é a palavra *up* que, de acordo com o *English Oxford Living Dictionary* (edição *on-line*), pode ser empregada com função de advérbio, preposição, adjetivo, substantivo ou mesmo verbo.

Outro exemplo, retirado de Biber *et al.* (1999, p. 407), mostra um caso em que a combinação “*come back*”, com sentido de “*recover; return an activity*”, pode ser interpretada tanto como combinação livre (quando *come* e *back* contribuem individualmente para o significado da expressão, como em “*When Jim went to the police station, officers told him to come back another day.*”), quanto como verbo frasal (quando o significado combinado das duas palavras é diferente daquele produzido por cada uma das partes isoladas, como em “*Everton came back from a goal down to beat Blackburn 2-1.*”). Assim, em razão das diferentes funções sintáticas das partículas nos mais variados contextos, saber reconhecer se um verbo é preposicionado ou se é um *phrasal verb*, por exemplo, nem sempre é tarefa fácil. Por isso, alguns autores propõem testes e/ou critérios que, apesar de não serem definitivos, podem ajudar na identificação das estruturas, como veremos a seguir:

- **Downing (2015)**

A distinção entre *phrasal verbs* e *prepositional verbs* para Downing (2015) está relacionada à posição, acentuação e inserção do advérbio na oração. Um dos critérios propostos é o de que o pronome é posicionado depois da preposição, mas precede a partícula do *phrasal verb* (PV), conforme mostram os exemplos abaixo:

*He broke with her.* (Prep. V)

*He broke it up.* (PV)

*\*He broke her with.*<sup>6</sup>

(DOWNING, 2015, p. 57)

*\*He broke up it.*

Outro critério de distinção é que a partícula dos verbos frasais é acentuada, especialmente quando esta se encontra no final da oração. Já a preposição, normalmente, não é acentuada. Em sentenças com verbos preposicionados, a acentuação recai sobre o verbo:

*He broke it UP.* (PV)

*Which party did he break UP?* (PV)

(DOWNING, 2015, p. 57)

*He has BROken with her.* (Prep. V)

*Which girl has he BROken with?* (Prep. V)

Um terceiro critério retoma a ideia de posição do advérbio, que nos *prepositional verbs* pode ser inserido entre um verbo e sua preposição. O mesmo não ocorre nos verbos frasais que, geralmente, não admitem advérbios entre o verbo lexical e sua partícula:

*\*He broke completely up the party.* (PV)

*He broke completely with his girl-friend.* (Prep. V)

(DOWNING, 2015, p. 58)

No que diz respeito a verbos frasais idiomáticos, Downing (2015) postula que a partícula é geralmente analisada como parte do verbo, ou seja, o verbo não existe sozinho (ex. o sentido idiomático é atribuído a “*peter out*” [parar ou desaparecer gradualmente], mas nunca a “*peter*” [falhar] isoladamente). Além disso, nas combinações livres com partícula adverbial de função direcional, o advérbio é analisado como complemento, de forma que pode ser posicionado antes do verbo para fins retóricos. Quando não tem função direcional, entretanto, a partícula se torna inseparável ao verbo e não pode ser posicionada antes dele (ex.: *The car broke down* / *\*Down broke the car*).

- **Greenbaum e Quirk (1999)**

O primeiro critério de distinção entre *phrasal* e *prepositional verbs* proposto por Greenbaum e Quirk (1999) está relacionado à mobilidade da partícula que, no verbo preposicionado deve preceder o objeto da preposição, mas no verbo frasal pode tanto preceder

---

<sup>6</sup> O símbolo (\*) indica que a sentença é agramatical.

quanto suceder o objeto direto. Para exemplificar a diferença, os autores usam o verbo preposicionado *call on* (visitar) e o verbo frasal *call up* (convocar):

*She called on her friends. (Prep. V)*                      *She called up her friends. (PV)*  
*\*She called her friends on.*                                      *She called her friends up.*  
 (GREENBAUM; QUIRK, 1999, p. 340)

Ainda a respeito da mobilidade, quando o objeto é um pronome pessoal, ele segue a partícula do verbo preposicionado, mas precede a partícula do verbo frasal:

*She called on them. (Prep. V)*                      *She called them up. (PV)*  
*\*She called them on.*                                      *\*She called up them.*  
 (GREENBAUM; QUIRK, 1999, p. 340)

Outro critério é o da inserção de um advérbio. Quando funciona como adjunto, o advérbio pode ser posicionado entre o verbo e a partícula de um verbo preposicionado. Nos verbos frasais, contudo, esse comportamento não ocorre:

*She called angrily on her friends. (Prep. V)*  
*\*She called angrily up her friends. (PV)*  
 (GREENBAUM; QUIRK, 1999, p. 340)

Diferentemente do que pode ocorrer com verbos preposicionados, a partícula de um verbo frasal não pode preceder um pronome relativo ou uma interrogativa formada com *wh-*:

*the friends on whom she called. (Prep. V)*                      *\*the friends up whom she called. (PV)*  
*On which friends did she call? (Prep. V)*                      *\*Up which friends did she call? (PV)*  
 (GREENBAUM; QUIRK, 1999, p. 340-341)

O último critério de Greenbaum e Quirk (1999) diz respeito à acentuação, e concorda com Downing (2015) quando estabelece que a partícula do verbo frasal é acentuada, em oposição aos verbos preposicionados, cuja partícula não é acentuada. Além disso, o núcleo tonal da expressão nos verbos preposicionados recai sobre o verbo:

*Which friends did she **CALL on**? (Prep. V)*

*Which friends did she **call UP**? (PV)*

(GREENBAUM; QUIRK, 1999, p. 341)

- **Biber, Conrad, Leech (2002)**

A distinção entre *phrasal verbs* e outros tipos de verbos formados por mais de uma palavra pode ser baseada em dois critérios principais: idiomaticidade e movimento da partícula. Antes, contudo, os autores ressaltam que é importante saber que a natureza dos MWV é determinada pela existência ou não de uma frase nominal. Assim, quando a expressão não é seguida por frase nominal (ex.: *Shut up!*), a interpretação é feita como verbo frasal intransitivo ou combinação livre de verbo + advérbio. Por outro lado, se a expressão é seguida por frase nominal (ex.: *Find out the meaning!*), a interpretação passa a ser feita entre três tipos de construção: verbo frasal transitivo, verbo preposicionado transitivo ou combinação livre de verbo + frase adverbial preposicionada.

O critério da idiomaticidade é mais adequado para situações em que o MWV não é seguido por frase nominal e a diferença é que os verbos frasais intransitivos são geralmente idiomáticos, enquanto as combinações livres mantêm os significados individuais das palavras que as compõem:

***Come on!** Tell us then. (intransitive PV)*

*He was afraid to **look back**. (intransitive FC)*

(BIBER; CONRAD; LEECH, 2002, p. 125)

O critério de movimento da partícula, por sua vez, é mais adequado para situações em que a combinação MWV é seguida por frase nominal e a diferença é que os verbos frasais transitivos permitem o movimento da partícula adverbial para antes ou depois do objeto da frase nominal, o que não é possível em verbos preposicionados, já que, nesses casos, a partícula deve vir sempre antes do objeto (frase nominal):

*I've got to **get this one back** to her mom. (transitive PV)*

*Well those kids are **waiting for the bus**. (Prep. V)*

(BIBER; CONRAD; LEECH, 2002, p. 125)

O QUADRO 2 exibe um resumo das teorias analisadas neste tópico no que concerne aos critérios de distinção dos verbos frasais em relação a outros tipos de MWV.

**QUADRO 2 - Testes e/ou critérios de distinção dos verbos frasais**

<b>Autores</b>	<b>Crítérios de distinção dos PVs</b>
<b>Downing (2015)</b>	Posição/mobilidade da partícula adverbial.
	Acentuação da partícula adverbial.
	Inserção de um advérbio na oração.
	Em PVs idiomáticos a partícula é analisada como parte do verbo.
<b>Greenbaum e Quirk (1999)</b>	Mobilidade da partícula adverbial.
	Inserção de um advérbio na oração.
	Posição da partícula em relação a pronomes relativos e <i>wh-</i> .
	Acentuação da partícula adverbial.
<b>Biber, Conrad e Leech (2002)</b>	Sentido idiomático (mais forte em PVs intransitivos).
	Mobilidade/posição da partícula adverbial.

Fonte: Elaborado pela autora.

### 2.3 *Phrasal verbs* e o ensino/aprendizagem da língua inglesa

Mesmo sendo consideradas marcas essenciais na linguagem cotidiana dos falantes da língua inglesa e, portanto, de grande importância para uma comunicação natural no idioma<sup>7</sup>, os MWV são combinações gramaticais de grande complexidade no que diz respeito à aprendizagem por estudantes de inglês como segunda língua (ESL/EFL).

Celce-Murcia e Larsen Freeman (1999) colocam que poucas línguas não germânicas apresentam formações como os *phrasal verbs*<sup>8</sup>, por isso, a maioria dos alunos acha esses verbos estranhos e também difíceis, o que faz com que muitos deles optem por evitar o uso dessas construções. Dentre os obstáculos que surgem na aprendizagem dessas combinações

<sup>7</sup> Sobre a potencialidade dos *multi-word-verbs* para a comunicação em língua inglesa de aprendizes não nativos, Bywater (1969 *apud* CORNELL, 1985, p. 270, tradução nossa) afirma: “O fato é que o que distingue a escrita e, acima de tudo, a fala de um bom estudante estrangeiro das de um falante nativo do inglês é que o que o falante nativo escreve ou diz está cheio dessas expressões [*multi-word verbs*], ao passo que a maioria dos aprendizes não nativos tem medo dessas estruturas, cuidadosamente as evitam e, conseqüentemente, soam não naturais quando as usam. Estudantes não nativos que gostam de ser lisonjeados por seu inglês podem alcançar isso usando corretamente um grande número desses verbos compostos”.

<sup>8</sup> Celce-Murcia e Larsen Freeman (1999) utilizam o termo *phrasal verbs* para falar dos verbos formados por mais de uma palavra de maneira geral.

verbais estão fatores sintáticos e semânticos, como a idiomaticidade, ou seja, o fato de o significado da locução não ser, muitas vezes, compreendido a partir do significado individual de cada uma das palavras que o compõem.

Acerca dessa característica, Dempsey, McCarthy e McNamara (2007) relatam que é comum aprendizes driblarem a idiomaticidade usando verbos “latinizados”, ou seja, palavras que, apesar de análogas, não têm a mesma conotação dos verbos frasais ou, quando têm, não soam tão naturais aos ouvidos de um falante nativo. Entretanto, uma vez que os verbos frasais são “ubíquos no inglês e ninguém pode falar ou entender o idioma, pelo menos no contexto informal, sem conhecer esse tipo de registro”<sup>9</sup> (CELSE-MURCIA; LARSEN FREEMAN, 1999, p. 425), isso, provavelmente, se torna um problema para o aprendiz.

Para Waibel (2007), por outro lado, os verbos latinizados podem parecer mais eruditos para os aprendizes não nativos e se tornam mais utilizados por soarem mais formais. Na visão da autora, os estudantes tendem a crer que a adoção de palavras mais sofisticadas faz deles, pelo menos em aparência, mais proficientes na comunicação em inglês, de forma que eles passam a ter *performances* semelhantes às de nativos da língua. Contudo, a consequência disso é que os aprendizes nem sempre são conscientes de que “o uso de verbos latinizados é inadequado em determinadas situações”<sup>10</sup> (WAIBEL, 2007, p. 37).

Em suma, a importância dos verbos frasais para a aquisição de fluência na língua inglesa e os elementos que permeiam seu ensino/aprendizagem são fatores que motivam pesquisadores em todo o mundo a utilizarem essas combinações como objeto de estudo, seja no domínio pedagógico ou no da descrição linguística. A maior parte dos trabalhos realizados nesse campo é internacional, como os de Dagut e Laufer (1985), que conduziram um estudo sobre o uso de *phrasal verbs* por aprendizes israelenses. Na pesquisa, trabalhou-se especificamente a questão da esquiva em relação aos verbos multipalavras e seus resultados revelaram, entre outras coisas, que os estudantes hebreus evitam o uso desse tipo de estrutura, dando preferência para formações de uma só palavra. Uma motivação é que, como não existem construções gramaticais equivalentes aos verbos frasais em hebraico, os estudantes optam por usar palavras que já conhecem.

Outro trabalho que também aborda o aspecto da esquiva é o de Liao e Fukuya (2004), que buscou explicar, nesse caso com base em aprendizes chineses, porque os estudantes tendem a evitar o uso de *phrasal verbs*. A análise focou nas diferenças estruturais entre

<sup>9</sup> No original: “Yet they are ubiquitous in English; no one can speak or understand English, at least informal register, without a knowledge of phrasal verbs”.

<sup>10</sup> No original: “[...] not being aware that in certain situations the use of a Latinate verb is inappropriate.”



primeira língua (L1) e segunda língua (L2) e apresentou, como conclusão, que evitar ou não evitar verbos frasais pode ser mais uma manifestação do desenvolvimento da interlíngua dos aprendizes, do que diferenças ou similaridades entre L1-L2.

O estudo de Waibel (2007) mostrou análises quantitativas e qualitativas sobre o uso de verbos frasais nos textos de língua inglesa escritos por estudantes alemães e italianos e descreveu os problemas que os aprendizes avançados têm em relação ao emprego dos *phrasal verbs*, como a influência da língua materna, por exemplo. Seus resultados contribuíram para a compreensão de aspectos gerais da linguagem de aprendizes avançados.

Dos trabalhos mais recentes, e também analisando *corpora*, Kamarudin (2013) examinou o nível de compreensão e de uso dos *phrasal verbs* por aprendizes malásios do inglês. Seu estudo indicou, entre outras coisas, que além do nível das condições de produção do material analisado, a natureza dos *phrasal verbs* e os fatores interlinguísticos têm papel importante na aprendizagem dos alunos da Malásia e que a dificuldade dos não nativos é agravada pela insuficiência e inadequação das informações difundidas nos livros didáticos e dicionários. Ryoo (2013) analisou os *phrasal verbs* na escrita de estudantes coreanos, identificando e comparando os verbos frasais mais frequentes nas composições de nativos e não nativos, e as conclusões indicaram baixa competência para formulação de *phrasal verbs* em aprendizes coreanos da língua inglesa.

No Brasil, podemos citar como exemplo o estudo de Fernandes (2012), que investigou a aquisição e o uso de *chunks* formados com o verbo *get* em construções de movimento por aprendizes brasileiros de inglês como segunda língua. Sua pesquisa foi baseada em *corpus* e demonstrou que, no processo de aprendizagem, os estudantes usaram a transferência de línguas como estratégia de aprendizado de verbos frasais e que isso gera erros na produção final.

A pesquisa de Fadanelli (2012) teve como objetivo determinar se os verbos frasais mais usados por aprendizes brasileiros eram também os mais usados por aprendizes nativos e se os significados aplicados por cada grupo coincidem entre si. O trabalho foi realizado com *corpora* de produções textuais e mostrou, entre outras coisas, que grande parte das estruturas usadas pelos aprendizes brasileiros não pertencia ao grupo de mais frequentes entre os falantes nativos de inglês.

Nos próximos tópicos, serão apresentados os principais conceitos relacionados à Linguística de *Corpus*, metodologia que serviu de base para a pesquisa aqui descrita.

## 2.4 Linguística de *Corpus*

### 2.4.1 Definição

A Linguística de *Corpus* é um dos campos de pesquisa que tem demonstrado grande relevância para os estudos do léxico e da linguagem. Na concepção de Sardinha (2004), ela:

[...] ocupa-se da coleta e exploração de *corpora*, ou conjunto de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador. (BERBER SARDINHA, 2004, p. 3).

Trata-se de uma metodologia cujo quadro conceitual é formado por uma abordagem baseada em dados contextualizados da língua e que entende a linguagem como um sistema probabilístico (BERBER SARDINHA, 2000), característica que evidencia a importância dos estudos baseados em *corpora*. Os dados de um *corpus* são oriundos de extratos da língua oral e escrita e o agrupamento de tais dados tem por finalidade propiciar a análise de fenômenos específicos da língua, permitir a observação de “padrões sintáticos de uso para as expressões linguísticas pré-definidas”<sup>11</sup> (BIBER, 2009, p. 276).

A terminologia “*corpus*” pode ser entendida como um “conjunto tão variado quanto possível de enunciados efetivamente emitidos por usuários da referida língua em determinada época” (DUCROT; TODOROV, 2001, p. 42). McEnery e Wilson (2001) afirmam que,

em princípio, qualquer coleção de mais de um texto pode ser chamada de *corpus*: o termo ‘*corpus*’ é o correspondente latino para ‘corpo, portanto um *corpus* pode ser definido como qualquer corpo de texto. [...] Mas o termo ‘*corpus*’ quando usado no contexto da linguística moderna tende mais frequentemente a ter conotações mais específicas do que prevê essa simples definição. Elas podem ser consideradas sob quatro categorias principais: amostragem e representatividade, tamanho finito, formato eletrônico e referência padrão.<sup>12</sup> (MCENERY E WILSON, 2001, p. 29).

A característica de formato eletrônico citada por McEnery e Wilson também aparece na concepção de Baker (1995):

<sup>11</sup> No original: “[...] the systematic patterns of use for those pre-defined linguistic features.”

<sup>12</sup> No original: “In principle, any collection of more than one text can be called a corpus: the term ‘corpus’ is simply the Latin for ‘body’, hence a corpus may be defined as any body of text. [...] But the term ‘corpus’ when used in the context of modern linguistics tends more frequently to have more specific connotations than this simple definition provides for. These may be considered under four main headings: sampling and representativeness, finite size, machine-readable form, a standard reference.”

Hoje, o *corpus* significa, principalmente, uma coletânea de textos agrupados em formato eletrônico e capazes de serem analisados automática ou semi-automáticamente de diversas formas; um *corpus* não é mais restrito à escrita, mas também inclui textos falados assim como textos escritos, e ele pode conter um grande número de textos de várias fontes, de vários escritores e sobre vários assuntos. O importante é que sejam compilados com um propósito e de acordo com critérios específicos de maneira a assegurar a representatividade de determinada área ou a amostra da língua a qual se pretende explicar.<sup>13</sup> (BAKER, 1995, p. 225).

Dessa forma, em concordância com Sinclair (2004), podemos concluir que um *corpus* é uma coleção de trechos de texto, em formato eletrônico, selecionados conforme critérios externos para representar, tanto quanto possível, uma língua ou variedade da língua como fonte de dados para a pesquisa linguística.

#### 2.4.2 Breve histórico

Embora tenha se tornado uma abordagem mais precisa após a evolução da tecnologia, a Linguística de *Corpus* não é considerada exatamente uma forma de estudo recente, como evidencia Berber Sardinha (2004) ao relatar que já na Antiguidade e Idade Média eram produzidos *corpora* com as palavras da Bíblia. Também McCarthy e O'Keeffe (2010) apontam que as primeiras concordâncias no campo da Linguística de *Corpus* foram realizadas na esfera dos estudos bíblicos e, conforme mencionam,

a etimologia de *concordantia* é a combinação do *cum* latino, que significa “com” e *cor*, que significa “coração”, que se amarra à base ideológica original desse cuidadoso esforço, mais especificamente para enfatizar a reivindicação de que a Bíblia era uma mensagem divina harmoniosa e não uma série de textos de múltiplas fontes. Anthony de Pádua (1195-1231) está associado à primeira concordância conhecida (anônima) da Bíblia, o *Concordantiae Morales*, baseado no *Vulgata* (a versão latina do século XV da Bíblia). Um trabalho bem documentado na mesma época foi do cardeal Hugo de StCaro (também referido como St Cher), que, em 1230, com a ajuda de uma equipe de 500 monges dominicanos do convento St James, em Paris, compilaram um “index de palavras” do *Vulgata* (Bromiley, 1997: 757; veja também esse volume de Tribble.)<sup>14</sup>. (MCCARTHY; O'KEEFFE, 2010, p. 3).

<sup>13</sup> No original: “*Corpus now means primarily a collection of texts held in machine-readable form and capable of being analyzed automatically in a variety of ways; a corpus is no longer restricted to ‘writings’ but includes spoken as well as written text, and a corpus may include a large number of texts from a variety of sources, by many writers and speakers and on a multitude of topics. What is important is that it is put together for a particular purpose and according to explicit design criteria in order to ensure that it is representative of the given area or sample of language for which it aims to account*”.

<sup>14</sup> No original: “*The etymology of concordantia is the Latin cum, meaning ‘with’, and cor meaning ‘heart’, which ties in with the original ideological underpinning of this painstaking endeavour, namely to underscore the claim that the Bible was a harmonious divine message rather than a series of texts from a multitude of sources. Anthony of Padua (1195–1231) is associated with the first known (anonymous) concordance of the bible, the Concordantiae Morales, based on the Vulgate (the fifth-century Latin version of the Bible). A well-documented work around the same time was by Cardinal Hugo of St Caro (also referred to as St Cher), who in 1230, aided*

Ao longo do século XX, muitos trabalhos foram feitos baseados em *corpora*, contudo, os procedimentos de coleta e análise de *corpus* demandavam equipes grandes de analistas, já que as tarefas eram realizadas manualmente. Combinado com o enorme número de dados de alguns trabalhos, isso favorecia o aumento das possibilidades de falhas procedimentais no estabelecimento de padrões de classificação. De igual maneira, os resultados eram comprometidos, tornando-se pouco confiáveis. Berber Sardinha (2000) cita que, em 1897, Käding utilizou um *corpus* de 11 milhões de palavras para analisar a distribuição ortográfica da língua alemã e que foi necessária a colaboração de cinco mil analistas para concretizar a análise. No ano de 1921, Edward Thorndike realizou um trabalho sobre a frequência de palavras da língua inglesa, o que foi considerado um estudo revolucionário para o ensino do idioma. A investigação foi baseada em um *corpus* de 4,5 milhões de palavras, realizada manualmente e revisada, anos depois, com a aplicação de um *corpus* ainda maior, contendo 18 milhões de palavras.

A tecnologia, representada especialmente pelo uso do computador, e a informatização dos dados na Linguística de *Corpus* inovaram e mudaram bastante a maneira como se perquire a linguagem, possibilitando a aferição de dados muito mais volumosos, em menos tempo e com exatidão e riqueza de detalhes antes inalcançáveis. Toda essa revolução deve-se, em grande parte, ao desenvolvimento de novos programas de computador e da aplicação de ferramentas, como os concordanciadores e etiquetadores, que se tornaram essenciais na atividade do linguista que trabalha com *corpora*, uma vez que dinamizam os processos de extração e análise das evidências baseadas em experiências no campo da exploração da linguagem.

O primeiro *corpus* linguístico eletrônico foi lançado em 1964, conforme aponta Berber Sardinha (2004). Denominado *Brown University Standard Corpus of Present-day American English*, ele continha um milhão de palavras (de inglês americano escrito) e foi considerado um marco na história da Linguística de *Corpus*, graças às contribuições que gerou para os estudos da linguagem. Apesar de ser considerado pequeno, ainda hoje é muito utilizado por analistas da língua.

No Brasil, a Linguística de *Corpus* é cada vez mais aliada da tecnologia e vem se tornando um importante recurso não só para as pesquisas relacionadas ao processamento da

---

by a 500-strong team of Dominican monks at St James' convent in Paris, put together 'a word index' of the Vulgate (Bromiley 1997: 757; see also Tribble this volume)."

linguagem natural, da lexicografia e da linguística computacional, mas também para as necessidades corporativas (BERBER SARDINHA, 2000).

### 2.4.3 Tipos de corpus

Com a Linguística de *Corpus* se tornando mais usual como campo de estudo, surgem a todo o momento novas compilações que servem a propósitos distintos de análise de dados. Esses *corpora* são classificados conforme critérios baseados na função e no conteúdo de cada coleção. O QUADRO 3, adaptado de Berber Sardinha (2004) e Teixeira (2008), resume as nomenclaturas empregadas para os principais tipos de *corpora* citados na literatura:

QUADRO 3 - Tipologia empregada na Linguística de *Corpus*

CRITÉRIO	TIPO DE <i>CORPUS</i>	DESCRIÇÃO
<b>Língua</b>	Monolíngue	Composto por uma língua ou variedade linguística.
	Bilíngue/Multilíngue	Composto por mais de uma língua.
	Multivarietais	Composto por variantes regionais ou nacionais de uma mesma língua.
<b>Modo</b>	Falado	Composto de porções de fala transcritas.
	Escrito	Composto de textos escritos, impressos ou não.
	Misto	Composto de textos escritos e falas transcritas.
<b>Tempo</b>	Sincrônico	Compreende um período de tempo.
	Diacrônico	Compreende vários períodos de tempo.
	Contemporâneo	Representa o período de tempo corrente.
	Histórico	Representa um período de tempo passado.
	Fechado	Não há possibilidade ou intenção de se acrescentar mais textos após a conclusão da compilação.
	Aberto	Há possibilidade ou intenção de se acrescentar mais textos após a conclusão da compilação.
<b>Seleção</b>	De amostragem	Composto por porções de textos ou de variedades textuais, planejado para ser uma amostra finita da linguagem como um todo.
	Monitor	A composição é reciclada para refletir o estado atual de uma língua. Opõe-se a <i>corpora</i> de amostragem.
	Dinâmico ou orgânico	O crescimento e diminuição são permitidos, qualifica o <i>corpus</i> monitor.
	Estático	Oposto de dinâmico, caracteriza o <i>corpus</i> de amostragem.
	Equilibrado	Os componentes (gêneros, textos, etc.) são distribuídos em quantidades semelhantes (por exemplo, mesmo número de textos por gênero).

Continua

		Conclusão
CRITÉRIO	TIPO DE <i>CORPUS</i>	DESCRIÇÃO
<b>Conteúdo</b>	Representativo da língua geral	Os textos são de tipos e gêneros textuais diversos.
	Representativo de parte da língua	Os textos são de tipos específicos (regionais, especializados etc.).
<b>Autoria</b>	Falantes nativos	Os autores são falantes nativos.
	Falantes não-nativos	Os autores dos textos não são falantes nativos (aprendizes, tradutores etc.).
	Individual	Textos elaborados por um autor.
	Coletiva	Textos elaborados por mais autores (ex. artigos científicos).
	Institucional	Textos de empresas, órgãos governamentais etc.
<b>Tamanho</b>	Pequeno	Até 80 mil palavras.
	Pequeno-médio	De 80 a 250 mil palavras.
	Médio	De 250 mil a um milhão de palavras.
	Médio-grande	De um milhão a 10 milhões de palavras.
	Grande	Acima de 10 milhões de palavras.
<b>Disposição Interna</b>	Comparáveis	Os textos são comparáveis (características semelhantes).
	Paralelos	Textos originais acompanhados das respectivas traduções. Estes também podem ser “Alinhados”, ou seja, quando a tradução e o texto original podem ser vistos lado a lado ou um abaixo do outro.
<b>Finalidade</b>	De estudo	O <i>corpus</i> que se pretende descrever.
	De referência	Usado para fins de contraste com o <i>corpus</i> de estudo.
	De treinamento ou teste	Construído para permitir o desenvolvimento de aplicações e ferramentas de análise.

Fonte: Elaborado pela autora com dados extraídos de Berber Sardinha, 2004, p. 20-21, e Teixeira, 2008, p. 161-163.

Neste estudo, foram utilizados *corpora* de dois tipos: um *corpus* de estudo e dois de referência. Todos serão apresentados no terceiro capítulo, que é a parte dedicada à metodologia de pesquisa.

#### 2.4.4 Ferramentas de apoio nas pesquisas baseadas em *corpus*

Como já visto anteriormente, a tecnologia ganhou papel fundamental nos trabalhos dentro do campo da Linguística de *Corpus*. Por apresentar dados cujo tratamento é realizado por meio eletrônico, o *corpus* necessita de ferramentas informatizadas para agilizar e conferir mais confiabilidade aos processos de coleta, armazenagem e análise. Na área da tecnologia

computacional, há disponíveis várias ferramentas eletrônicas, desenvolvidas com a finalidade de se facilitar o trabalho com *corpora*. Além disso, programas de computador especializados, como os concordanciadores, proporcionam o aprimoramento das pesquisas, estabelecendo processos mais automatizados e rápidos, bem como a obtenção de resultados mais fidedignos. A seguir, são apresentadas duas das ferramentas mais conhecidas pelos pesquisadores.

#### 2.4.4.1 Concordanciadores

São chamados concordanciadores os programas que extraem as ocorrências de uma palavra (ou nódulo) de busca em um *corpus* juntamente com o seu contexto, apresentando-as na forma de uma concordância (TAGNIN, 2010). Esses *software* criam listas que agrupam todas as ocorrências das palavras ou expressões previamente indicadas pelo usuário. Tais listas geradas exibem o nódulo em posição central, em uma formatação que, na maioria dos concordanciadores, recebe o nome de KWIC (*Key Word In Context*). Os concordanciadores permitem ao usuário buscar por palavras em um *corpus*, fornecendo exaustivas listas da palavra em um contexto (SANTOS, 2008) e, por esse motivo, são essenciais para investigações baseadas em *corpus*, principalmente quando o propósito é observar a colocação e a padronização do léxico.

#### 2.4.4.2 Etiquetadores

Outra ferramenta relevante para a investigação em *corpus* são os etiquetadores. Trata-se de um programa que insere no *corpus*, automaticamente, códigos referentes à classificação gramatical das palavras (nos planos sintático, semântico, morfológico ou discursivo), estabelecendo para o usuário a indicação funcional da palavra conforme os padrões da gramática de uma língua. Berber Sardinha (2004) informa que o processo de etiquetagem pode ocorrer de forma automática ou semiautomática (interativa) e que pode pertencer a vários tipos, como a etiquetagem semântica (que indica o sentido das palavras), sintática (também conhecida como *parsing* e que faz indicações conforme a estrutura sintática das frases), morfossintática (considerada a mais comum e também conhecida como *pos tagging*) e discursiva (que trabalha no plano do discurso, marcando características como tópicos discursivos, por exemplo).

Embora seja muito útil no processo de análise linguística baseado em *corpus*, em alguns tipos de investigação, como o da presente pesquisa, esse tipo de ferramenta

demonstrou necessitar de auxílio manual para produzir os resultados esperados. Afinal, algumas palavras desempenham funções sintáticas e semânticas diferentes dependendo do contexto em que se encontram, e a anotação automática pode não ser fiel na detecção do sentido das palavras no contexto. Resultados de experiências pessoais utilizando etiquetagem morfossintática para a análise de *multiword verbs* mostraram que a anotação automática não foi confiável na extração de verbos frasais do inglês, já que a etiquetagem não foi capaz de diferenciar as partículas adverbiais e preposicionadas com eficácia. Verbos que faziam combinação com palavras como “*down*”, por exemplo, que é classificada pelo *English Oxford Living Dictionary* (edição *on-line*) com as funções de advérbio, preposição, verbo, adjetivo e substantivo, tiveram que ser analisadas de forma semiautomática, pois, em um número considerável de casos, elas foram marcadas como advérbio quando exerciam função de preposição, e vice-versa. Por isso, o uso unicamente de anotação automática deve ser avaliado em cada caso.



### 3 METODOLOGIA

Neste capítulo, apresentaremos as decisões metodológicas que conduziram a realização deste trabalho. Para tal, em um primeiro momento apresentaremos as ferramentas que serviram de suporte para a investigação, permitindo o tratamento e o exame dos dados linguísticos. Em seguida, descreveremos os *corpora* de estudo e de referência utilizados na pesquisa (Br-ICLE, LOCNESS e BNC). Por fim, discutiremos os procedimentos adotados na etapa de análise.

#### 3.1 Ferramentas utilizadas na pesquisa

Ao longo deste trabalho, usamos como instrumentos de suporte os seguintes programas de computador: **AntConc** (*software* concordanciador), que foi usado para analisar os dados linguísticos dos *corpora* e que será apresentado no item 3.1.1; **Microsoft Excel** (planilhas eletrônicas), utilizado para a análise e organização dos dados estatísticos obtidos, e **Bloco de Notas** (editor de textos), que serviu para transformar os dados em textos e para anular formatações prévias (limpeza) na transferência de dados do AntConc para o Excel.

##### 3.1.1 AntConc

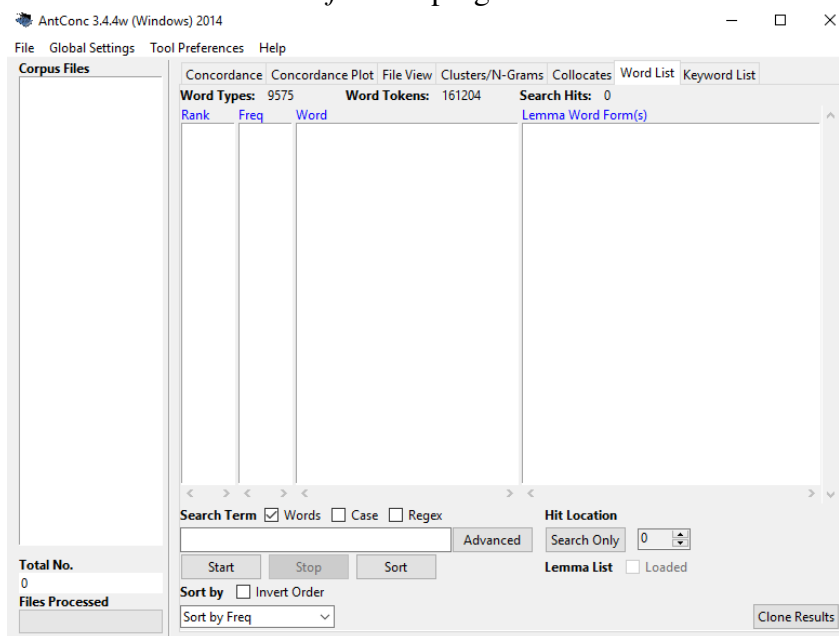
O AntConc é um *software* concordanciador desenvolvido pelo pesquisador britânico Laurence Anthony, da Universidade de Waseda, no Japão. Trata-se de uma ferramenta que permite analisar *corpora* a partir de funções tais como linhas de concordâncias, lista de palavras-chaves, listas de frequências, listas de palavras individuais e múltiplas, colocações, dentre outras. O programa é *freeware*, ou seja, é um *software* gratuito, que está disponível para *download*<sup>15</sup> na internet. Atualmente, o AntConc é oferecido para as plataformas Windows, Linux e Macintosh. Apesar de estar na versão 3.5.0 (2016), que fora atualizada em dezembro de 2016, o programa também pode ser adquirido em versões anteriores, que permanecem no *site* do autor à disposição dos usuários. Para esta pesquisa, foi utilizada a

---

<sup>15</sup> O programa AntConc está disponível para ser baixado gratuitamente no site de Laurence Anthony, criador do programa. Disponível em: <<http://www.laurenceanthony.net/software/antconc/>>. Acesso em: 14 ago 2016.

versão 3.4.4w (2014), que, até o dia de realização das análises, era a única - dentre as mais recentes - compatível com o sistema Windows 10 (FIG. 1).

**FIGURA 1 - Interface do programa AntConc 3.4.4w**



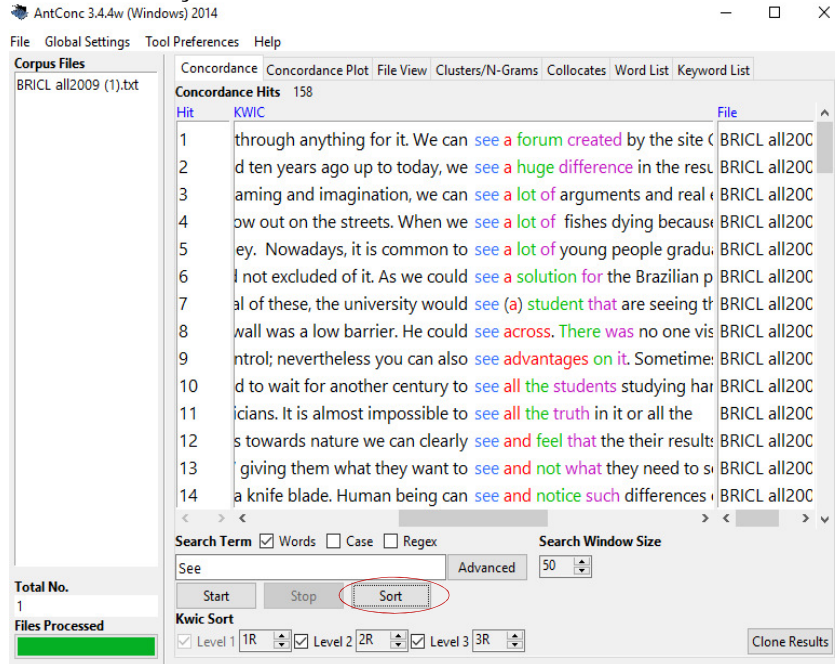
Fonte: AntConc (ANTHONY, 2014).

São funções/ferramentas disponíveis no AntConc:

- **Collocates:** essa função produz uma lista com todos os colocados de um termo de busca, ou seja, ela destaca as palavras que ocorrem ao redor do nóculo, em posições preestabelecidas pelo usuário (a primeira, a segunda e a terceira palavra à esquerda ou à direita, por exemplo), favorecendo a investigação de padrões não sequenciais de linguagem.
- **Concordance Plot Tool:** esse instrumento proporciona ao usuário observar a distribuição do nóculo dentro do *corpus* por meio de um gráfico, ou seja, o usuário vê o resultado da busca a partir de uma imagem gráfica, cujo formato é semelhante a um código de barras, que mostra imageticamente onde ocorre a palavra buscada em cada um dos textos contidos no *corpus*.
- **Concordance Tool:** essa ferramenta produz linhas de concordâncias e mostra o resultado das buscas no formato KWIC (*KeyWord In Context*), ou seja, as palavras de busca são apresentadas acompanhadas dos textos que as cercam, permitindo ao usuário verificar seu uso no contexto em que aparecem no *corpus*. A função “Sort” (sortir ou selecionar, FIG. 2),

que pode ser usada juntamente com a *Concordance Tool*, cria destaque para os termos de busca e para as palavras próximas a eles por meio de cores diferentes, facilitando a visualização do nódulo e das palavras que os cercam.

**FIGURA 2 - Função “Sort” em linha de concordância do AntConc 3.4.4w**

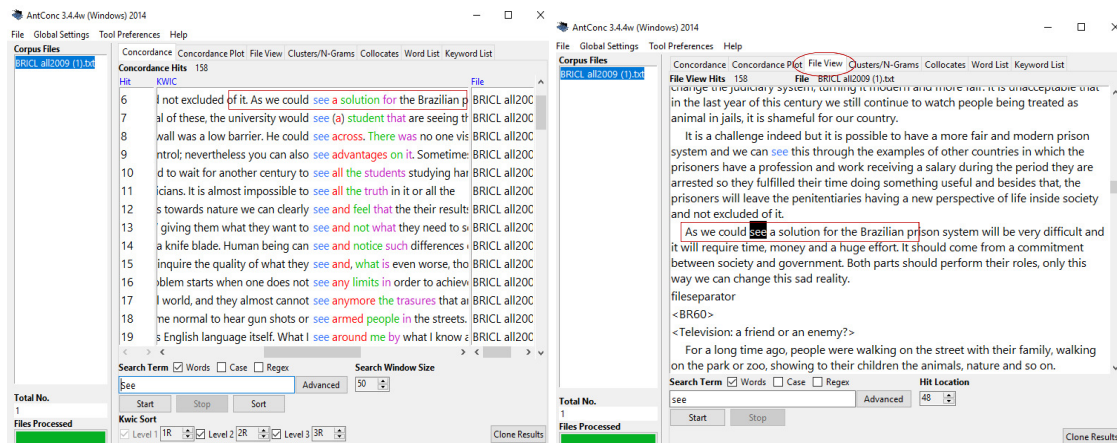


Fonte: AntConc (ANTHONY, 2014).

- **Word List:** esse mecanismo gera uma listagem contendo todas as palavras do *corpus* selecionado, individualmente, e elas são exibidas juntamente com o número no *ranking* das palavras e sua frequência bruta (ou absoluta). É possível ao usuário observar as palavras em ordem alfabética (que pode ser reversa, para uso em estudos de sufixos, por exemplo) ou pela frequência.
- **Keyword List:** essa ferramenta possibilita a comparação das palavras entre o *corpus* de estudo e o de referência. Por meio dela, cria-se uma lista de palavras-chaves que podem ser confrontadas com as palavras do *corpus* de referência, possibilitando ao usuário identificar palavras peculiares em um *corpus*.
- **Clusters:** essa ferramenta produz listas de agrupamentos lexicais (ou *clusters*) e exhibe as palavras que se agrupam ao nódulo usado na busca, permitindo ao usuário encontrar no *corpus* expressões e/ou multipalavras comuns de uma língua e recorrentes nas concordâncias do *corpus*.

- **File View Tool:** esse mecanismo exibe a palavra de busca dentro do texto ao qual ela pertence. Assim, ao buscar por uma palavra específica no *corpus* e selecioná-la usando essa função, o usuário obtém como resultado o texto completo do qual foi extraída a palavra (FIG. 3), favorecendo uma investigação mais detalhada do resultado gerado nas outras ferramentas do AntConc.

**FIGURA 3 - A ferramenta File View Tool no AntConc 3.4.4w**



Fonte: AntConc (ANTHONY, 2014).

Nesta pesquisa, foram usadas quatro das ferramentas disponíveis no programa: *Wordlist*, *Concordance Tool*, *File View Tool* e *Clusters*, conforme será mostrado mais adiante.

## 3.2 Os dados utilizados na pesquisa

### 3.2.1 *Corpus de estudo*

O *corpus* de estudo adotado para a realização deste trabalho é um componente brasileiro do *International Corpus of Learner English* (ICLE), intitulado *Brazilian subcorpus of the ICLE* (Br-ICLE). A coleção é formada por composições completas de alunos brasileiros que estudam a língua inglesa e que já possuem um conhecimento avançado do idioma. O objetivo do projeto é propiciar a identificação e a descrição de padrões léxico-gramaticais recorrentes, bem como traços de regularidades que permitam aos pesquisadores sugerir e compreender as escolhas dos alunos brasileiros no que concerne à escrita em língua inglesa. O trabalho pode colaborar ainda para a tomada de decisões pedagógicas sobre ensino e aprendizagem do inglês entre os brasileiros.

Ainda em processo de construção, sob a coordenação de Antônio Berber Sardinha, professor e pesquisador da PUC-SP, o Br-ICLE será composto por 250 mil palavras. Sua compilação reúne ensaios produzidos por alunos de diferentes universidades brasileiras, tais como a Universidade Federal do Rio de Janeiro (UFRJ), a Universidade Federal de Minas Gerais (UFMG) e a Pontifícia Universidade Católica de São Paulo (PUC- SP).

Para a realização desta pesquisa, foi usada a versão do Br-ICLE com textos produzidos até o ano de 2009, que contém aproximadamente 160 mil palavras. Cada redação contida no *corpus* de estudo foi produzida em formato argumentativo e apresenta no mínimo 500 e no máximo 1000 palavras, seguindo os mesmos critérios de produção (tipo textual, tamanho, temática etc.) do *corpus* que originou o projeto. Os textos foram redigidos com base em temáticas sugeridas pelos organizadores do ICLE, conforme mostra o QUADRO 4:

---

**QUADRO 4 - Títulos e/ou temas sugeridos pelo projeto ICLE para redações acadêmicas**

---

**Assuntos sugeridos para as composições acadêmicas**

---

- (1) Crime does not pay.
- (2) The prison system is outdated. No civilized society should punish its criminals: it should rehabilitate them.
- (3) Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.
- (4) A man/woman's financial reward should be commensurate with their contribution to the society they live in.
- (5) The role of censorship in Western society.
- (6) Marx once said that religion was the opium of the masses. If he was alive at the end of the 20th century, he would replace religion with television.
- (7) All armies should consist entirely of professional soldiers: there is no value in a system of military service.
- (8) The Gulf War has shown us that it is still a great thing to fight for one's country.
- (9) Feminists have done more harm to the cause of women than good.
- (10) In his novel Animal Farm, George Orwell wrote "All men are equal: but some are more equal than others". How true is this today?
- (11) In the words of the old song "Money is the root of all evil".
- (12) Europe
- (13) In the 19th century, Victor Hugo said: "How sad it is to think that nature is calling out but humanity refuses to pay heed." Do you think it is still true nowadays?
- (14) Some people say that in our modern world, dominated by science technology and industrialisation, there is no longer a place for dreaming and imagination. What is your opinion?

---

Fonte: Br-ICLE<sup>16</sup>.

---

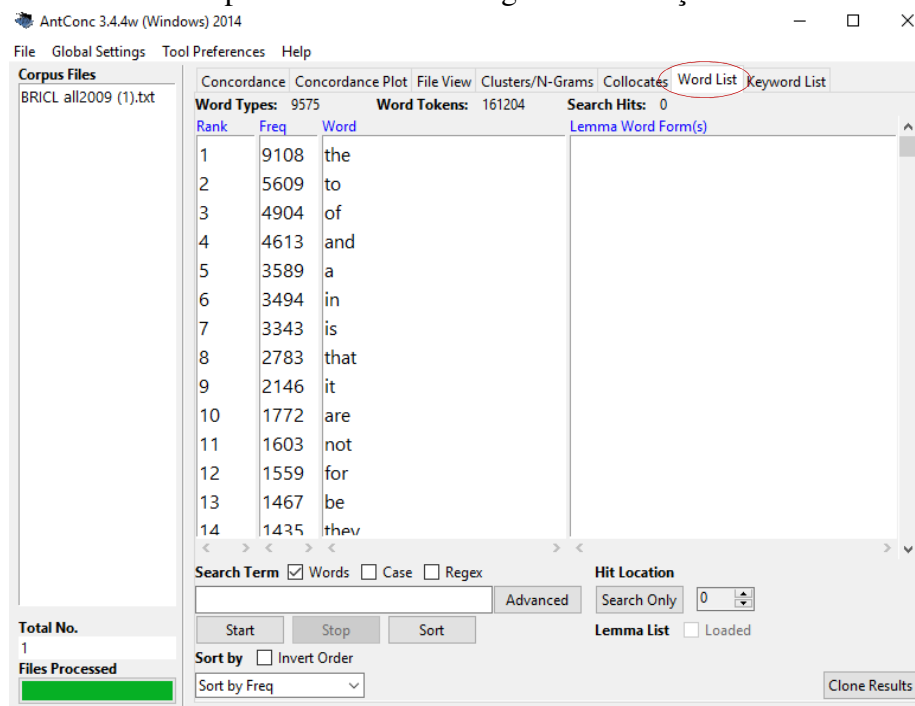
<sup>16</sup> Informações retiradas do *site* do Br-ICLE. Disponível em: <<http://www2.lael.pucsp.br/corpora/bricle/>>. Acesso em: 12 jul. 2016.

As condições de produção dos textos podem ser variáveis, ou seja, pode haver no *corpus* textos escritos, por exemplo, sob tempo demarcado ou não, com ou sem o auxílio de material de consulta, entre outros fatores. Por isso, ao contribuir com uma redação, cada aprendiz preenche o formulário *Learner Profile* (ANEXO A), que além de registrar as informações pertinentes às circunstâncias do processo de escrita de cada ensaio, também coleta dados acerca do estudante, tais como sexo, idade, nacionalidade, formação educacional, língua materna, que outras línguas domina, entre outras características.

### 3.2.1.1 Dados do corpus de estudo

A fim de conhecer melhor o *corpus* de estudo, manipulamos os dados disponíveis com a ajuda do concordanciador AntConc. O primeiro procedimento adotado foi a criação de uma lista de palavras, que nos permitiu observar a relação de vocábulos contidos no Br-ICLE. Todos os dados (armazenados em formato de texto – txt) foram migrados para o *software* e, utilizando a ferramenta *Wordlist*, carregamos a lista de palavras, conforme mostra a FIG. 4:

**FIGURA 4** - Lista de palavras do Br-ICLE gerada na função *Wordlist* do AntConc



Fonte: AntConc (ANTHONY, 2014).

O procedimento realizado gerou os seguintes dados acerca do Br-ICLE:

**TABELA 1** - Dados estatísticos do *corpus* de estudo

Fonte	Número de Types/Vocábulos	Número de Tokens/Ocorrências
Br-ICLE	9.575	161.204

Fonte: Elaborado pela autora.

### 3.2.2 *Corpus de referência para a análise quantitativa*

O *corpus* de referência escolhido para esta pesquisa foi o *Louvain Corpus of Native English Essays* (LOCNESS), que, assim como o *corpus* de estudo, também é composto por textos acadêmicos produzidos por estudantes do inglês. Os colaboradores são aprendizes nativos (britânicos e americanos) da língua inglesa e estudantes de instituições como as Universidades da Carolina do Sul e Michigan, nos Estados Unidos, e de Surrey, na Inglaterra. Esse *corpus* foi escolhido para validar os padrões encontrados no Br-ICLE devido à sua semelhança com o *corpus* de estudo (ambos estão em linguagem escrita, contêm ensaios e textos argumentativos e foram produzidos em meio acadêmico).

O LOCNESS foi compilado pela equipe da professora Sylviane Granger (Universidade Católica de Louvain, Bélgica) - que também dirigiu o projeto ICLE - e, atualmente, ele conta com 324.304 palavras, oriundas de 322 textos diferentes, conforme mostra a tabela a seguir:

**TABELA 2** - *Corpora* presentes no LOCNESS e quantidades de textos em cada *corpus*

Fonte	Número de textos
<i>Argumentative essays written by American university students</i>	176
<i>Literary-mixed essays written by American university students</i>	56
<i>Argumentative and literary essays written by British university students and A-level students</i>	90
<b>Total</b>	322

Fonte: LOCNESS<sup>17</sup>.

A maior parte dos textos reunidos nesse projeto é datada de 1995 e as circunstâncias de escrita dos textos, assim como no *corpus* de estudo, também apresentam variáveis, como o

<sup>17</sup> Informações retiradas do site do LOCNESS. Disponível em: <<https://www.uclouvain.be/en-cecl-locness.html>>. Acesso em: 12 jul. 2016.

uso de dicionários, livros e outros materiais de referência, o tempo cronometrado ou não para redigir o ensaio, a idade dos participantes, entre outros. As condições de produção das composições podem ser vistas no ANEXO B. As redações do LOCNESS abordam assuntos similares ou temáticas relacionadas aos temas propostos nas redações do projeto ICLE (e, conseqüentemente, também do Br-ICLE), o que reforça seu caráter de *corpus* propício para o estudo comparativo do fenômeno analisado nesta pesquisa. O QUADRO 5 mostra os assuntos sugeridos nas composições que formam o *corpus* de referência:

**QUADRO 5 - Temas/assuntos sugeridos para as composições do LOCNESS**

<b>ASSUNTOS SUGERIDOS PARA AS COMPOSIÇÕES ACADÊMICAS</b>	
<ul style="list-style-type: none"> <li>- euthanasia</li> <li>- controversy in the classroom</li> <li>- capital punishment</li> <li>- does affirmative action work?</li> <li>- yoga</li> <li>- nuclear power</li> <li>- values and consequences of school interaction</li> <li>- pride or segregation</li> <li>- surrogate motherhood</li> <li>- can we afford wellness</li> <li>- prozac: the wonder drug</li> <li>- homosexuality</li> <li>- animal testing</li> <li>- prayer in schools</li> <li>- praying for a miracle</li> <li>- roles, feminism, etc</li> <li>- US government</li> </ul>	<ul style="list-style-type: none"> <li>- sex equality</li> <li>- teenagers</li> <li>- aids</li> <li>- orphanages</li> <li>- profit: good or evil</li> <li>- freedom of the press</li> <li>- sex in schools</li> <li>- welfare reforms needs a return to family values</li> <li>- the cost of grass</li> <li>- abortion</li> <li>- ethics</li> <li>- would anyone care for a drink</li> <li>- cheating in colleges</li> <li>- O. J. Simpson</li> <li>- suicide</li> <li>- Football</li> </ul>

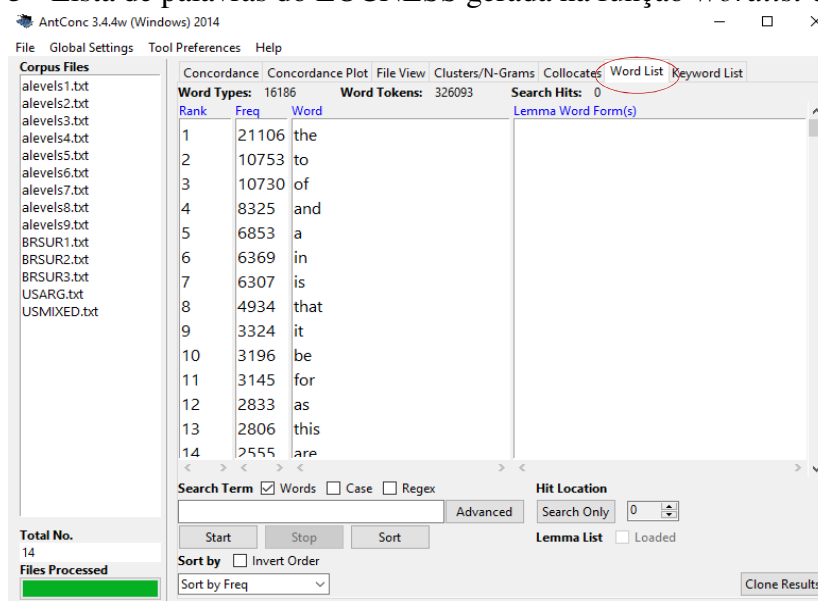
Fonte: LOCNESS.

*3.2.2.1 Dados do corpus de referência*

Assim como ocorreu com o *corpus* de estudo, o *corpus* de referência também fora manipulado no AntConc, para que pudéssemos realizar uma primeira observação dos dados contidos na compilação. Os procedimentos adotados nesta etapa foram os mesmos usados no Br-ICLE, sendo o primeiro passo o carregamento de uma lista de palavras (ferramenta *Wordlist*, FIG. 5).



**FIGURA 5** - Lista de palavras do LOCNESS gerada na função *Wordlist* do AntConc



Fonte: AntConc (ANTHONY, 2014).

Os resultados gerados no procedimento são os seguintes:

**TABELA 3** - Dados estatísticos do *corpus* de referência

Fonte	Número de Types/Vocábulos	Número de Tokens/Ocorrências
LOCNESS	16.186	326.093

Fonte: elaborado pela autora.

A partir dos resultados obtidos na análise dos dados dos *corpora*, percebeu-se que o Br-ICLE tem cerca de 60% do número de vocábulos do LOCNESS. A informação confirma a validade do *corpus* de referência como material de controle na análise, já que, conforme explica Sardinha (2004), para que os resultados da comparação sejam significativos, é ideal que o *corpus* de referência tenha entre duas e cinco vezes o tamanho do *corpus* de estudo.

### 3.2.3 *Corpus de referência para a análise qualitativa*

Para a etapa da análise qualitativa dos dados desta pesquisa também foi utilizado como *corpus* de referência o *British National Corpus* – BNC, que contém 100 milhões de palavras e sua composição é 90% de parte escrita e 10% de parte oral. A seção escrita reúne extratos de textos de jornais, periódicos especializados, cartas, memorandos, livros de ficção, além de ensaios acadêmicos e outros tipos textuais. Já a seção oral, consiste de transcrições de conversas informais e formais, entrevistas, discursos políticos e religiosos, reuniões, dentre

outros subgêneros. O trabalho de construção do BNC teve início em 1991 e foi finalizado em 1994. Entretanto, o *corpus* passou por processos de revisão e a última versão revista da compilação foi lançada em 2007.

Nesta pesquisa, o BNC foi usado para fins de verificação da aceitabilidade dos usos de verbos frasais conforme os padrões da língua inglesa. Por isso, na etapa de análise consideraram-se somente textos cujos tipos eram semelhantes aos verificados no *corpus* de estudo, ou seja, somente textos argumentativos foram analisados. Em casos isolados, por ausência de estrutura válida para comparação nos termos desta pesquisa, utilizou-se a amostra de outro tipo textual, como será mostrado no capítulo sobre os resultados.

### 3.3 Procedimentos gerais

Para o procedimento de análise, foi adotada uma metodologia que considerou duas diferentes perspectivas: uma de caráter quantitativa, que foi aplicada para se testarem as hipóteses de preferência pelo uso ou não uso dos *phrasal verbs* por estudantes brasileiros, o que foi verificado por meio de análises de frequência e de padrões lexicais; e outra de caráter qualitativo, que foi aplicada para se responder às questões relacionadas aos motivos de ocorrência ou não do fenômeno estudado e do comportamento dos aprendizes em suas escolhas, considerando, para isso, uma análise aprofundada das ocorrências e seus contextos.

A combinação dessas duas perspectivas é sugerida por Mair (1991), que explica que a análise de *corpus* nunca deve ser somente estatística ou qualitativa.

O papel do *corpus*, afinal, não é só proporcionar uma base de dados limitada e representativa para a análise estatística, mas também fornecer dados autênticos e realistas, a leitura atenta daquilo que permitirá ao linguista abordar a gramática de uma perspectiva funcional e discursiva<sup>18</sup>. (MAIR, 1991, p. 77).

A fim de tornar a análise dos dados mais abrangente e segura, especialmente considerando as limitações que uma análise puramente automática pode configurar nesse tipo de pesquisa e de fenômeno (cf. WAIBEL, 2007, p. 67), o tratamento dos dados foi realizado por processo semiautomático, ou seja, parte dos verbos frasais foi analisada com o auxílio do AntConc e parte foi examinada manualmente, conforme será visto adiante.

---

<sup>18</sup> No original: “*The role of the corpus, after all, is not only to provide a limited and representative data-base for statistical analysis, but also to provide authentic and realistic data, the close reading of which will allow the linguist to approach grammar from a functional and discourse perspective*”.

### 3.4 Seleção, extração e validação dos *phrasal verbs*

A extração dos *phrasal verbs* constitui uma das mais problemáticas questões de estudos que analisam essa estrutura gramatical, especialmente quando os *corpora* são volumosos, porque, além da grande quantidade de verbos que podem ser encontrados formando combinações com partículas, a tarefa de analisar, contextualmente, cada um desses verbos multpalavras para se verificar se, de fato, desempenham função de *phrasal verb* torna-se bastante exaustiva. Visando a estabelecer uma maneira mais viável de selecionar os verbos frasais, testamos ao longo desse estudo duas metodologias diferentes, que serão descritas a seguir: uma com foco nos verbos (que foi descartada por demonstrar não ser muito eficaz para alcançarmos os objetivos da pesquisa) e outra com foco nas partículas (esta, por ser mais assertiva, foi adotada como procedimento definitivo).

Como já demonstrado, o objeto de pesquisa estabelecido nesta análise são os verbos frasais do inglês que, para fins desta investigação, serão considerados conforme a definição proposta, principalmente, por Downing e Locke (2006) e Downing (2015) - ou seja, combinações formadas por um verbo lexical e uma partícula adverbial que, juntas, apresentam uma extensão de sentido diferente do significado individual das palavras que as compõem. Posto isso, o primeiro passo na realização do estudo foi a seleção dos verbos lexicais componentes dessas combinações.

Através do procedimento realizado com o AntConc, foram geradas listas de palavras (*wordlist*) de ambos os *corpora* (cf. APÊNDICES A e B), das quais foram selecionados os verbos mais frequentes de cada um deles. Como o número de verbos encontrados foi muito grande, optamos por fazer uma triagem que pudesse tornar a análise mais viável e decidimos estabelecer novas listas contendo apenas os 84 verbos mais usados nos *corpora*. Dessa lista, foram extraídos os verbos lexicais de maior frequência, bem como eliminados os verbos auxiliares<sup>19</sup> e também os modais que, por não formarem combinação com partículas adverbiais, não se adequam ao nosso estudo. Como resultado, obtivemos listas de 40 verbos lexicais (cf. APÊNDICE C para a lista do *corpus* de estudo), que foram analisados individualmente nas listas de concordância a fim de se identificar as combinações com partículas adverbiais e de se realizar a extração dos *phrasal verbs*. Esse procedimento, entretanto, mostrou-se pouco eficaz, pois, ao longo das análises, vimos que verbos

---

<sup>19</sup> Apesar de formarem verbos frasais em alguns casos, o “do” e o “have” não foram considerados nessa pesquisa em função de aparecerem nos *corpora*, predominantemente, exercendo função de verbo auxiliar.

importantes para a composição de verbos frasais, como o “*put*” e o “*come*” - que são, inclusive, apresentados nas seleções de verbos produtivos na formação de *phrasal verbs* propostas por Biber et. al (1999, p. 413) e também por Carter e McCarthy (2006, p. 430) - não apareciam na lista, embora formassem muito mais composições com partículas do que muitos dos verbos selecionados na lista final. Dessa forma, optamos por abandonar essa metodologia e reiniciar o estudo utilizando outro método, dessa vez, baseado nas partículas.

O segundo método aplicado, assim como o primeiro, também partiu da lista de palavras (*Wordlist*) do AntConc, procedimento que nos permitiu conhecer melhor os *corpora* (cf. itens 3.2.1.1 e 3.2.2.1). Porém, dessa vez, decidimos fazer a busca dos verbos frasais partindo das partículas adverbiais, já que, ao contrário dos verbos, estas são limitadas, o que torna a análise menos exaustiva e igualmente (ou até mais) eficaz. Cabe, então, ressaltar que, por ter demonstrado conferir mais credibilidade ao processo de análise, este método foi adotado como preferencial e, logo, foi utilizado como procedimento definitivo na extração dos verbos frasais e na condução da pesquisa.

A seleção das partículas adverbiais, por sua vez, foi baseada na lista proposta por Waibel (2007, p. 65), que considera uma combinação dos advérbios mencionados por Quirk et. al (1985) e Johansson e Hofland (1989), conforme mostra o QUADRO 6.

**QUADRO 6** - Lista de partículas adverbiais

Aback	Ahead	Around	Back	Down	In	Out	Round	Through
About	Along	Aside	Behind	Forth	Off	Over	Under	Together
Across	Apart	Away	By	Forward	On	Past	Up	Without

Fonte: Elaborado pela autora com dados extraídos de Waibel, 2007.

A partir dessa lista, analisamos os *corpora* utilizando a ferramenta *Clusters*, que nos permitiu verificar se as palavras mais próximas dos termos de busca (partículas adverbiais) eram verbos lexicais que formavam *phrasal verbs*. Dessa forma, com o auxílio dessa análise, foi possível eliminar aquelas composições que não se encaixavam em nossa definição, como, por exemplo, as formadas por verbos modais ou auxiliares, ou ainda as expressões que ocorriam em usos nominais (ex. “[...] *buying them make-up*”, em que a combinação verbo + partícula funciona como substantivo, significando “maquiagem”). Em função da estrutura dos verbos frasais (verbo lexical anterior à partícula adverbial), a configuração dessa etapa considerou como parâmetro a seleção de *clusters* com no mínimo e no máximo cinco palavras. Tal configuração foi adotada para favorecer que os verbos frasais pudessem ser

capturados mesmo que as partículas estivessem posicionadas a certa distância do verbo (no limite de até quatro palavras antes da partícula), procedimento bastante importante se considerarmos que as partículas podem se movimentar na oração. O termo de busca foi posicionado à direita, ou seja, seriam visualizadas todas as palavras à esquerda do termo de busca, respeitando a estrutura dos verbos frasais (verbo seguido de partícula), como mostra a FIG. 6. O formato de observação favoreceu a captura de estruturas que contêm também pronomes e objetos diretos entre o verbo e a partícula (característica dos verbos frasais transitivos), como em “*She switched it off*” e “*Set the alarm off*”.

**FIGURA 6** - Exemplo de análise da partícula *up* usando a ferramenta *Clusters*

The screenshot displays the AntConc 3.4.4w (Windows) 2014: User Settings interface. The 'Clusters/N-Grams' tool is active, showing a search for the term 'up'. The 'Cluster Size' is set to 5. The results table is as follows:

Rank	Freq	Range	Cluster
19	1	1	as well, to sum up
20	1	1	at all, to sum up
21	1	1	bad guys always end up
22	1	1	be someone to round up
23	1	1	brave enough to walk up
24	1	1	can make them grow up
25	1	1	children, we have grown up
26	1	1	coll people", teens look up
27	1	1	complex sense of society up
28	1	1	developed into the most up
29	1	1	dichotomy, this logical split-up
30	1	1	do not only popped up
31	1	1	dreams keep being built up
32	1	1	dreams, for then go up
33	1	1	efforts, another question comes up
34	1	1	end, they usually gave up
35	1	1	famous square, to sum up
36	1	1	food, you picked it up
37	1	1	for the life gave up

The interface also shows search options: Search Term (up), Words checked, Case unchecked, Regex unchecked, N-Grams unchecked, Cluster Size (Min: 5, Max: 5), Sort by (Freq), and Search Term Position (On Left unchecked, On Right checked).

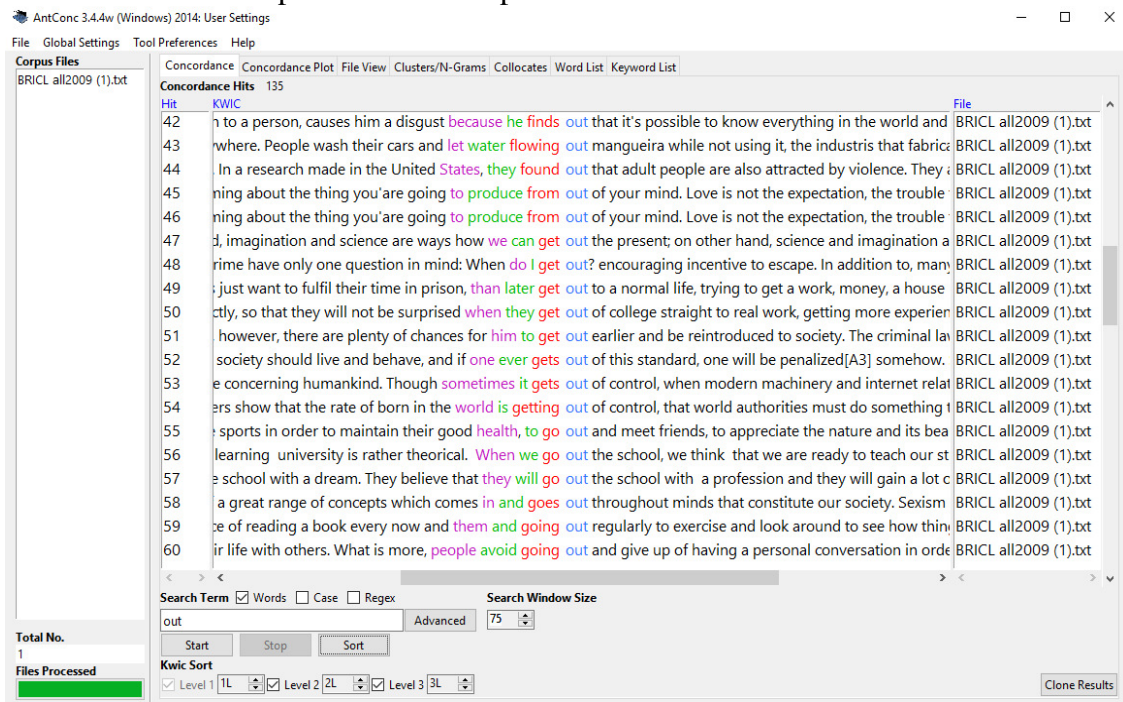
Fonte: AntConc (ANTHONY, 2014).

Como resultados dessa etapa, foram extraídas todas as combinações de verbo + partícula com potencial para a formação de verbos frasais. Os dados foram transferidos para planilhas eletrônicas e cada *phrasal verb* encontrado foi individualmente analisado também com o auxílio da ferramenta *Concordance*. Essa última etapa teve o objetivo de validar a seleção anterior, já que favoreceu a observação dos verbos frasais em seus contextos de uso. A partir desse procedimento, foi possível verificar se as combinações encontradas possuíam valor real de *phrasal verb* ou se configuravam-se como combinações livres entre verbos e advérbios, conforme concepção de Biber, Conrad e Leech (2002):

Em contraste com os *multi-word verbs*, as combinações livres consistem de um verbo lexical formado por uma única palavra seguido de um advérbio. Ou preposição com um significado distinto (por exemplo, *come down, go back*). Na prática, é difícil fazer uma distinção absoluta entre as combinações livres e os *multi-word verbs*. É melhor pensar em um contínuo no qual algumas utilizações de verbos são relativamente idiomáticas. [...] As palavras em combinação livre retêm seus próprios significados.<sup>20</sup> (BIBER; CONRAD; LEECH, 2002, p. 124-125).

Em outras palavras, nessa etapa foi possível verificar se as combinações encontradas constituíam-se, na verdade, de um verbo combinado a um sintagma preposicionado (um adjunto ou complemento locativo), como em “*She went down the hill*”, em que *down* funciona como uma preposição e a expressão “*down the hill*” como um complemento locativo (ou adjunto), que modifica o verbo e todo o sintagma preposicionado, não somente a partícula “*down*”. A FIG. 7 ilustra a utilização da ferramenta *Concordance*, do AntConc.

**FIGURA 7** - Exemplo de análise da partícula *out* usando a ferramenta *Concordance*



Fonte: AntConc (ANTHONY, 2014).

A etapa de análise das linhas de concordância foi essencial para a extração definitiva dos verbos frasais, já que revelou a ocorrência de combinações verbais iguais expressando valores diferentes dentro do mesmo *corpus*. É o caso de “*live in*”, que, apesar de aparecer

<sup>20</sup> No original: “*In contrast to multi-word verbs, free combinations consist of a single-word lexical verb followed by an adverb or preposition with a separate meaning (e.g. come down, go back). In practice, it is hard to make an absolute distinction between free combinations and multi-word verbs. It is better to think of a continuum where some uses of verbs are relatively free and others relatively idiomatic. [...]the words in free combination retain their own meanings*”.

predominantemente como combinação livre, com significado literal de viver em algum lugar (ex. “*Unfortunately, we **live in** a society where a lot of people still think that [...]*” - Br-ICLE), também ocorreu como verbo frasal, com significado de “viver no mesmo local onde se estuda e ou trabalha” (ex. “[...] *many stores had several floors of apartments above them to accommodate families to **live in.***” - LOCNESS). Em casos como esse, as combinações livres foram descartadas da análise, mas as ocorrências que tinham função de verbo frasal foram mantidas.

A etapa de análise por linha de concordância também colaborou para a eliminação de algumas combinações como “*depend on*” que, apesar de figurar entre as combinações com o maior número de ocorrências em ambos os *corpora*, não aparece exercendo função de verbo frasal, mas sim de verbo preposicionado<sup>21</sup>, como pode ser visto nos exemplos a seguir, retirados das linhas de concordância do Br-ICLE e do LOCNESS e exibidos no QUADRO 7.

**QUADRO 7** - Linhas de concordâncias para o lema “*depend*” + partícula “*on*”

<b>Combinação:</b> <i>depend on</i> = ser controlado ou determinado por algo ou alguém; ser capaz de confiar em algo ou alguém; precisar do apoio ou suporte de algo ou alguém. Sinônimos: <i>rely on, according to</i> . ( <i>English Oxford Living Dictionaries</i> , 2016)	
<b>Concordância</b>	<b>Estrutura</b>
[...] <i>since no man is equal to the other, it will <u>depend on</u> the power that this person has over the community</i> . (Br-ICLE)	Verbo + preposição
[...] <i>Although it <u>depends on</u> the kind of job the student want to get, the majority of jobs expect students background knowledge in the field</i> . (Br-ICLE)	Verbo + preposição
[...] <i>As long as unwed mothers know that they can <u>depend on</u> the government for financial support, what incentive do they have to stop having illegitimate kids?</i> (LOCNESS)	Verbo + preposição
[...] <i>The president could only exercise his power <u>depending on</u> the mood of the parliament</i> . (LOCNESS)	Verbo + preposição

Fonte: Elaborado pela autora.

Acerca das decisões metodológicas, é importante ressaltar que, apesar de haver diversos recursos de etiquetagem disponíveis atualmente, o que agilizaria o trabalho de extração de verbos e partículas pela automatização dos dados, optamos por manter na

<sup>21</sup> Haja vista que algumas palavras podem exercer tanto função de preposição, quanto de advérbio, uma questão importante é a realização de testes que nos permitam distinguir cada combinação verbal em uso. No caso de “*depend on*”, foram utilizados os testes de mobilidade da partícula adverbial, propostos por Downing (2015, p. 57) e Quirk et. al (1989, p. 1157), descritas no tópico 2.2.2 deste trabalho, e o de passivização de sentenças, proposto por Celce-Murcia e Larsen-Freeman (1999, p. 430). Outras combinações de alta frequência nos *corpora*, como “*focus on*” e “*base on*” também foram analisadas por meios desses testes e, por não corresponderem à estrutura de verbos frasais - conforme as propostas de Downing e Locke (2006) e Downing (2015) -, adotadas neste trabalho, também foram eliminadas da lista de *multi-word verbs* a serem analisados.

pesquisa um processo de análise semiautomático, ou seja, com o uso de ferramentas eletrônicas para a manipulação dos dados, mas com o exame manual das linhas de concordância. A decisão, que tem base em experiências de pesquisas anteriores (COSTA, 2012), foi tomada a fim de evitar erros na classificação das palavras, já que nem sempre é fidedigno o resultado proporcionado pelos etiquetadores, especialmente quando se trata de termos referentes a unidades gramaticais relacionadas com os *multi-word verbs* (cf. item 2.4.4.2). A seleção manual nos permitiu identificar as funções gramaticais de cada uma das partículas, individualmente, em seu contexto de uso, favorecendo a correta classificação dos verbos multipalavras.

Sobre a etapa de seleção dos verbos, é relevante comentar ainda que apenas a forma infinitiva dos verbos lexicais foi considerada. Durante a extração e produção dos resultados, entretanto, levou-se em consideração também a inclusão das outras formas verbais existentes para as palavras escolhidas. As formas negativas, por não aparecerem de forma sistemática nos *corpora*, não foram consideradas.

Em suma, os procedimentos metodológicos adotados foram:

1. Observação das estatísticas gerais do *corpus* de estudo e relação da lista de palavras na ferramenta *Wordlist*;
2. Seleção das partículas adverbiais mais propensas à formação de *phrasal verbs*;
3. Obtenção e análise da lista de *Clusters* (usando como base as partículas selecionadas no passo anterior) para extração inicial dos *phrasal verbs*;
4. Extração dos verbos lexicais combinados com partículas no *corpus* de estudo;
5. Extração dos verbos lexicais combinados com partículas no *corpus* de referência;
6. Obtenção e exame das linhas de concordância para extração definitiva dos *phrasal verbs*;
7. Validação dos *phrasal verbs* encontrados nos *corpora*;
8. Comparação e análise dos *phrasal verbs* nos *corpora*.

Os resultados dos procedimentos aqui descritos serão mostrados e discutidos nos capítulos 4 (análise quantitativa) e 5 (análise qualitativa).



### 3.5 Procedimentos para a análise qualitativa dos *phrasal verbs*

Para a análise qualitativa, que verificou aspectos como competência e desempenho dos aprendizes em relação ao uso dos verbos frasais, o procedimento metodológico foi, basicamente, o exame das linhas de concordância para apreciação contextual e semântica dos *phrasal verbs* em uso. Haja vista o grande número de verbos frasais encontrados, e considerando que seria inviável a análise individual de cada um deles em termos da proposta deste trabalho, decidimos utilizar como amostra apenas alguns dos *phrasal verbs* mais problemáticos no *corpus* de estudo, a fim de exemplificar os comentários e observações relativos ao uso de verbos frasais em textos argumentativos de nativos e não nativos do inglês.

Para essa etapa, utilizamos como suporte o *British National Corpus* (BNC) e também os dicionários de língua inglesa *English Oxford Living Dictionary* (edição *on-line*, 2016), *Longman Dictionary of Contemporary English* (Pearson Education Limited, nova edição, 2003), *Oxford Phrasal Verbs Dictionary for Learners of English* (Oxford University Press, 2001). Essas ferramentas serviram como parâmetro para a comparação dos dados encontrados e nos ajudaram a descobrir se a utilização dos verbos frasais por parte dos aprendizes coincidia ou não com o uso indicado nos padrões do inglês nativo, auxiliando na averiguação de aspectos como o da naturalidade ou não no emprego de PVs, bem como para a sugestão de possíveis motivações para os problemas encontrados (influência – interferência e/ou transferência da língua materna, criatividade, uso dos verbos frasais como desvios colocacionais etc.).

## 4 ANÁLISE E DISCUSSÃO DOS RESULTADOS: ABORDAGEM QUANTITATIVA

Neste capítulo, apresentamos os resultados encontrados a partir da análise quantitativa dos dados. Em um primeiro momento, serão mostrados os dados estatísticos dos *corpora* de estudo e de referência. Em seguida, serão apresentados e discutidos aspectos como a frequência das partículas adverbiais e dos verbos frasais em ambos os *corpora* (Br-ICLE e LOCNESS), bem como alguns achados acerca da produção e da utilização de *phrasal verbs* na escrita de aprendizes nativos e não nativos da língua inglesa.

### 4.1 Dados estatísticos dos *corpora* e parâmetro de apresentação dos resultados

Conforme apresentado anteriormente (cf. tópicos 3.2.1.1 e 3.2.2.1), os *corpora* utilizados nesta pesquisa apresentam as seguintes dimensões:

**TABELA 4** - Dados estatísticos dos *corpora* de estudo e de referência

Fonte	Número de <i>Types/Vocábulos</i>	Número de <i>Tokens/Ocorrências</i>	<i>Type-Token ratio</i>
Br-ICLE	9.575	161.204	5.93
LOCNESS	16.186	326.093	4.96

Fonte: Elaborado pela autora.

Esse quadro mostra que, no que concerne ao número de palavras (*tokens*), o *corpus* de referência é duas vezes maior que o *corpus* de estudo, como recomendado por BERBER SARDINHA (2004, p. 102). Por serem de grandezas numéricas diferentes, entretanto, foi necessário se fazer a normalização dos dados, a fim de se estabelecer a razão/proporção entre eles. O procedimento, que torna as contagens comparáveis, foi realizado considerando-se o cálculo simples de razão, que consiste na divisão de um número a por um número b, ou seja, em nossa pesquisa, dividimos o valor de *tokens* do LOCNESS pelo número de *tokens* do Br-ICLE e obtivemos o valor de 2,02 palavras<sup>22</sup>. Isso indica que, para cada palavra do *corpus* de estudo, temos 2,02 palavras no *corpus* de referência. Da mesma forma, todas as ocorrências

<sup>22</sup> Cálculo da razão dos *corpora*:  $326.093 / 161.204 = 2.022$ .

apresentadas no estudo também foram normalizadas, porém, haja vista que trabalhamos com dois *corpora* de tamanhos distintos (um classificado como pequeno-médio e outro como médio, na perspectiva de Berber Sardinha, 2004), optamos por adotar, na normalização<sup>23</sup>, a proporção de 100.000 palavras. Essa base foi escolhida para facilitar a visualização dos resultados, já que, se usássemos a proporção de 10.000, por exemplo, teríamos números muito pequenos, o que poderia afetar visualmente a frequência das palavras. Por outro lado, se usássemos a proporção de 1.000.000, as frequências seriam visualmente aumentadas, o que poderia causar certa estranheza na análise.

Para que pudéssemos verificar se os dados resultantes da comparação dos *corpora* eram estatisticamente significativos, em alguns procedimentos, aplicamos também o teste estatístico *Log-likelihood* (LL)<sup>24</sup> que, segundo Rayson (2002), é um dos que apresenta melhores resultados na comparação de frequências de itens ou expressões entre dois *corpora*. Trata-se de uma regra estatística que calcula a probabilidade de a diferença entre as ocorrências ser aleatória ou não. Em outras palavras, o teste nos permite observar se a diferença apresentada é casual ou não, o que é útil para a visualização do caráter de sobreuso e subuso dos verbos frasais.

## 4.2 Frequência e produtividade das partículas adverbiais

Conforme já apontado anteriormente, para a análise do uso de verbos frasais entre aprendizes nativos e não nativos do inglês proposta neste trabalho, foram selecionadas 27 partículas adverbiais (PAVs). Apesar de muitas delas também desempenharem função sintática de preposição, foram avaliadas nos resultados apenas as partículas cuja função é adverbial, já que nosso objeto de estudo é o verbo frasal definido como sendo a combinação entre um verbo lexical e um advérbio. Entretanto, neste primeiro tópico, consideramos todas as ocorrências das partículas a fim de estabelecermos um panorama de uso dessas palavras nos *corpora*.

---

<sup>23</sup> A fórmula usada no cálculo para normalização dos *corpora* foi regra de três: Frequência encontrada \* 100000/Valor de *tokens* do *corpus* = x.

<sup>24</sup> O ponto de corte do teste *Log-likelihood* é  $LL > 3.84$  para um  $p < 0.05$ , o que representa um grau de 95% de confiança. Ou seja, se o resultado do teste for maior que 3.84, há menos de 5% de chances de a diferença entre os *corpora* comparados serem aleatórias, o que torna o resultado significativo. Assim, quanto mais alto for o valor apresentado no LL, mais significativa será a diferença entre os *corpora*. Para a realização do teste *Log-likelihood* foi usada como referência a calculadora do UCREL (*University Centre for Computer Corpus Research on Language*), da Universidade de Lancaster. Disponível em: < <http://ucrel.lancs.ac.uk/llwizard.html> >. Acesso em: 12 ago. 2016.

Após a extração dos nódulos, verificamos que as PAVs representam 4% do número de palavras contidas tanto do *corpus* de estudo (CE), quanto do *corpus* de referência (CR). Em valores brutos, e sem considerar a função que desempenham no contexto (se com valor de preposição ou de advérbio), vemos que elas aparecem 6.630 vezes no Br-ICLE e 13.707 no LOCNESS, conforme mostra a TAB. 5.

**TABELA 5 - Frequência de partículas adverbiais nos corpora**

Br-ICLE					LOCNESS			
#	Partícula	Freq. Bruta	Freq. 100 mil	% norm. no CE	Partícula	Freq. Bruta	Freq. 100 mil	% norm. no CR
1	<i>In</i>	3474	2155	52,41%	<i>In</i>	6371	1954	46,51%
2	<i>On</i>	778	483	11,75%	<i>On</i>	1797	551	13,12%
3	<i>By</i>	680	422	10,26%	<i>By</i>	1703	522	12,43%
4	<i>About</i>	453	281	6,83%	<i>About</i>	578	177	4,21%
5	<i>Without</i>	192	119	2,89%	<i>Out</i>	569	174	4,14%
6	<i>Up</i>	162	100	2,43%	<i>Up</i>	481	148	3,52%
7	<i>Through</i>	145	90	2,19%	<i>Over</i>	333	102	2,43%
8	<i>Out</i>	135	84	2,04%	<i>Through</i>	295	90	2,14%
9	<i>Over</i>	120	74	1,80%	<i>Without</i>	232	71	1,69%
10	<i>Around</i>	85	53	1,29%	<i>Back</i>	178	55	1,31%
11	<i>Together</i>	67	42	1,02%	<i>Away</i>	151	46	1,09%
12	<i>Past</i>	66	41	1,00%	<i>Down</i>	150	46	1,09%
13	<i>Under</i>	49	30	0,73%	<i>Around</i>	147	45	1,07%
14	<i>Back</i>	41	25	0,61%	<i>Past</i>	143	44	1,05%
15	<i>Away</i>	39	24	0,58%	<i>Under</i>	128	39	0,93%
16	<i>Behind</i>	31	19	0,46%	<i>Off</i>	116	36	0,86%
17	<i>Off</i>	24	15	0,36%	<i>Together</i>	85	26	0,62%
18	<i>Down</i>	23	14	0,34%	<i>Along</i>	60	18	0,43%
19	<i>Forward</i>	18	11	0,27%	<i>Behind</i>	47	14	0,33%
20	<i>Along</i>	16	10	0,24%	<i>Forward</i>	34	10	0,24%
21	<i>Apart</i>	9	6	0,15%	<i>Across</i>	33	10	0,24%
22	<i>Ahead</i>	6	4	0,10%	<i>Apart</i>	25	8	0,19%
23	<i>Across</i>	5	3	0,07%	<i>Round</i>	18	6	0,14%
24	<i>Aside</i>	4	2	0,05%	<i>Ahead</i>	14	4	0,10%
25	<i>Forth</i>	3	2	0,05%	<i>Forth</i>	11	3	0,07%
26	<i>Round</i>	3	2	0,05%	<i>Aside</i>	8	2	0,05%
27	<i>Aback</i>	2	1	0,02%	<i>Aback</i>	-	-	0,00%
<b>TOTAL</b>		<b>6.630</b>	<b>4.113</b>	<b>100,00%</b>		<b>13.707</b>	<b>4.201</b>	<b>100,00%</b>

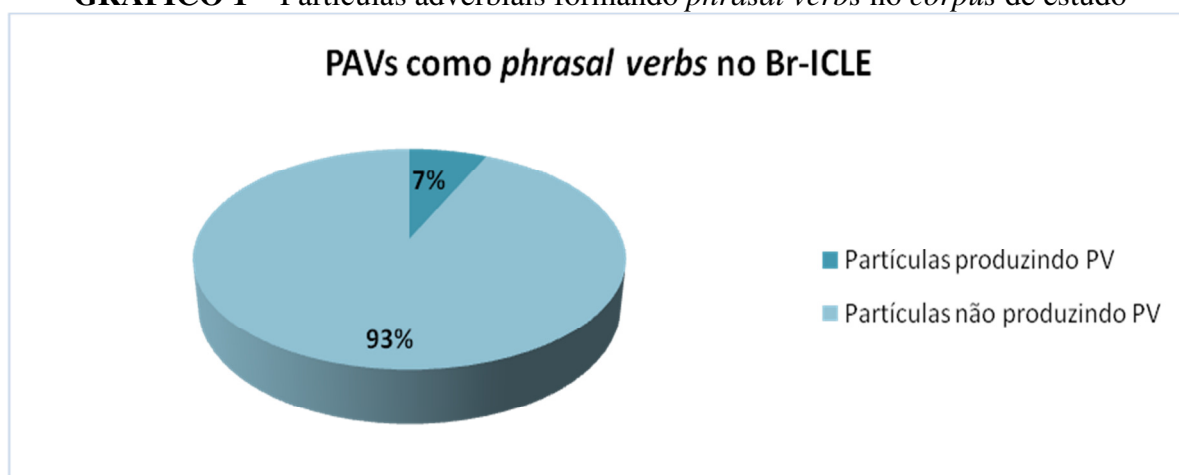
Fonte: Elaborado pela autora.

Os valores normalizados são próximos entre os dois *corpora* (CE = 4.113 / CR = 4.201) e indicam que as quantidades de diferentes partículas usadas entre os aprendizes

nativos e não nativos é similar. Ou seja, os valores após a normalização mostram que em cada 100 mil palavras, as partículas adverbiais ocorrem 4.113 vezes no Br-ICLE e 4.201 vezes no LOCNESS. Com exceção de “*aback*”, que não ocorre no CR, todas as demais partículas estão presentes também no CE e, das dez mais frequentes neste, nove também estão no CR, ainda que haja diferença entre elas no que toca à posição em que se encontram na lista de frequência (*without*, por exemplo, é a quinta partícula com maior número de ocorrências no Br-ICLE, mas a nona no LOCNESS. Da mesma forma, *out* é a oitava mais frequente no CE, mas a quinta no CR).

Considerando que o objetivo deste estudo é verificar o uso de verbos frasais na escrita acadêmica de estudantes brasileiros, analisamos as linhas de concordância de todas as partículas extraídas a fim de saber quantas delas produzem combinações com verbos lexicais produzindo PVs. O GRÁFICO 1 exibe o resultado dessa análise no CE.

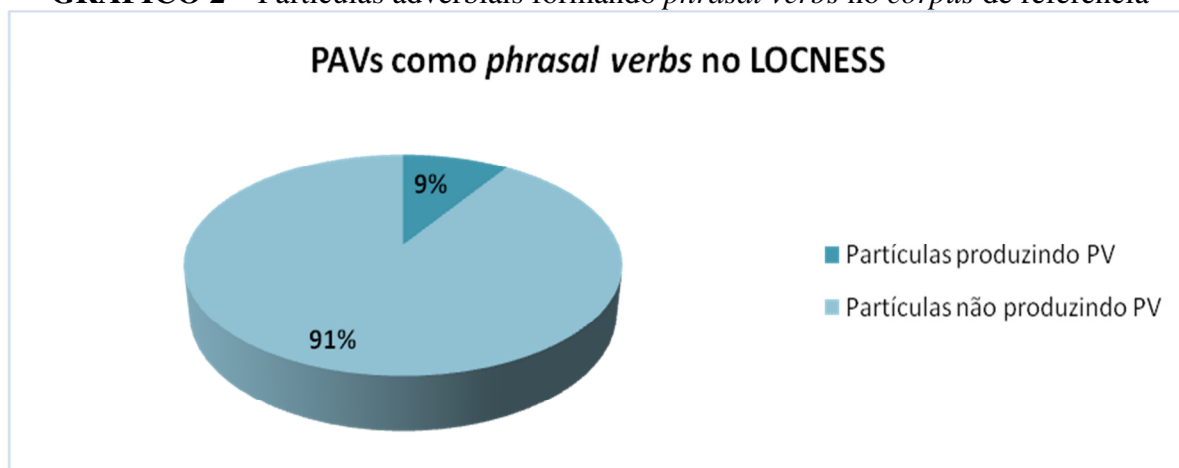
**GRÁFICO 1** – Partículas adverbiais formando *phrasal verbs* no *corpus* de estudo



Fonte: Elaborado pela autora.

Como é possível verificar, do número total de partículas encontradas no Br-ICLE, 7% forma verbos frasais. Em termos de frequência bruta, temos partículas produzindo 463 combinações, como *give up* e *go on*, no CE. Já no LOCNESS, são 1.288 combinações, o que representa 9% do total de PAVs encontradas no *corpus*, conforme mostra o GRÁFICO 2.

**GRÁFICO 2** – Partículas adverbiais formando *phrasal verbs* no *corpus* de referência



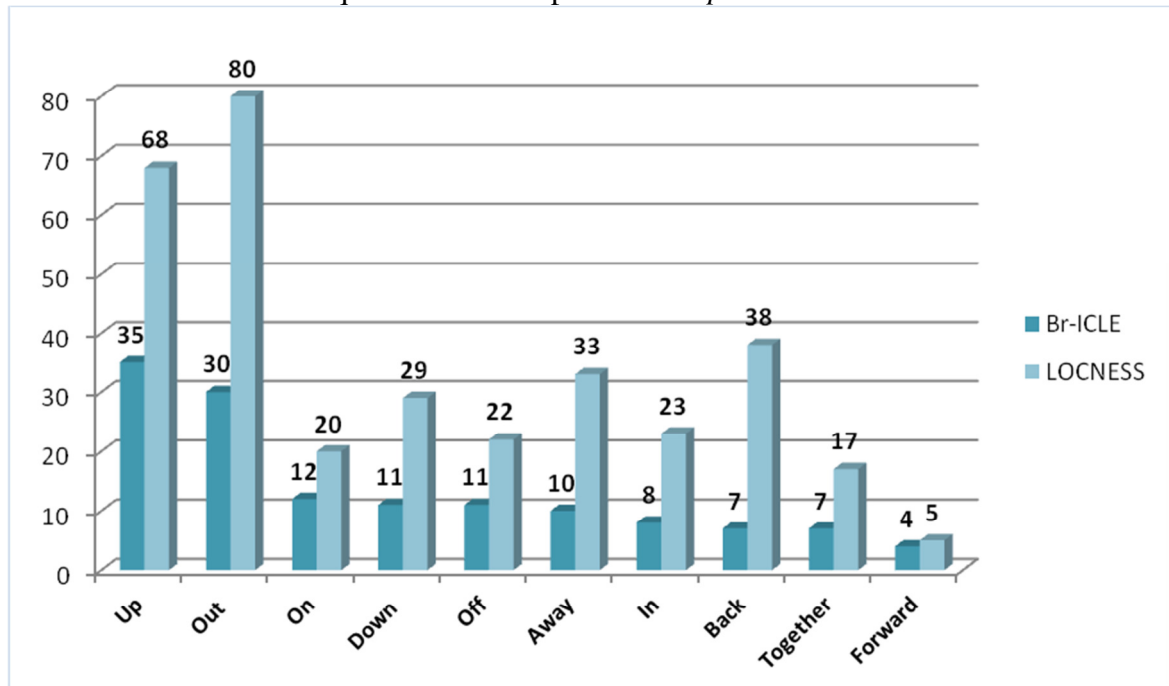
Fonte: Elaborado pela autora.

Esses dados indicam que, de maneira geral, os aprendizes nativos produzem mais combinações entre verbo + partícula adverbial do que os aprendizes não nativos, o que parece ser completamente natural, já esse tipo de estrutura gramatical é um traço peculiar da língua inglesa e os alunos nativos têm contato com esse tipo de construção desde cedo, vivenciando os *phrasal verbs* tanto em contextos de língua escrita, quanto de língua falada, o que não ocorre - salvo em casos fortuitos - com estudantes não nativos. Por outro lado, se analisarmos os valores da frequência normalizada, vemos que a diferença de produção entre os grupos não é tão proeminente. Enquanto os valores brutos apontam para uma diferença de 64%, os valores normalizados mostram que a diferença no uso de partículas formando PVs é de 27%, ou seja, no montante de 100 mil palavras, os aprendizes brasileiros produzem 105 combinações a menos do que os alunos nativos (cf. APÊNDICE D para a lista completa de frequência de PAVs formando PVs). Essa análise sugere que, embora os aprendizes nativos utilizem uma quantia maior de partículas compondo PVs, a produção dos não nativos também é bastante considerável, especialmente se ponderarmos que o emprego das partículas está, na maioria dos casos, vinculado a um terceiro significado, ou seja, os aprendizes brasileiros demonstram compreender que esse tipo de combinação produz sentidos diferentes daqueles contidos nas palavras individuais, mesmo que eles não saibam exatamente como usá-las (cf. tópico 4.5 e capítulo 5).

Ambos os aprendizes, nativos e não nativos, empregam praticamente os mesmos tipos de partículas em seus PVs. Alguns tipos, contudo, ocorreram apenas no *corpus* de estudo (*aback* e *round*), enquanto outro, somente no *corpus* de referência (*forth*). As partículas *past* e *without* não ocorreram em nenhum dos *corpora*. O número de verbos com os quais as

partículas selecionadas combinam também varia entre CE e CR. O GRÁFICO 3 mostra as dez partículas mais produtivas no Br-ICLE.

**GRÁFICO 3** - As dez partículas adverbiais mais produtivas no CE e o número de verbos com os quais combinam para formar *phrasal verbs*



Fonte: Elaborado pela autora.

Em uma análise baseada na frequência bruta e em combinação, respectivamente, com 33 e 30 verbos diferentes, *up* e *out* são as partículas mais produtivas no CE e, juntos, os verbos frasais formados a partir dessas PAVs representam 47% do total de PVs encontrados no *corpus*. As duas partículas também são as mais produtivas no CR, porém, diferentemente do que acontece no CE, neste, *out* ocorre em maior frequência, combinando com 80 verbos distintos, enquanto *up* combina com 68. A produtividade de *out* na combinação com verbos diversos no Br-ICLE é 63% menor do que no LOCNESS, sendo essa a partícula que apresenta a maior diferença bruta de combinações na comparação CE e CR (são 50 combinações a mais no CR do que as identificadas no CE). A PAV *up* também apresenta um percentual de diferença alto (49%) na comparação entre os *corpora*, assim como *back*, que aparece como a oitava partícula mais produtiva no CE (juntamente com *together*), mas a terceira no CR. A PAV *back*, aliás, é a que apresenta a maior diferença percentual entre os grupos analisados, combinando com 82% menos tipos de verbos lexicais no CE do que no CR, o que mostra sua alta produtividade entre os nativos.

A terceira partícula mais frequente no Br-ICLE é *on*, que produz 12 combinações com verbos lexicais. Embora tenha aparecido nos *corpora* formando mais *prepositional verbs* (o

que pode ser explicado pelo tipo de texto analisado e pelas expressões que são típicas desse gênero, como *base on*, *focus on*, *depends on* etc.), ela também ocorre formando muitos *phrasal verbs* e está entre as dez partículas mais produtivas do CR, no qual representa a oitava maior frequência. Assim como *up* e *out*, *on* também aparece na lista de Biber *et al.* (1999, p. 413), que reúne as PAVs mais comuns na formação de PVs. Em uma análise comparativa, a partícula *on* produz no CE oito combinações a menos do que no CR, o que pode ser considerado como uma variação pequena em relação às demais apresentadas no gráfico. Essa diferença só não é menor do que a estabelecida para *forward*, que aponta apenas uma combinação a menos entre um *corpus* e outro (ou seja, uma diferença de 20%). Isso faz de *forward*, inclusive, a partícula que mais apresenta semelhança de produtividade na relação LOCNESS e Br-ICLE.

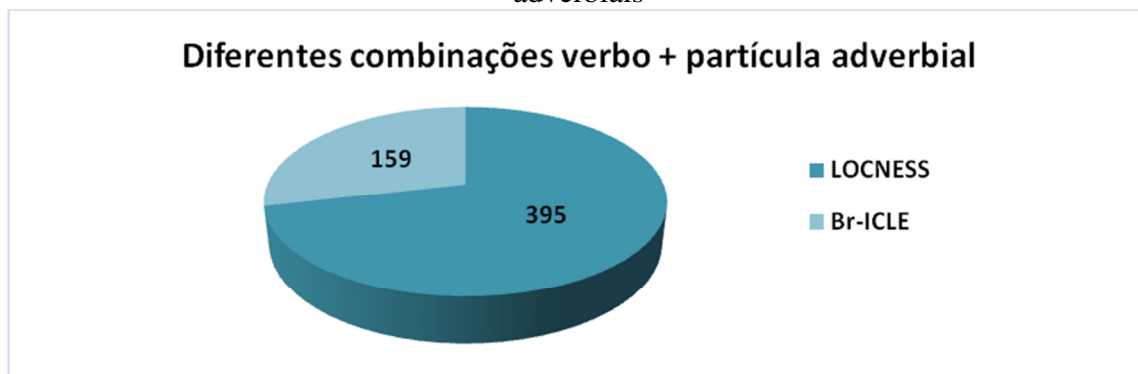
Embora apresente 23 combinações a mais no CR do que no CE, a partícula *away* no *corpus* de não nativos forma praticamente as mesmas combinações identificadas no *corpus* de nativos. Das dez combinações encontradas no CE, oito estão também no CR. Além disso, os dois verbos mais produtivos no Br-ICLE com essa partícula (*take* e *go*), são também os mais produtivos no LOCNESS, o que mostra a semelhança no uso entre um grupo e outro.

Ainda no que toca à produtividade com tipos distintos de verbos lexicais, as PAVs *down*, *off*, *in* e *together*, de maneira geral, apresentam frequências com poucas e/ou médias discrepâncias entre um *corpus* e outro. As diferenças entre as produções de cada grupo variam entre 10 combinações (para *together*) e 18 combinações (para *down*). É importante aqui ressaltar que os valores relativos à produtividade neste tópico são mostrados na forma bruta porque os resultados gerados a partir de probabilidade estatística poderiam parecer estranhos nesse caso, haja vista que não se pode afirmar, ao certo, que as mesmas combinações identificadas seriam reproduzidas também após a normalização. É indiscutível que o tamanho do *corpus* exerce bastante influência no número de frequência, mas a ampliação do *corpus* de estudo também poderia implicar em uma reação inesperada, como a identificação de combinações com verbos lexicais diferentes dos já apresentados nos *corpora*, por exemplo. Assim, a opção por exibir os dados em valores brutos nesse quesito visa a reduzir a possibilidade de julgamento errôneo em relação à produção dos grupos envolvidos.

Por fim, no que concerne à produtividade geral, conforme mostra o GRÁFICO 4, as partículas adverbiais selecionadas formaram 159 diferentes combinações de PVs no Br-ICLE. No LOCNESS, foram formadas 395 combinações distintas.



**GRÁFICO 4** - Número de verbos lexicais individuais formando PVs com partículas adverbiais



Fonte: Elaborado pela autora.

Em valores normalizados, podemos dizer que a cada 100 mil palavras foram produzidos 99 tipos diferentes de PVs no *corpus* de estudo, e 121 no *corpus* de referência, o que podemos considerar como uma diferença moderada entre os grupos. As condições de escrita dos textos (exposição à língua, tempo total de estudos do idioma, influência da língua materna etc.) podem ser destacadas como possíveis fatores de interferência na produtividade das partículas por parte dos aprendizes não nativos, afinal, o contato com a língua alvo propicia a compreensão das palavras em contextos diversificados, bem como permite o conhecimento da ampliação de significados atribuídos às combinações de verbos multipalavras, o que favorece o uso de novas construções em substituição aos PVs mais usuais ou ainda a verbos únicos, contribuindo, por exemplo, para a coesão textual. A lista contendo a distribuição da produtividade das partículas pode ser vista no APÊNDICE E.

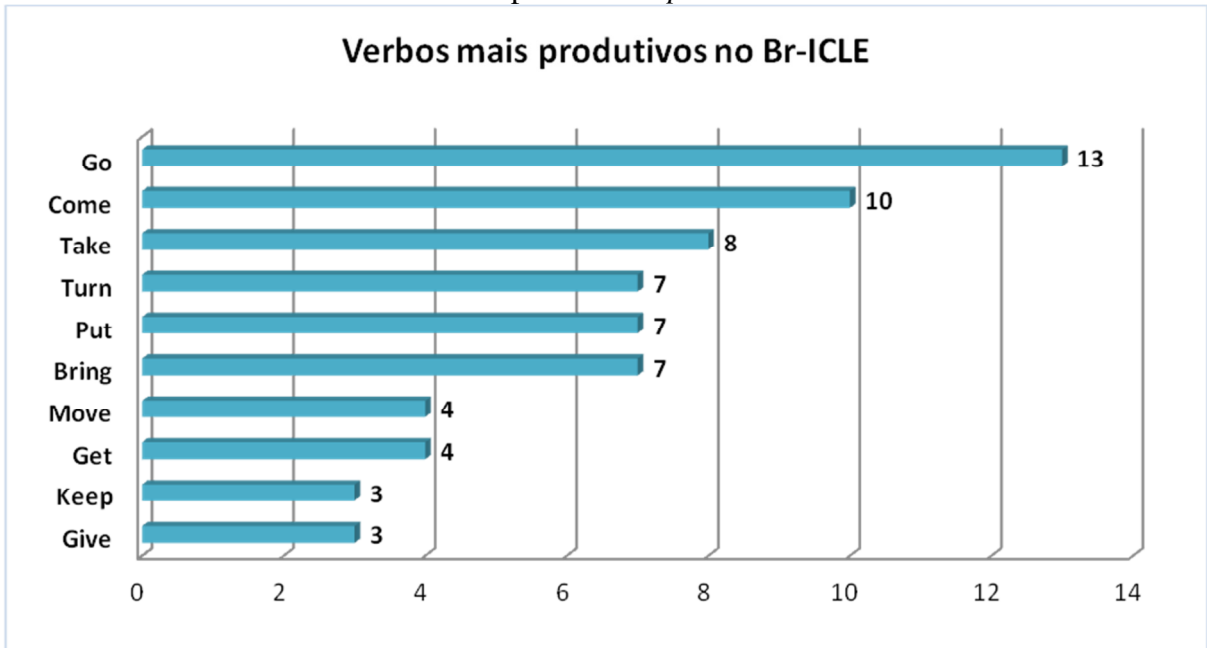
### 4.3 Frequência e produtividade dos verbos lexicais

De todos os verbos lexicais<sup>25</sup> identificados no Br-ICLE, 80 formaram *phrasal verbs*. Já no LOCNESS, foram 209 verbos. Apesar de a diferença parecer grande, os valores normalizados mostram que, a cada 100 mil palavras, 50 formam PVs no CE, e 64 no CR, o que significa uma margem de apenas 22% entre os dois grupos. No que concerne à produtividade desses verbos, ou seja, à capacidade de essas palavras se juntarem a outras para formar verbos frasais, alguns se destacaram mais que os outros. Enquanto verbos como *set* - que aparece na lista dos mais produtivos de Biber *et al* (1999) - formaram combinações com pelo menos sete partículas no *corpus* de referência, no *corpus* de estudo ligou-se a apenas

<sup>25</sup> Para essa contagem, especificamente, cabe pontuar, somente os verbos no infinitivo foram considerados. As formas flexionadas foram usadas nas análises de verbos frasais.

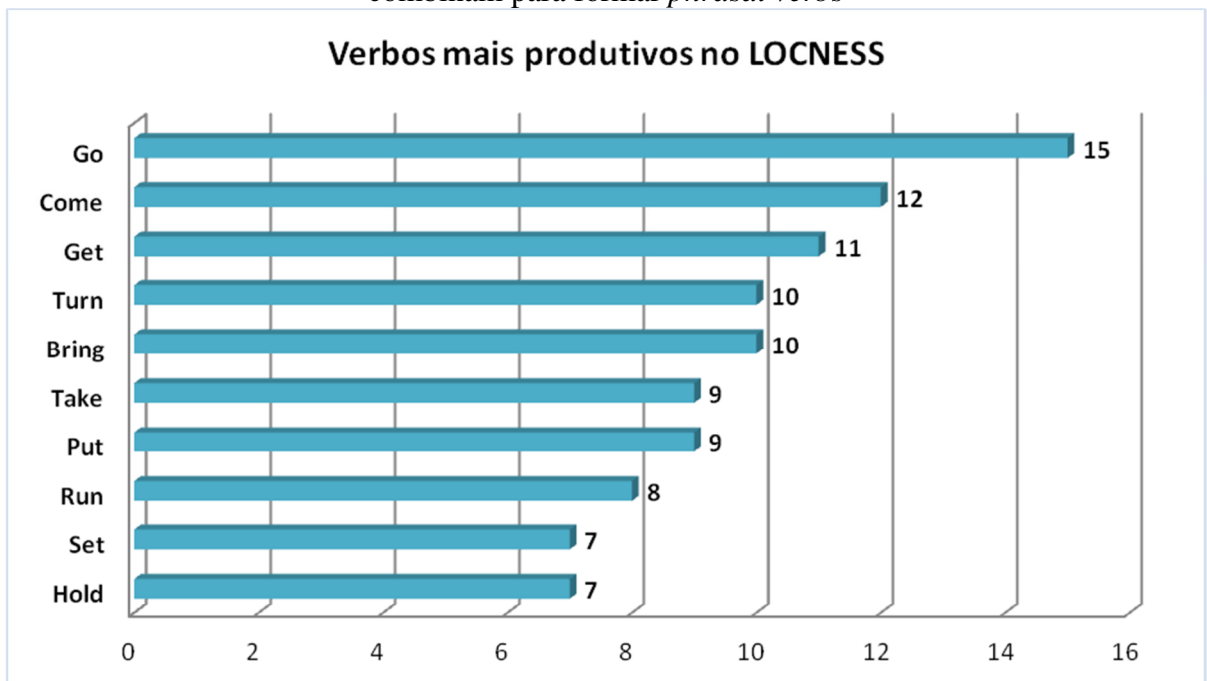
uma, o que pode ser considerado um indício de repertório limitado dos aprendizes não nativos. Os gráficos a seguir (GRÁFICOS 5 e 6) mostram os dez verbos mais produtivos em ambos os *corpora*.

**GRÁFICO 5** - Os dez verbos mais produtivos no CE e o número de partículas com as quais combinam para formar *phrasal verbs*



Fonte: Elaborado pela autora.

**GRÁFICO 6** - Os dez verbos mais produtivos no CR e o número de partículas com as quais combinam para formar *phrasal verbs*



Fonte: Elaborado pela autora.

Os resultados dos gráficos, que exibem valores brutos, mostram que sete dos verbos mais produtivos no *corpus* de estudo também são os mais produtivos no *corpus* de referência. Dos sugeridos por Biber *et. al* (1999, p. 413), todos constam no CR e apenas *set* não está no CE. *Go* e *Come* são os dois mais produtivos em ambos os grupos, com dez ou mais ocorrências em cada um deles. *Take* é o terceiro mais frequente no Br-ICLE e o sexto no LOCNESS. Apesar disso, a diferença entre eles é pequena, com apenas uma combinação discrepante. Dos verbos que não ocorrem em ambos, temos no CE *move* (que forma quatro combinações distintas), *keep* e *give* (que formam três combinações cada), e no CR *run* (formando oito combinações), *set* e *hold* (formando sete combinações cada).

Um verbo que chama a atenção é *get*, que apresenta sete combinações a mais no CR do que no CE. Uma explicação possível para a disparidade na produtividade pode estar no fato de esse ser um verbo, normalmente, de significado muito amplo ou geral, o que o torna de difícil aprendizagem e/ou compreensão para os brasileiros e não favorece a construção de novas expressões, dando lugar às alternativas de origem latina. Um exemplo é a formação “*get back*”, que ocorre no LOCNESS, mas não no Br-ICLE. Neste último, inclusive, o PV é substituído por “*return*” e também por “*go back*”, um verbo frasal de assimilação mais fácil para aprendizes não nativos, provavelmente, em função da opacidade reduzida.

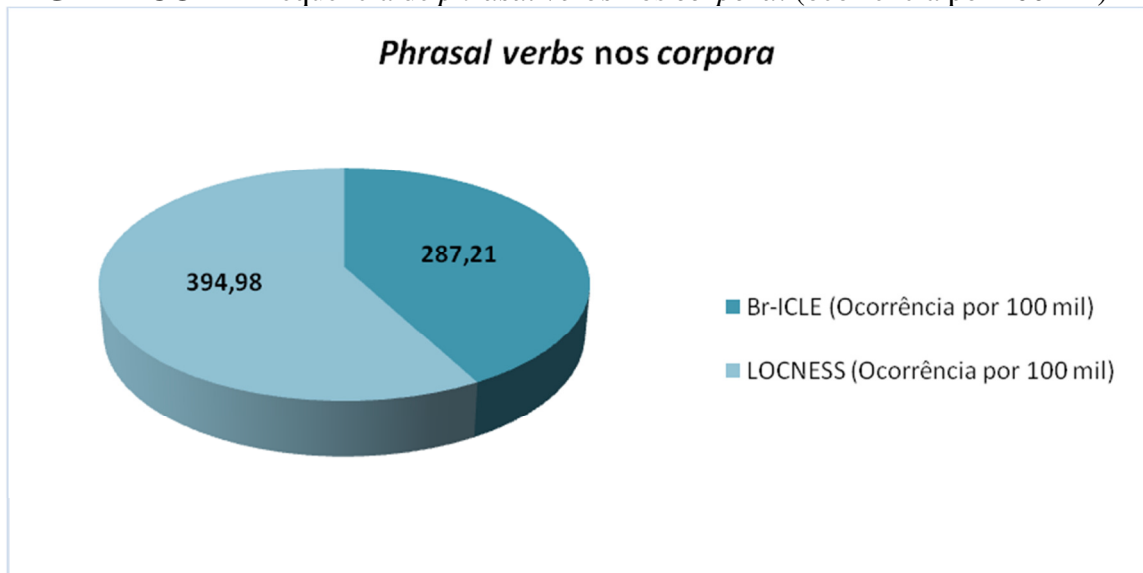
Por fim, de maneira geral, os dados apresentados sugerem que a produtividade dos verbos lexicais é semelhante nos dois grupos e demonstra que grande parte das palavras que constroem um maior número de combinações no CE, também são as mesmas no CR. As combinações formadas a partir dos verbos produtivos serão mostradas no tópico que se segue.

#### **4.4 Os *phrasal verbs* nos *corpora* - visão geral**

Conforme mostrado no tópico anterior, foram encontrados 159 tipos diferentes de verbos frasais no CE. Dentre eles, muitos tiveram mais de uma ocorrência no *corpus*, gerando um total bruto de 463 PVs produzidos nos textos acadêmicos dos brasileiros. Da mesma forma, no CR também houve ocorrências repetidas de PVs. Para os 395 tipos identificados, foram empregadas pelos nativos 1.288 combinações. Em valores normalizados, conforme mostra o GRÁFICO 7, isso significa que, para uma ocorrência de 100 mil palavras, há 287,21 verbos frasais no Br-ICLE, e 394,98 no LOCNESS, ou seja, uma diferença de apenas 27% entre os grupos. Na comparação entre a produção dos aprendizes, esses dados sugerem que, ainda que o emprego dessas estruturas gramaticais nos textos de nativos seja maior, os não

nativos também produzem um número notável de PVs nas composições textuais de caráter mais formal.

**GRÁFICO 7** – Frequência de *phrasal verbs* nos *corpora*. (ocorrência por 100 mil)



Fonte: Elaborado pela autora.

Para determinar se a diferença encontrada na ocorrência de *phrasal verbs* nos *corpora* é estatisticamente significativa aplicamos o teste *Log-likelihood*. Consideramos na análise o valor de  $LL > 3,84$  para um  $p < 0,05$ , ou seja, para que o dado fosse tido como significante, o resultado alcançado no teste deveria ser maior que 3,84, o que geraria um grau de 95% de confiança de que esse saldo não ocorreu por questões aleatórias, mas sim por motivações específicas. A TAB. 6 mostra o resultado do teste e, nele, verificamos que o valor de LL é maior do que o estabelecido como valor crítico. Logo, a diferença de frequência dos PVs nos grupos analisados é sim considerada um fenômeno relevante estatisticamente.

**TABELA 6** - Teste *log-likelihood*

<i>Corpus</i>	# PVs encontrados (valor bruto)	Valor de LL	Probabilidade (p)
<b>Br-ICLE</b>	463	36,24	<0.05
<b>LOCNESS</b>	1288		

Fonte: Elaborado pela autora.

De uma perspectiva mais geral, por outro lado, os números relativos à frequência dos PVs apontam para uma baixa produtividade de verbos frasais na escrita acadêmica, o que também foi relatado por Biber *et. al* (1999) e por outros pesquisadores que avaliaram esse tipo

de construção gramatical no gênero argumentativo. Apesar disso, a amostra gerada na análise dos *corpora* é representativa para fins desta pesquisa.

#### 4.5 Frequência e produtividade dos *phrasal verbs* na escrita acadêmica

Depois de identificados, todos os verbos frasais dos *corpora* foram organizados em planilhas e dispostos conforme o número de ocorrências. Como a quantidade encontrada foi muito grande (cf. APÊNDICE F), optamos por facilitar a visualização dos dados mostrando no QUADRO 8 apenas os 25 PVs mais frequentes no CE e CR. Os verbos sombreados em cinza escuro são os dez com maior ocorrência no CE e a posição em que aparecem no CR.

**QUADRO 8 - Os 25 verbos frasais mais frequentes nos *corpora***

	Br-ICLE		LOCNESS	
	<i>Phrasal verb</i>	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência por 100 mil
1	<i>end up</i>	14	<i>go on</i>	16
2	<i>give up</i>	14	<i>carry out</i>	15
3	<i>go on</i>	14	<i>point out</i>	13
4	<i>sum up</i>	12	<i>take away</i>	9
5	<i>point out</i>	9	<i>bring up</i>	9
6	<i>come back</i>	9	<i>bring about</i>	8
7	<i>grow up</i>	9	<i>take on</i>	8
8	<i>find out</i>	8	<i>end up</i>	8
9	<i>keep on</i>	8	<i>grow up</i>	8
10	<i>put forward</i>	8	<i>give up</i>	8
11	<i>take away</i>	6	<i>find out</i>	6
12	<i>call out</i>	6	<i>make up</i>	6
13	<i>go out</i>	6	<i>set up</i>	5
14	<i>go by</i>	6	<i>go back</i>	5
15	<i>turn on</i>	5	<i>break down</i>	4
16	<i>get together</i>	5	<i>cut off</i>	4
17	<i>leave behind</i>	5	<i>come about</i>	3
18	<i>come up</i>	4	<i>bring in</i>	3
19	<i>wake up</i>	4	<i>run up</i>	3
20	<i>put together</i>	3	<i>carry on</i>	3
21	<i>stand up</i>	3	<i>go out</i>	3
22	<i>get out</i>	2	<i>turn out</i>	3
23	<i>go back</i>	2	<i>bring back</i>	3
24	<i>turn off</i>	2	<i>fit in</i>	3
25	<i>go through</i>	2	<i>get out</i>	3

Fonte: Elaborado pela autora.

Com 14 ocorrências por 100 mil palavras, o verbo frasal *end up* é o mais frequente no *corpus* de estudo. Em todos os textos analisados, a combinação expressa sentido de se tornar ou terminar em uma situação ou lugar inesperado (*Longman Dictionary of Contemporary English*, 2003), como nos exemplos a seguir:

- a) “*More than often these people **end up** suffering or even dying in these foreign places because they have to live in very poor and violent areas*” (Br-ICLE).
- b) “*“On television, the bad guys always **end up** behind bars, children are never neglected by their parents, and no one is never left out.”*” (Br-ICLE).
- c) “*I mean, if criminals **ended up** performing their crimes I believe that they were pushed into it in most cases and it usually has to do with lack of opportunities in life.*” (Br-ICLE).

Nesse contexto, ele poderia ser substituído pelo verbo único<sup>26</sup> “*finish*” (com sentido de “*come to an end*”) ou pelos *phrasal verbs* “*finish up*” ou “*wind up*” que, aliás, não ocorrem no CE. As ocorrências de *finish* (*finish, finishes, finishing* e *finished*), por sua vez, somam 24 (valor bruto) no CE, entretanto, seu uso está relacionado ao significado de “completar algo”, como em “*It needs years of work to finally achieve what it supposed to happen when the studies are **finished***” (Br-ICLE), e nunca com o mesmo sentido de *end up*, o que sugere que os aprendizes brasileiros utilizaram o verbo frasal conscientes de seu sentido e que, nesse caso, preferiram o uso do MWV em detrimento do verbo individual.

*End up* também está os entre os 25 mais frequentes no *corpus* de referência, porém, em menor quantidade (47% menos ocorrências do que no CE, o que o coloca na oitava posição do ranque). A análise das concordâncias mostra que a combinação é empregada pelos nativos com o mesmo significado usado pelos não nativos, como nos exemplos (d) e (e), e reforça a afirmação de que, para esse verbo frasal, há similaridade de uso entre os aprendizes dos dois grupos.

- d) “[...] *therefore their prejudice may not actually **end up** in discrimination against minorities*” (LOCNESS).

---

<sup>26</sup> Por “verbo único” entende-se aqui o verbo formado por uma única palavra (*one-word verb*), em oposição a *multi-word verb* (verbo multipalavra), que é o caso dos *phrasal verbs*.

e) “*Michael Watson ended up in a wheelchair, paralysed and with brain damage.*” (LOCNESS).

*Give up* é a segunda combinação mais frequente no Br-ICLE e tem apenas uma ocorrência bruta a menos que *end up* nesse grupo (são 23 ocorrências brutas para *end up* e 22 para *give up*). O verbo frasal também ocorre no *corpus* de referência, porém, como pouco mais da metade das frequências identificadas no CE. Mesmo assim, em ambos os *corpora* ele é empregado semelhantemente com os valores de “abrir mão de algo” (f), “abandonar um hábito” (g) e “desistir” (h). Na comparação com “*abandon*”, foi possível verificar que a equivalente latina foi menos usada por ambos os grupos, ocorrendo quatro vezes a cada 100 mil palavras nos dois *corpora*. Aqui, novamente, vê-se que o MWV foi preferido em relação ao verbo único.

f) “*The green house effect is a major issue to most citizens but very few of them are prepared to give up their car to show it.*” (LOCNESS).

g) “*If a lot of people did give up beef products, this will have had effects on the agricultural industry.*” (LOCNESS).

h) “*According the academic standards it was a crap but she didn't gave up.*” (Br-ICLE).

Apesar de ser menos comum na prosa acadêmica do que na produção oral ou no gênero de ficção (BIBER *et al.*, 1999), a combinação *go on* está no topo de ocorrências no CR e é a terceira mais frequente no CE. Seus significados nas linhas de concordância dos *corpora* são variados e o mais comum dentre eles (que ocorre em ambos) está relacionado à ação de dar continuidade a algo (Longman Dictionary of Contemporary English, 2003), como em (i), (j) e (k), em que o PV poderia ser substituído pelo verbo “*continue*” ou pelo *phrasal verb* “*carry on*” que, aliás, não aparece no *corpus* de estudo, mas ocupa a posição 20 no ranque do *corpus* de referência.

i) “*The real thing is the relationship that may go on with or without marriage, the real problem is the coexistence.*” (Br-ICLE).

j) “*Sartre was particularly against letting oneself be ruled by the past: what's done is done, life must go on.*” (LOCNESS).

k) “*In the first text and the movie, what made the characters keeping **going on** and not desist, was the desire of learning and improve [...]*” (Br-ICLE).

Outro significado em comum nos *corpora* atribuído a *go on* é o de “acontecer”, cujo verbo único equivalente é “*happen*”, e que aparece predominantemente na forma progressiva (*going on*) em ambos os grupos, como mostram os exemplo (l) e (m):

l) “*It says that parents and teachers should pay attention on what is **going on** with TV programs, movies and publications that enthusiasm the young [...]*” (Br-ICLE).

m) “*There is wide debate **going on** about how much power the European Community should have.*” (LOCNESS).

As concordâncias (n) e (o) mostram o terceiro significado mais usual para *go on* nos *corpora*, que é relativo a tempo (passar o tempo) e tem como verbo equivalente o *pass* (*Longman Dictionary of Contemporary English*, 2003):

(n) “*As time **goes on** vicissitudes of life make people forget commitments and promises made by their representatives once running for a leadership career.*” (Br-ICLE).

(o) “*As time **goes on** certain things fade away and other things become more important to us.*” (LOCNESS).

Por fim, o quarto sentido identificado para o PV *go on* nos textos dos aprendizes é “durar por um período de tempo” (*Longman Dictionary of Contemporary English*, 2003), que poderia ser substituído pelo correspondente “*last*”, como nos exemplos a seguir:

(p) “*They still happen and they will probably **go on** for a long time.*” (Br-ICLE).

(q) “*This **went on** for about a year.*” (LOCNESS).

A partir desses exemplos podemos perceber que, além de ser bastante usado também pelos aprendizes não nativos, a combinação *go on* é usada com variedade de significados pelos dois grupos. No caso dos estudantes brasileiros, isso sugere que há compreensão, a princípio, da diferença na função desse PV e que os não nativos conseguem aplicá-lo em pelo menos quatro contextos diferentes, assim como os nativos. Além disso, a alta frequência de um verbo frasal tão coloquial como esse indica que os textos de ambos os grupos foram



produzidos em estilo pouco formal. O contraste com palavras equivalentes (QUADRO 9) sugere que os não nativos têm preferência pelo uso do verbo único para três dos significados, mas não para “*pass*” que, apesar da origem latina e da semelhança com o “passar” - tradução direta no português -, não foi priorizado pelos brasileiros, mostrando que a influência da primeira língua não foi fator preponderante na escolha. De forma semelhante, os nativos só preferem o uso do verbo único quando o sentido de *go on* é o de duração (*last*).

**QUADRO 9** - Comparação entre “*go on*” e seus verbos únicos equivalentes

Verbos correspondentes para <i>go on</i>			
Verbo em inglês	Origem	Frequência CE (100 mil)	Frequência CR (100 mil)
<i>Go on</i>	Inglês antigo (ano 1777)	14	16
<i>Continue</i>	Latim ( <i>cotinuare</i> )	36	57
<i>Happen</i>	Inglês médio tardio ( <i>hap</i> )	112	47
<i>Last</i>	Inglês antigo ( <i>lǣstan</i> )	24	8
<i>Pass</i>	Latim ( <i>passare</i> )	7	34

Fonte: Elaborado pela autora. Dados etimológicos: HARPER, 2016.

Outro ponto interessante a ser destacado na comparação de *go on* e verbos correlatos é a preferência dos aprendizes pelo equivalente “*happen*”. Tanto nativos quanto não nativos produzem mais essa alternativa do que o MWV, indicando uma inclinação pelo uso da palavra mais formal em detrimento da mais coloquial. Além disso, é notória a diferença de 58% entre a produção dos dois grupos. O fato do CE produzir 65 ocorrências a mais que o CR pode ser atribuído, por exemplo, à ausência de outros PVs similares como alternativas nos textos dos brasileiros, o que faz com que haja um aumento ou concentração no uso de determinada palavra. No *corpus* dos nativos há mais opções. Além de *happen* e *go on*, ocorre também uma frequência considerável de *come about* e suas variantes, que também podem ser usadas como palavras substitutas.

O quarto verbo frasal mais frequente no *corpus* de estudo, *sum up*, chama a atenção por estar em subuso no *corpus* de referência, não compondo a lista dos 25 mais frequentes. Ele ocorre 12 vezes a cada 100 mil palavras no CE, é apresentado em um padrão predominantemente precedido de preposição (*to* + infinitivo) e sua alta frequência pode ser explicada por dois fatores principais: a influência da língua nativa, devido à semelhança com “*suma*” no português (do latim, *summa*.), e por ser um verbo muito usado na escrita acadêmica em razão de sua função retórica. Outra explicação também possível pode ser o fato

de esse verbo frasal compor parte de um enunciado, o que, considerando que os aprendizes tendem, muitas vezes, a repetir parte dos enunciados em seus textos, ocasionaria a frequência elevada do PV. A baixa frequência no CR, por sua vez, sugere que os nativos utilizem outros verbos e expressões para introduzirem uma conclusão. O QUADRO 10 mostra uma comparação entre *sum up* e expressões com funções similares.

**QUADRO 10** - Comparação entre “*sum up*” e termos equivalentes

<b>Expressões</b>	<b>Freq. Br-ICLE (100 mil)</b>	<b>Freq. LOCNESS (100 mil)</b>
<i>To sum up</i>	12,4	0,31
<i>Summarizing</i>	1,2	0,31
<i>In sum</i>	1,2	-
<i>Summing up</i>	-	0,61
<i>In summary</i>	1,2	-
<i>To summarize</i>	-	-
<i>To conclude</i>	12,4	3,37
<i>In conclusion</i>	17,9	5,83
<i>Concluding</i>	3,1	0,31

Fonte: Elaborado pela autora.

Na comparação, sombreada de azul, vemos que a forma mais usada por ambos os grupos é “*in conclusion*”, o que sugere a preferência dos aprendizes nativos e não nativos por um elemento latinizado. Mesmo CE, que apresenta alta frequência de *sum up*, mostra grande produção de “*in conclusion*”, com uma diferença de 31% entre o PV e o verbo único. A variante “*to conclude*” também tem frequência significativa no CE e é a segunda maior frequência no CR. Reunindo todas essas informações, podemos inferir que a opção pelo uso das formas de verbos únicos e latinizados manifesta a predileção pelos termos de caráter mais formal na escrita acadêmica, o que ocorre especialmente no grupo de aprendizes brasileiros. Contudo, não é possível depreender a partir desses dados se a escolha pelos termos menos coloquiais é realizada de forma consciente pelos aprendizes ou se, no caso dos brasileiros, o comportamento reflete uma influência da língua materna. A diferença entre os valores do CE e do CR para essas alternativas pode indicar também que os não nativos tendem a seguir uma estrutura bem marcada em seus textos, posicionando claramente a conclusão das ideias, o que

não parece ocorrer no grupo dos nativos. Por outro lado, aqui também a língua nativa pode estar exercendo influência, já que a estrutura formal também é bastante marcada no português.

Assim como *sum up*, outro verbo frasal comumente empregado em textos acadêmicos (cf. Biber *et al*, 1999, p. 410) e que também aparece na lista dos 25 mais frequentes do CE e do CR, é *point out*. Ranqueado entre os dez de maior ocorrência (nove vezes a cada 100 mil palavras no Br-ICLE e 13 no LOCNESS), esse PV é utilizado por ambos os grupos com o significado de “indicar”, “mostrar” algo (*Longman Dictionary of Contemporary English*, 2003). Sua frequência, relativamente alta nesse tipo de texto, está relacionada ao fato de ele ser considerado um *phrasal verb* mais formal (OPVD, s. 16-17), o que torna seu uso favorável para a escrita acadêmica.

O sexto verbo frasal de maior ocorrência no CE é *come back*, cuja frequência normalizada é de nove vezes a cada 100 mil palavras. A composição, que é usada expressando valor de “voltar” ou “retornar” (*Longman Dictionary of Contemporary English*, 2003), não ocorre entre os 25 mais comuns do LOCNESS. Nesse grupo, a alternativa empregada é *go back*, o 14º PV mais frequente no CR. *Grow up* e *find out* também ocorrem em ambos os corpora. *Grow up* apresenta uso similar nos dois grupos e seus significados estão ligados à ação de se tornar adulto, como nos exemplos (r) e (s), ou de se desenvolver gradualmente (t, u). Sua alta frequência está relacionada, provavelmente, às temáticas propostas para os textos, como a adolescência, a censura na sociedade, os valores familiares, entre outros.

(r) “[...] *a generation which **grows up** without having listened to fairy tales, without having created their own games with their friends.*” (Br-ICLE).

(s) “*As children **grow up**, they learn morals from their religious community.*” (LOCNESS).

(t) “*In that time civilization was **growing up** and we have that process of industrialization, there is the technology, the inventions [...].*” (Br-ICLE).

(u) “*Candide also grows in this novel by learning of other people and cultures. He starts out **growing up** sheltered in the castle of Boron thunder-ten-trunckh, until he was booted from it after his romantic meeting with Cunegonde.*” (LOCNESS).

No que diz respeito ao significado atribuído, *find out* também é usado de maneira semelhante entre os dois grupos. Os padrões formados, entretanto, não são muito similares, já que, enquanto no grupo dos nativos o *phrasal verb* ocorre normalmente aliado a palavras típicas de perguntas (*how, what, who* etc.), à conjunção *if* ou a *that* (introduzindo oração

subordinada), no grupo dos não nativos não há regularidade de uso, sendo o PV empregado ora seguido por pronomes pessoais, ora por substantivos ou por *that* etc. O principal sentido atribuído pelos aprendizes em geral é o de “encontrar ou apreender uma informação”, como em (v) e (w).

(v) “*At this age, teenagers are trying to understand what's going on the world, they want to take point in it and they need socialize with other people their age, they want to **find out** where they belong to.*” (Br-ICLE).

(w) “*As more genes are `mapped' we **find out** more about the way our life is defined.*” (LOCNESS).

*Keep on* é outra combinação que consta no ranque das dez mais recorrentes no CE, mas que não ocorre entre as 25 mais frequentes do CR. O motivo, hipoteticamente, é que os aprendizes nativos utilizem outras combinações ou verbos únicos como alternativa a esse PV, o que pode ser reforçado pela ocorrência significativa de *continue* (cf. QUADRO 9) e *carry on*, por exemplo. O PV *put forward*, apesar de não aparecer no quadro do CR, é 28º mais frequente entre os nativos. Em ambos os *corpora*, é utilizado com a função de “sugerir uma ideia” e sua frequência relativamente alta no CE (é o décimo de maior ocorrência) pode ser explicada pelo fato de ser considerado um verbo frasal mais formal e comumente empregado em ensaios e relatórios (OPVD, s. 16-17).

Um fato interessante a ser observado é a ausência de verbos frasais formados com *bring* e *carry* entre os 25 mais frequentes do *corpus* de estudo. Na comparação com o CR, vemos que os dois verbos formam seis combinações diferentes dentre as mais frequentes produzidas por nativos, representando, juntos, 24% dos PVs da lista (cf. QUADRO 8). Formações como *carry out*, *inclusive*, são consideradas menos coloquiais e mais comuns na escrita acadêmica (BIBER *et al.*, 1999, p. 410), e o fato de não aparecerem formando PVs no CE sugere que os aprendizes brasileiros não têm muita familiaridade com esse verbo e com as possibilidades que apresenta na construção de verbos multipalavras.

Com base nos dez verbos mais frequentes no *corpus* de estudo, aplicamos o teste *Log-likelihood* para verificar a significância dos resultados e as condições de sobreuso e subuso dos PVs identificados. O QUADRO 11 mostra o resultado.

**QUADRO 11** – Sobreuso e subuso dos verbos frasais no Br-ICLE

<i>Phrasal verb</i>	Br-ICLE Frequência bruta	Br-ICLE Frequência por 100 mil	LOCNESS Frequência bruta	LOCNESS Frequência por 100 mil	LL	Sobreuso/ subuso
<i>end up</i>	23	14	26	8	4,03	+
<i>give up</i>	22	14	25	8	3,79	+
<i>go on</i>	22	14	53	16	0,49	-
<i>sum up</i>	20	12	1	0	37,01	+
<i>point out</i>	14	9	41	13	1,51	-
<i>come back</i>	14	9	7	2	9,86	+
<i>grow up</i>	14	9	26	8	0,07	+
<i>find out</i>	13	8	19	6	0,80	+
<i>keep on</i>	13	8	0	0	28,76	+
<i>put forward</i>	13	8	9	3	6,22	+

Fonte: Elaborado pela autora.

Conforme mostram os dados, oito dos PVs classificados dentre os dez de maior ocorrência no CE apresentam sobreuso em relação ao grupo de nativos, dentre eles *sum up* e *keep on*, que são os que apresentam as maiores diferenças entre os *corpora*. Os valores de log acima de  $LL < 3,84$  indicam que as diferenças encontradas não são aleatórias, mas sim linguisticamente significativas.

Finalmente, a análise quantitativa dos dados mostrou que os verbos frasais são, de maneira geral, pouco usados na escrita acadêmica tanto de nativos, quanto de não nativos. Os resultados de COSTA (2012) apontaram para o uso significativo de *multi-word verbs* nos *corpora*, o que de fato ficou atestado, porém, a verificação das linhas de concordância nos permitiu detectar que, dos três tipos de verbos formados por mais de uma palavra (cf. tópico 2.2), os *phrasal verbs* são os menos empregados nesse tipo de texto e os *prepositional*<sup>27</sup> são os de maior ocorrência.

Os resultados também apontaram para semelhanças e diferenças no emprego das combinações verbo + partícula na escrita acadêmica de nativos e não nativos do inglês. A primeira semelhança está relacionada à frequência dos verbos frasais, que é similar em quantidade para ambos os grupos. Os verbos frasais de maior ocorrência no *corpus* de estudo estão também, em sua maioria, no *corpus* de referência, o que também demonstra similaridade entre os grupos. Outra similaridade é que, em algumas situações, verbos de uma

<sup>27</sup> A afirmação de que os *prepositional verbs* são os que mais ocorrem nos *corpora* é baseada em minha observação das linhas de concordância. Entretanto, para que obtenhamos dados precisos e a comprovação efetiva desse achado, é necessário que seja feita uma investigação específica com o grupo dos verbos preposicionados.

palavra só ou construções com verbos latinizados foram priorizados para ambos os *corpora*, provavelmente em função do tipo de texto produzido e da formalidade das palavras. Isso sugere que, no caso de verbos frasais, não necessariamente os aprendizes brasileiros tenham predileção ou tendência por palavras de origem latina. Por outro lado, algumas ocorrências (como o caso dos verbos *carry* e *bring*) indicaram que pode haver uma tendência dos não nativos em evitar certos verbos frasais, apesar de ser necessário um estudo semântico mais aprofundado para descobrir o motivo desse evento. Em relação às diferenças no uso dos verbos frasais, vimos que são relativas, especialmente, ao significado que é atribuído a alguns PVs por parte dos não nativos, bem como à preferência do emprego de palavras de verbo único em detrimento de verbos multipalavras, como ocorreu com *happen* (em oposição a *go on*, que foi o escolhido pelos nativos).

No próximo tópico, trataremos da análise qualitativa dos dados, a fim de discutir alguns dos problemas encontrados e ponderar sobre as possíveis causas desses achados.

## 5 ANÁLISE E DISCUSSÃO DOS RESULTADOS: ABORDAGEM QUALITATIVA

Neste capítulo, serão apresentados os resultados encontrados a partir da análise qualitativa dos dados da pesquisa. Apesar de algo referente à qualidade já ter sido apresentado no capítulo 4, retomaremos neste tópico alguns dos pontos discutidos, a fim de investigar alguns aspectos, como a naturalidade, a criatividade e o desempenho dos aprendizes ao empregarem verbos frasais em textos acadêmicos.

### 5.1 Influências da primeira língua na aprendizagem do inglês

Como verificamos em exemplos apresentados na etapa de análise quantitativa, uma das diferenças entre o uso dos verbos frasais por aprendizes nativos e não nativos está relacionada à interferência/transferência da primeira língua na aprendizagem da segunda língua. O exemplo (1) mostra um trecho encontrado no CE que ilustra como essa influência da língua materna pode impactar na produção e no desempenho dos aprendizes:

- (1) *“On the whole, thinking of equal opportunities, authorization for contraceptive methods, women protection laws, maternity leave and abortion I strongly believe that feminism movement has dramatically **moved up** the society structure, mainly women’s life, and has done more good to the cause of women than harm, specially for all those that born in the past fifty years.” (Br-ICLE).*
- (2) *“No-one challenged the assumptions which ran throughout the lecture. Teaching did provide an opportunity for young people to **move up** the social scale.” (BNC/CLY).*

Vemos em (1) que o aprendiz utiliza o *phrasal verb* “*move up*” para falar de mudança no sentido de tornar algo diferente, transformar. Entretanto, esse não corresponde ao real significado do PV, que expressa, na realidade, uma mudança no sentido de movimento, como de um nível mais baixo para outro mais alto (2), de uma velocidade mais baixa para outra mais alta. Houve, portanto, nesse caso, uma interferência do português, gerando uma confusão com a tradução direta do verbo “mudar”. Em outras palavras, a análise contextual nos mostra que, na tentativa de explicar que o feminismo “mudou” drasticamente o mundo, o aprendiz utilizou o verbo “*move*” (mudança = movimento) inconsciente de que o verbo adequado era

“*change*” (mudança = alteração). Além disso, a análise das linhas de concordância do BNC mostrou que a palavra “*dramatically*”, usada pelo aprendiz antes do PV é um colocado de “*change*”, mas não de “*move*”, o que reforça a questão da inadequação por mostrar um desvio em relação ao padrão usado pelos nativos e, conseqüentemente, uma produção pouco natural para os falantes que têm o inglês como língua materna.

Outro exemplo de interferência de L1 é mostrado no exemplo (3):

(3) “*This complexity in society made the inequality <sic> **grows up**. Then, not only one man could be a leader, as in those basic society from the earliest time*”. (BRICLE).

(4) “*He comments: The adopted child can count upon help and friendship from a larger circle than the ordinary child. They **grow up** to respect both sets of kinfolk. The two families, usually related, are drawn closer together.*” (BNC/ALM).

O PV *grow up*, que tem valor principal de “crescer” no sentido de desenvolvimento ou amadurecimento de um estágio infantil para um estágio adulto, foi usado com função de “aumentar” para falar de crescimento da desigualdade. Nesse caso, vemos que o sentido atribuído reflete a tradução direta do *phrasal verb* para o português (crescer). Uma vez que, nesse contexto, *grow up* tem valor de “ficar maior (em quantidade)” e não de “se tornar adulto”, como é comumente empregado pelos nativos (4), seria mais adequada a sua substituição pelo verbo único “*grow*” ou ainda por “*increase*” que, aliás, é colocado<sup>28</sup> de “*inequality*”, ou seja, as palavras “desigualdade” e “aumento” aparecem frequentemente juntas em contextos similares aos apresentados em (3). Da mesma forma, os colocados de *grow up* (*child, children, kids, young* etc.) no BNC indicam inclinação para o uso do PV com função de crescer = amadurecer. Mesmo as outras possibilidades de uso de *grow up* (ex. desenvolver gradualmente como uma planta), não serviriam para tornar o emprego do PV totalmente natural nesse contexto, de forma que é clara a inadequação causada pela interferência de L1 na produção em L2, já que o aprendiz parece fazer uso das expressões sem ter a consciência do que é certo ou errado em relação aos padrões da língua inglesa.

---

<sup>28</sup> Na concepção de Berber Sardinha (2004), colocados são “associações entre itens lexicais, ou entre o léxico e campos semânticos”.



Haja vista a diversidade de significados atribuídos a um mesmo *phrasal verb*, faz-se necessário um estudo semanticamente mais aprofundado, com análise individual das linhas de concordância, para que seja possível um resultado conclusivo.

## 5.2 Criatividade na produção de verbos frasais

Outro fator importante da produção textual de aprendizes não nativos da língua inglesa diz respeito à criatividade no emprego de verbos frasais, ou seja, à criação de combinações inexistentes ou sem função fixada na língua alvo, mas que seguem um padrão ou regra gramatical similar às que estão disponíveis no sistema linguístico de L2 e que são geradas como forma de suprir uma deficiência ou mesmo a falta de uma expressão existente em L1, mas não em L2.

A criação de novos *phrasal verbs* geralmente é realizada por meio de analogias e o valor aspectual das partículas, ou seja, a maneira como elas podem expressar valores tais quais completude, intensidade, duração, reciprocidade etc., é um grande influenciador. No inglês, a partícula *on* em *carry on*, *hang on*, *go on* e *stay on*, por exemplo, marca continuidade. Ao tê-la como referência nesses verbos, o aprendiz pode criar combinações com novos verbos por analogia a esses modelos já existentes e, conseqüentemente, pode produzir expressões que não tem valor de verbo frasal em L2.

Os exemplos (5) e (6) mostram casos de possíveis usos criativos de combinação entre verbo lexical e partícula adverbial:

(5) “*The problem is that such a system is impossible. Man’s desire to rise ahead of others is too strong, and money is not the only thing that man **competes over**.*” (Br-ICLE).

(6) “*Women and men who have never even studied the feminist agenda are, and have been, **passing forward** its ideological thinking, having understood that both genders are equal and can take part in the same social, economical and political activities, as opposed to the understanding of the past generations, which considers women and men to have different roles that should be respected and not surpassed.*” (Br-ICLE).

(7) “*May I ask er, er Mr (-----) that (unclear) it would, whether she would now er **pass forward** the petition.*” (BNC/JWA).

Apesar de “*compete over*” e “*pass forward*” não existirem como *phrasal verbs* em inglês, não podemos considerar que as construções sejam, de fato, um erro produzido por aprendizes. O modo como foram elaboradas demonstra que houve consciência acerca da conotação das partículas adverbiais empregadas (*over* indicando oposição - como em *run over* -, e *forward*, posição) ou que, pelo menos, os aprendizes conhecem e memorizam o valor aspectual dos verbos frasais produzidos a partir de *over* e *forward*. A combinação no exemplo (5) apresenta o que seria um verbo frasal intransitivo, ou seja, que não requer objeto. Porém, a formação criada não é adequada e/ou gramaticalmente suficiente para se transmitir a ideia de que o dinheiro é um dos fatores pelos quais os homens competem uns com os outros. Além disso, nenhum verbo frasal similar consta nos dicionários consultados nesta pesquisa e a análise das colocações no BNC mostra que “*compete*” não forma colocado com “*over*”, mas sim com “*against*”, o que contribui para a identificação de um desvio.

A frase (6), por sua vez, mostra a junção entre o verbo “*pass*” e o advérbio “*forward*” formando algo como “avançar” e, apesar de não existir nos dicionários consultados, a combinação apresenta características similares às dos verbos frasais, como a mobilidade da partícula. Por outro lado, há apenas uma ocorrência dessa expressão no BNC (7), que foi encontrada no *corpus* de língua falada, ou seja, é uma produção rara na língua nativa e, se analisarmos semanticamente a construção, veremos que no contexto do CR ela tem um significado bastante transparente, o que a tornaria um PV não idiomático (DOWNING, 2015) ou uma combinação livre, já que ambas as partes que formam a expressão mantêm seus significados individuais (*pass forward the petition* = passar a petição adiante). A análise da lista de colocados para *pass* mostra que esse verbo é normalmente acompanhado por *through*, o que reforça o raro uso da combinação analisada.

De maneira geral, a incidência de novos verbos frasais no *corpus* de estudo é bem pequena, o que pode sinalizar que os aprendizes brasileiros só utilizam os PVs cujos significados já são conhecidos ou com os quais eles já estão familiarizados. Por outro lado, isso pode ser um reflexo do gênero textual analisado - uma vez que seu caráter mais formal não favorece o emprego dessas expressões -, da baixa frequência de *phrasal verbs* nos textos investigados ou ainda de fatores desconhecidos na produção dos textos, como o tempo de contato dos aprendizes com a L2 e o nível de proficiência desses estudantes, o que influenciaria no uso consciente dos PVs.

### 5.3 Naturalidade no idioma

Fatores semânticos, sintáticos e lexicais podem revelar a falta de naturalidade do aprendiz em relação a uma língua, ou seja, aspectos como a escolha errada de palavras, a falta de coerência textual e a falta de conhecimento das propriedades e regras gramaticais podem ressaltar a condição de falante não nativo dos estudantes, fazendo com que seu desempenho no idioma soe estranho para os falantes nativos.

No domínio dos verbos frasais da língua inglesa, a falta de naturalidade pode estar relacionada a desvios colocacionais, à escolha inapropriada de um verbo ou de uma partícula, ou ainda à simplificação dos significados das combinações entre verbo lexical e advérbio, alterando, inclusive, o caráter idiomático dessas estruturas. No *corpus* de estudo, pudemos observar alguns casos em que o emprego dos PVs representou falta de naturalidade no idioma, como no exemplo (8):

(8) *“Considering this situation in our country, we would face to a huge catastrophic environment because our Governmental System wouldn't support the over population. You are able to imagine São Paulo, which already have **turn around** car as solution for traffic jam problems, living with even more people”.* (Br-ICLE).

(9) *“It's a disease and people will **turn around** and say you can have controlled drinking and do this and that.”* (BNC/CEN).

Na frase (8), a intenção do aprendiz é informar que uma das soluções para reduzir os problemas com o tráfego de São Paulo é o rodízio de carros. Para isso, ele usa a expressão *turn around*, que significa, na verdade “virar na direção oposta”. Logo, vê-se que o uso do PV escolhido não se adequa ao contexto aplicado, uma vez que seus significados possíveis não são compatíveis com o que o aprendiz tentou comunicar. A princípio, podemos considerar que a escolha desse PV ocorreu sob a influência da L1, já que, semanticamente, os componentes da combinação produzem um significado que remete à ação de rodízio - movimento em alternância (nesse caso, o verbo *turn* indica o movimento → girar, e *around* a alternância ou a falta de uma finalidade/rumo → aproximadamente, ao redor, por toda parte etc.). Porém, se analisarmos o emprego desse PV pelos nativos (9), veremos que ele é aplicado predominante com o sentido de virar-se, de ir no sentido oposto e que a expressão usada para se referir a rodízio de carros é *“road space rationing”*. Há, portanto, diferença de uso entre os grupos, o que sugere não só uma inadequação na escolha da expressão por parte dos aprendizes

brasileiros, mas também a falta de naturalidade em relação ao padrão utilizado pelos nativos. Além disso, a análise das linhas de concordância no BNC para *turn around* mostra que *say* é seu principal colocado (em contextos como “*turn around and say something*”), não havendo ocorrências de *car* nas proximidades no nóculo de pesquisa.

Outro caso da influência de L1 pode ser visto no exemplo (10):

(10) “*They are seen as “ex-criminals” and for this reason the companys <sic> do not hire them officialy <sic>, just informaly <sic>. This is one reason for the crime keep continuing because society **put away** the ex-criminals.*” (Br-ICLE).

(11) “*She then goes home to begin preparing something for dinner. Once dinner has been started she rounds up the dirty clothes so that they can be washed, dried, and **put away**.*” (LOCNESS).

(12) “*I **put away** the money to pay that bill and if I haven't got anything left over to buy food with, then we have to manage.*” (BNC/FST).

(13) “*They let me out for the day and then, when he regained consciousness five days later, they let me out again to see him. Then John got done for burglary and he was **put away** five days before I got out.*” (BNC/EDC).

A expressão *put away* tem significado de prender ou matar alguém e também de guardar algo em seu devido lugar ou “deixar algo de lado”, como mostram os exemplos (11), (12) e (13), retirados dos *corpora* de referência. O trecho extraído do *corpus* de estudo, por outro lado, mostra que o aprendiz brasileiro utilizou o PV para dizer que ex-criminosos são afastados ou isolados da sociedade, ou seja, “alguém é deixado de lado”, um significado muito próximo semanticamente da tradução literal das palavras que compõem o verbo frasal (*put* = colocar / *away* = longe, à distância). Apesar de não ter sido encontrada dessa forma nos dicionários pesquisados, essa ocorrência pode não constituir necessariamente um erro, haja vista a proximidade com o equivalente mais adequado “*put aside*”, mas, a partir da análise das linhas de concordância dos *corpora* de referência, vemos que ela não ocorre ou é rara nos textos de nativos, o que faz com que seu uso seja considerado pouco natural. Além disso, a análise dos colocados no BNC mostra que a expressão *put away*, quando aplicada para indicar que “algo” foi deixado de lado, nunca está relacionada a pessoas (*somebody*, *someone*), mas sempre a coisas e objetos (*something*), o que contribui para a inadequação no que toca aos padrões de uso da expressão em L1.

A sentença apresentada em (14) mostra mais um caso de desempenho pouco natural de aprendizes:

(14) *“For many years the belief of an university, as space discussion and knowledge, graduating students for the life **gave up place for** the formation for the labor. Space of arguments and above that, researches, to a place that general population starts plans to find a job, simply was lowered to a tutorial course.” (Br-ICLE).*

(15) *“This type of grave-marker, however, now **gave place to** marble monuments, and the masterpieces of sixth-century black-figure are on a more modest scale.” (BNC/FPW).*

(16) *“In bronze the great tripod-cauldrons **give place to** a new fashion: tall, conical stands supporting big bowls with incurving rim, their ring-handles held by attachments in the form of human-headed birds (sirens); and, standing round the shoulder between these, long-necked griffin-heads.” (BNC/FPW).*

O exemplo (14) acima mostra a aplicação desnecessária da partícula adverbial *up*. Ao mencionar que a universidade, instituição que antes formava o aluno para a vida, passa a dar lugar para à formação visando ao mercado de trabalho, o aprendiz utiliza o verbo frasal *give up* - cujo significado real é de desistir, render-se, parar de fazer algo -, inconsciente de que a escolha não é adequada ao propósito do contexto. A análise de *give up* no BNC mostra que os nativos utilizam a mesma expressão apenas empregando o verbo “*give*”, como em (15) e (16), o que representa a forma natural na fala. Além disso, a análise de colocados para o PV indica que ele aparece predominantemente seguido da preposição *to* e não de *for*, o que sugere também um desvio colocacional por parte do aprendiz.

O erro apresentado em (17) também mostra um caso de uso inadequado da partícula adverbial alterando o significado da frase:

(17) *“Considering the fact that in Brazil, according to a recent report about education published in “Revista Já”, from “Diário Popular” journal, 17,2% of our population don't know to read, neither to write, it's evident that the illiterate people, or even those who studied only in the high school, are more and more restricted to **find out** a job, even though, if this job is to guarantee a payment just to survive.” (Br-ICLE).*

(18) *“Usually, couples which are in love, say that they could not live without the other, but what we have seen through the decades, is that they replace the beloved and*

*incomparable partner as soon as they **find out** another, and all of those promises they had done just disappeared.”* (Br-ICLE).

No exemplo (17), o aprendiz utiliza o PV *find out*, que significa “descobrir uma informação” em uma contextualização na qual apenas o verbo único *find* seria suficiente, já que a mensagem expressa é a restrição de “encontrar um emprego”. Nesse seguimento, o emprego da partícula adverbial alterou o sentido da frase, já que, em combinação com o verbo lexical, formou uma expressão idiomática. O mesmo acontece no exemplo (18), em que *find out* é usado com sentido de “encontrar um novo amor/parceiro”. O que ocorre nesses dois casos é que o verbo frasal empregado pelos estudantes foi utilizado sem a consciência de que não se adequava aos contextos apresentados. *Find out* está geralmente ligado às descobertas no campo do conhecimento, à aquisição de novas informações e não deve ser usado como substituto de *find*. De qualquer forma, não podemos afirmar que há de fato um erro nesses empregos, já que sentenças similares foram encontradas (em poucas ocorrências) também no *corpus* de nativos. Contudo, certamente, trata-se de sentenças pouco naturais em relação ao que é produzido por falantes da língua inglesa.

Como visto nos exemplos, os textos analisados expuseram que, na escrita acadêmica, os aprendizes não nativos soam pouco naturais em razão, principalmente, de erros relacionados ao emprego de palavras inadequadas e à escolha errada de verbos ou partículas. Com menor frequência, a simplificação semântica também ocorre, contribuindo para a formulação de sentenças estranhas em L2. Ademais, o desvio colocacional é um dos problemas mais recorrentes entre os aprendizes brasileiros, provavelmente impactando menos na compreensão dos nativos (já que alguns dos erros também ocorrem nas produções de falantes cujo inglês é a língua materna), mas comprometendo, da mesma forma, o desempenho dos aprendizes no aspecto da naturalidade.

### 5.3.1 Desvios colocacionais

Como já mencionado anteriormente, os desvios colocacionais são também responsáveis pela expressão pouco natural dos aprendizes em L2. As sentenças a seguir exemplificam casos de inadequações encontradas no *corpus* de estudo.

(19) *“This position **goes on assert** that technology and industrialization become people’s life effortless and pleasurable.”* (Br-ICLE).

(20) “*This position **goes on to assert** that due to the pressures <sic> of modern life and the rat race that constitutes corporal world leave people no time to do any kind of creative work.*” (Br-ICLE).

(21) “*The tenth-century Arab document **goes on to assert** that the sect from whom the Nazarean text issued is still in existence, and is regarded as an elite amongst Christians.*” (BNC/EDY).

(22) “*He **went on to assert** his first mature, outrageous, filmmaking style with a D. H. Lawrence adaptation, *Women in Love* (1969), *The Music Lovers* (1970), covering the life of Tchaikovsky [...]*”. (BNC/A7L).

(23) “*At this age, teenagers are trying to understand what's **going on the world**, they want to take point in it and they need socialize with other people their age.*” (Br-ICLE).

A análise de colocados para *go on* em textos acadêmicos do BNC indicou que essa expressão é frequentemente seguida por “*assert*”. Esse tipo de construção também ocorre no *corpus* de estudo, porém, como é possível verificar em (19), as linhas de concordância mostram que nem sempre os aprendizes sabem associar o PV com as palavras adequadas. A relação entre *go on* e *assert*, como se vê em (20), é realizada por meio da preposição *to*. Os exemplos (21) e (22), extraídos do BNC, corroboram essa afirmação. A falta da preposição no exemplo (23) também mostra um desvio colocacional, contudo, nesse caso, não é possível identificar se se trata de um erro gramatical por falta de atenção do aprendiz ou se é um caso de influência da língua materna, com a tradução direta de “*on the world*” como “no mundo” afetando a produção final, e a não consciência de que *going on* só funciona como “acontecendo” se partícula e verbo lexical estiverem unidos semanticamente. O fato é que todos esses desvios relacionados a colocados interferem no desempenho do aprendiz, especialmente no do não nativo, que acaba produzindo sentenças que não soam naturais para os falantes nativos.

Outro erro comum percebido em relação aos colocados nos textos de não nativos é que a ausência de preposições transforma em verbos frasais expressões já estabelecidas nos dicionários (e entre os falantes nativos) como verbos frasais preposicionados. O exemplo (24) ilustra um caso no CE:

(24) “*In the teen years, the person is looking for what he/she is. They can't see what they are for themselves, so they try to look like someone they think is cool, they imitate that person: wearing the same kind of clothes, talking alike... Nowadays the midia*

*creates these "coll <sic> people", teens **look up at** them and start imitating them, and that means they buy what the companies want."* (Br-ICLE).

No trecho acima, o aprendiz informa que os adolescentes são influenciáveis e que a mídia cria pessoas “legais” que são admiradas e imitadas pelos jovens. No inglês, a palavra “admiradas” é representada pelo *phrasal prepositional verb* “*look up to*”, uma construção cuja estrutura já é fixada no idioma. Como é possível verificar em (24), um desvio colocacional ocorre quando o aprendiz utiliza a preposição *at*, possivelmente sem saber que para aplicar o significado esperado na frase, a preposição adequada é *to*. Algo similar acontece no exemplo (25), porém, nesse caso, o aprendiz não utilizou preposição alguma, causando um problema no *phrasal prepositional verb* e criando um *phrasal verb* semanticamente inapropriado para o contexto aplicado.

(25) *“Although, prisons are important tools, society must **face up** the fact that each year, larger numbers of people are being incarcerated and effective measures must be taken in order to deal with this issue.”* (Br-ICLE).

(26) *“Children and adults come to learn and **face up to** the facts of caring and sharing on the principal that we can not survive alone on Earth -- we must share it with all the other animals and plants.”* (BNC/KA1).

As sentenças (25) e (26) mostram ambas o uso de uma combinação com o significado de “encarar”. Porém, enquanto a frase produzida pelo nativo exibe uma colocação com a preposição, o que dá sentido completo e torna natural a construção como um verbo frasal preposicionado para os falantes de inglês, a sentença extraída do CE manifesta um desvio, afinal, *face up* sozinho não constitui significado idiomático.

É importante ressaltar, por fim, que apesar de todos os exemplos deste tópico retratarem situações pertinentes à escrita de não nativos, o exame das linhas de concordância mostra que os nativos também cometem erros de colocação, mas em quantidades bem inferiores, o que é mais do que natural, considerando-se as circunstâncias.

#### **5.4 Estilo e formalidade na escrita acadêmica**

Aprendizes de inglês como segunda língua nem sempre estão atentos ou têm consciência das diferenças existentes entre a escrita formal e a informal, o que faz com que



muitas das palavras e/ou expressões empregadas em seus textos sejam consideradas inadequadas dentro de um determinado gênero textual. Um exemplo é o uso de verbos frasais considerados extremamente coloquiais em textos acadêmicos, como vemos em (27), (28) e (29):

(27) “*You read newspaper articles about how people are leaving busy cities and **heading off** to the countryside to live the so-called simple life.*” (Br-ICLE).

(28) “*I remember I had games that stimulated my creativity, when I was a teenager I used to **hang out** with my friends every week, we got together, we played games, we danced.*” (Br-ICLE).

(29) “*They will have to compete with their colleagues <sic> to get a job, it doesn't matter if the job will be good or not, once they have just been **kicked out** into the market. And the struggle goes on.*” (Br-ICLE).

O PV *head off* significa “começar uma viagem” ou “deixar algum lugar”, que é o sentido indicado em (27). Entretanto, essa é uma combinação considerada bastante informal, o que pode afetar o registro textual. Nesse caso, a formalidade exigida pelo gênero acadêmico poderia ser alcançada pela substituição do PV pelo verbo único “*leave*”, que carrega o mesmo significado, mas que contribui para a manutenção do estilo formal. *Hang out* (28), que significa “sair para passear”, também é uma expressão mais coloquial e, a fim de se criar uma adequação ao tipo de texto, poderia ser substituída por “*spend time*”. Da mesma forma, *kick out* (expulsar) em (29) é um indicador de registro informal e, ao ser usado no texto acadêmico, pode sinalizar uma inadequação ao estilo textual. Nesse caso, o aprendiz poderia ter usado uma expressão equivalente e menos coloquial, como “*expel*” ou ainda “*force to leave*”.

No que diz respeito às características de estilo e formalidade, cabe ressaltar que as temáticas sugeridas para os textos aqui analisados parecem permitir um grau menor de seriedade, diferentemente do que ocorreria com um texto de cunho científico, por exemplo. Dessa forma, o uso de certos verbos frasais parece ser bastante aceitável entre os aprendizes. Como mostrou a análise quantitativa, esse tipo de combinação foi encontrada tanto nos textos de nativos, quanto nos de não nativos, indicando que o registro predominante é semiformal. Além disso, percebeu-se em ambos os *corpora*, mas em especial no *corpus* de estudo, o emprego de vários PVs que parecem ser típicos dos textos argumentativos e que são usados para marcar a ordenação das ideias, dos argumentos e da estrutura do texto, como é o caso de *sum up*, *point out* e *put forward*.

## 6 CONSIDERAÇÕES FINAIS

Este trabalho investigou o uso dos *phrasal verbs* na escrita acadêmica de aprendizes brasileiros da língua inglesa, tendo como base para comparação as produções textuais de falantes nativos do inglês. A pesquisa que deu origem a esse documento foi realizada por meio da análise de *corpora* (que permitiu o estudo de fenômenos da língua em uso) e teve como objetivo principal descrever e comparar o emprego de verbos multipalavras em textos argumentativos através de análises quantitativas e qualitativas.

A primeira parte da observação pretendeu responder às seguintes questões: 1) Quais são os *phrasal verbs* mais frequentes em ambos os *corpora*? 2) No emprego de *phrasal verbs*, quais são as similaridades e as diferenças percebidas na escrita de nativos e de aprendizes brasileiros? 3) Quais são as tendências e/ou preferências de uso dos aprendizes em relação aos nativos? Há evidências que apontem que os aprendizes brasileiros prefiram verbos latinizados e/ou evitem o uso de *phrasal verbs*? A segunda parte, por sua vez, teve como norte as seguintes perguntas: 4) No que diz respeito ao emprego dos *phrasal verbs* e seus significados, quais são os principais problemas encontrados nos textos de não nativos em comparação com os de nativos? 5) Em relação aos problemas, quais são os possíveis motivos? Pode-se considerar que há ocorrências de novos significados atribuídos a verbos frasais já existentes e/ou formações de novas combinações verbo-partícula?

A resposta para a primeira pergunta, relativa à análise quantitativa, é que os verbos frasais do inglês são, de maneira geral, pouco empregados na escrita acadêmica, entretanto, dos mais frequentes, vários são considerados típicos de textos mais formais, como *sum up* e *point out*. O resultado dos *phrasal verbs* mais frequentes nos *corpora* pode ser melhor visualizado no capítulo 4, que expõe detalhadamente a análise quantitativa dos dados.

A resposta para a segunda pergunta é que sim, existem semelhanças e diferenças no emprego das combinações entre verbo + partícula adverbial nos textos produzidos por nativos e não nativos do inglês. As semelhanças estão relacionadas à frequência dos verbos frasais, especialmente no que diz respeito à quantidade de PVs empregados, à similaridade entre as ocorrências (os verbos frasais mais frequentes no *corpus* de estudo também estão presentes, em grande parte, no *corpus* de referência) e ao comportamento dos indivíduos no que concerne à utilização de verbos multipalavras e de verbos únicos.

Para a terceira pergunta, a resposta é que não é possível afirmar que os aprendizes brasileiros apresentam inclinação ou predileção pelo uso de palavras de origem latina, já que,

em determinados momentos, verificou-se que houve a preferência por elementos latinizados por parte dos dois grupos. Tal preferência, provavelmente, foi fomentada pela estrutura do tipo textual em questão e da formalidade que lhe é pertinente. Por outro lado, a alta frequência de alguns verbos frasais no *corpus* de referência e a ausência destes no *corpus* de estudo sinalizou que pode sim haver uma tendência dos não nativos para evitar certas construções, entretanto, é necessário um estudo semântico mais aprofundado para se descobrirem as motivações do fenômeno. As diferenças percebidas entre os grupos, por sua vez, são relativas, especialmente, ao aspecto semântico e implicam no significado que é atribuído a alguns PVs por parte dos aprendizes brasileiros. Além disso, em alguns casos, os grupos diferem também na preferência do emprego de palavras de verbo único, que, em alguns momentos, são escolhidos em detrimento de verbos multipalavras.

Em relação à análise qualitativa, a resposta para a quarta pergunta é que os principais problemas nos textos dos não nativos estão vinculados a inadequações sintáticas e semânticas, que impactam diretamente na naturalidade dos aprendizes em L2 e no desempenho que obtêm na língua. A resposta para a quinta pergunta complementa a resposta da quarta questão e aponta que os motivos para os problemas dos aprendizes no uso de verbos frasais partem, em sua maioria, de escolhas inapropriadas de partículas e ou verbos lexicais, de desvios colocacionais (provavelmente pelo desconhecimento desse tipo de evento na língua inglesa) e do emprego de verbos frasais cujos significados não são adequados aos contextos em que são aplicados, gerando problemas de sentido nas sentenças e certo estranhamento na comunicação escrita no idioma. Além disso, no *corpus* de estudo, há sim alguns casos de formação de novas combinações verbo-partículas e de novos significados dados a verbos frasais já estabelecidos na língua nativa, porém, a ocorrência é bastante baixa, o que não nos permite tirar conclusões acerca desse aspecto no comportamento de não nativos.

Finalmente, cabe aqui ressaltar que esta pesquisa não tem o objetivo de ser conclusiva. Ao contrário, os resultados mostraram que muito ainda pode ser observado sobre esse fenômeno, especialmente no campo da semântica. Entretanto, este estudo chama a atenção para a relevância dos estudos baseados em *corpus* de aprendizes como ferramenta para a descrição da língua em uso, bem como para a observação de aspectos importantes do comportamento de aprendizes de inglês como segunda língua. A investigação aqui descrita pretende ter contribuído não só para as pesquisas envolvendo os *phrasal verbs* e suas particularidades no que toca à escrita de não nativos, mas também para as pesquisas que envolvem processos de aquisição da língua inglesa.

## REFERÊNCIAS

- ANTHONY, L. *AntConc*. Disponível em <<http://www.antlab.sci.waseda.ac.jp/software.html>>. Acesso em: 15 dez. 2016.
- BAKER, M. Corpora in Translation Studies: An overview and some suggestions for Future Research. *Target: International Journal of Translations Studies*. 7 (2), p. 223-243, 1995.
- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri, São Paulo. Editora Manole, 2004.
- \_\_\_\_\_. Linguística de Corpus: Histórico e Problemática. *DELTA – Documentação de Estudos em Linguística Teórica e Aplicada*. São Paulo, v. 16, n. 2, 2000. Disponível em: <<http://goo.gl/GtyHz1>>. Acesso em: 23 jan. 2016.
- BIBER, D. A Corpus Driven Approach to formulaic Language in English. Multi-word Patterns in Speech and Writing. In: *International Journal of Corpus Linguistics* 14:3. 235-311. John Benjamins Publishing Company, 2009.
- BIBER, D; JOHANSSON, S; LEECH, G; CONRAD, S; FINEGAN, E. *Longman Grammar of Spoken and Written English*. London: Longman, 1999.
- BRAZILIAN SUB-CORPUS OF THE INTERNATIONAL CORPUS OF ENGLISH (BR-ICLE). Disponível em: <<http://www2.lael.pucsp.br/corpora/bricle/>>. Acesso em: 24 jan. 2016
- BYWATER, F. V. *A proficiency course in English*. London: University London Press, 1969.
- CARTER, R; MCCARTHY, M. *Cambridge Grammar of English*. A Comprehensive Guide: Spoken and Written English Grammar and Usage. Cambridge: Cambridge University Press, 2006.
- CELCE-MURCIA, M; LARSEN-FREEMAN, D. *The Grammar Book: An ESL/EFL Teacher's Course*. Boston, MA: Heinle and Heinle Publishing Company, 1999.
- CORNELL, A. Realistic goals in teaching and learning phrasal verbs. *IRAL – International Review of Applied Linguistics in Language Teaching*, 23 (4), p. 269-280, 1985.

COSTA, P. T. *O uso dos multi-word verbs em textos acadêmicos de estudantes brasileiros do inglês: um estudo comparativo baseado em corpora*. 2012. 61 f. Monografia (Graduação em Letras) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

DAGUT, M.; LAUFER, B. Avoidance of phrasal verbs – a case for contrastive analysis. *Studies in Second Language Acquisition*, Vol. 7, p. 73-80, 1985.

DARWIN, C. M.; GRAY, L. S. Going after the phrasal verb: An alternative approach to classification. *TESOL Quarterly*, Vol. 33, n. 1, p. 65-83, 1999.

DEMPSEY, K. B.; MCCARTHY, P. M.; MACNAMARA, D. S. Using Phrasal Verbs as an Index to Distinguish Text Genres. *The Florida AI Research Society Conference - FLAIRS*, p. 217-222, 2007.

DOWNING, A. *English grammar: a university course*. 3ed. Revised Edition. New York: Routledge, 2015.

DOWNING, A.; LOCKE, P. *A university course in English grammar*. Revised Edition. London; New York: Routledge, 2006.

DUCROT, O; TODOROV, T. *Dicionário enciclopédico das ciências da linguagem*. Perspectiva, 2001.

ENGLISH OXFORD LIVING DICTIONARY. Edição *on-line*. Oxford University Press, 2017.

FADANELLI, S. B. A Corpus study on Brazilian learners' usage of English phrasal verbs. *Revista Estudos Anglo-Americanos*, n. 37, 2012. Disponível em: <<http://goo.gl/2xxTh8>>. Acesso em: 27 jan. 2016.

FARIA, S. B.; BERNARDO, D. C.; SILVA, F. X. G. Phrasal Verbs em inglês: aprender é o melhor remédio. *Cadernos de Letras*, n. 23, p. 81-95. UFRJ, Rio de Janeiro, jan./dez. 2007.

FERNADES, G. G. M. *O uso de chunks formados pelo verbo get por aprendizes de inglês como L2*. 2012. 222 f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2012. Disponível em: <<https://goo.gl/nVqUxc>>. Acesso em: 27 jan. 2016.

FLETCHER, B. Register and phrasal verbs. *MED Magazine*. The Monthly Webzine of McMillan English Dictionaries, v. 33, p. 45-54, 2005.

GARDNER, D.; DAVIES, M. Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL Quarterly*, Vol. 41, n 2, Jun. / 2007, p. 339-359.

GRANGER, S. From CA to CIA and back: An integrated approach to computerized bilingual corpora and learner corpora. In: AIJMER, K.; ALTENBERG, B.; JOHANSSON, M. *Languages in Contrast*. Lundi: Lundi University Press, 1996, p. 37-51.

\_\_\_\_\_. On identifying the syntactic and discourse features of participle clauses in academic English: native and non-native writers compared. In: AARTS, J.; DE MÖNNINK, I.; WEKKER, H. (eds). *Studies in English Language and Teaching*. Rodopi: Amsterdam & Atlanta, 1997, p. 185-198.

GREENBAUM, S. QUIRK, R. *A Student's Grammar of the English Language*. Longman, 1990.

HALLIDAY, M. A. K. *An introduction to functional grammar*. Revised Edition. London: Hodder Arnold Publication, 2004.

HARPER, D. *Online Etymology Dictionary*. Copyright 2001-2017. Disponível em: <<http://www.etymonline.com/>>. Acesso em: 23 dez. 2016.

KAMARUDIN, R. *A Study on the Use of Phrasal Verbs by Malaysian Learners of English*. 2013. 381 f. Thesis (Ph.D. in Philosophy) - University of Birmingham, Birmingham, 2013. Disponível em: <<http://goo.gl/iEyksN>>. Acesso em: 21 dez. 2015.

LIAO, Y.; FUKUYA, Y. J. Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning*, 2004, p. 193-226.

LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH. New Edition. Pearson Education Limited: Harlow, 2003.

LOUVAIN CORPUS OF NATIVE ENGLISH (LOCNESS). Disponível em: <<https://www.uclouvain.be/en-cecl-locness.html>> Acesso em: 5 dez. 2015.

MAIR, C. Quantitative or qualitative corpus analysis? Infinitival complement clauses in the Survey of English Usage corpus. In: JOHANSSON, Stig; STENSTRÖM, Anna-Brita (Ed.).

*English computer corpora. Selected papers and research guide.* Vol. Topics in English linguistics: 3. Berlin & New York: Mouton de Gruyter, 1991, p. 67-80.

MCCARTHY, M.; O'DELL, F. *English phrasal verbs in use.* Cambridge University Press, 2004.

MCCARTHY, M.; O'KEEFFE, A. Historical Perspective: What are corpora and how have they evolved? In: *The Routledge handbook of Corpus Linguistics.* Routledge, 2010.

MCENERY, T; WILSON, A. *Corpus linguistics: An introduction.* Edinburgh University Press, 2001.

OXFORD PHRASAL VERBS DICTIONARY FOR LEARNERS OF ENGLISH. 1ª Edição. Oxford University Press: Oxford, 2001.

PALMER, F. R. *The English Verb.* Second edition. London: Longman, 1988.

QUIRK, R; GREENBAUM, S; LEECH, G; SVARTVIK, J. *A comprehensive grammar of the English language.* London: Longman, 1985.

RAYSON, P. *Matrix: A statistic method and software tool for linguistic analysis through corpus comparison.* 2002. 207 f. Thesis (Ph.D. in Computer Science). Computer Department, Lancaster University, Lancaster, 2002. Disponível em: <<https://goo.gl/KhDr91>>. Acesso em: 8 ago. 2016.

RYOO, M. Corpus-based Study of the Use of Phrasal Verbs in Korean EFL Students' Writing. *The Journal of Asia TEFL*, Vol.10, No. 2, 2013, p. 63-89.

SANTOS, E. M. *Engenharia Lingüística: Uma tecnologia para apoiar as decisões gerenciais na era da Internet.* Rio de Janeiro: Editora E-papers, 2008.

SIDE, R. Phrasal Verb: Sorting Them out. *ELT Journal*, v. 44 (2), 1990, p. 144-152.

SINCLAIR, J. *Developing linguistic corpora: a guide to good practice – Corpus and text – basic principles.* Versão de texto eletrônico. 2004. Disponível em: <<http://goo.gl/EK3Oax>>. Acesso em: 8 jan. 2016.

TAGNIN, S. E. O. *Glossário de Linguística de Corpus.* São Paulo: HUB Editorial, 2010.

THE BRITISH NATIONAL CORPUS (BNC). Online Edition. Oxford University Computing Services on behalf of the BNC Consortium. Disponível em: <<http://corpus.byu.edu/bnc/>>. Acesso em: 28 out. 2016.

TEIXEIRA, E. D. *A linguística de corpus a serviço do tradutor: proposta de um dicionário de culinária voltado para a produção textual*. 2008. 439 f. Tese (Doutorado em Letras) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2008. Disponível em: <<https://goo.gl/xOMm6S>>. Acesso em: 29 dez. 2016.

WAIBEL, B. *Phrasal verbs in learner English: A corpus-based study of German and Italian students*. 2007. 212 f. Dissertation (Doktorwürde der Philologischen) - Fakultät der Albert-Ludwigs, Universität Freiburg, Alemanha, 2007. Disponível em: <<http://www.freidok.uni-freiburg.de/volltexte/3592/>>. Acesso em: 21 dez. 2015.

WALKOVÁ, M. *The aspectual function of particles in phrasal verbs*. 2013. 237 f. Dissertation (Ph.D. in Linguistics) - Faculty of Arts, University of Groningen, Países Baixos, 2013. Disponível em: <<http://goo.gl/e7V9Eb>>. Acesso em: 21 ago. 2016.



**APÊNDICE A - As 270 palavras mais frequentes Br-ICLE (Freq. bruta).**

Rank	Palavras	Freq.	Rank	Palavras	Freq.	Rank	Palavras	Freq.
1	the	9036	44	society	479	87	make	258
2	to	5587	45	which	477	88	like	249
3	of	4864	46	an	473	89	just	248
4	and	4590	47	them	467	90	very	244
5	a	3573	48	because	465	91	work	239
6	in	3474	49	about	453	92	such	237
7	is	3300	50	so	451	93	nowadays	226
8	that	2778	51	many	434	94	us	224
9	it	2132	52	time	426	95	think	221
10	are	1742	53	you	424	96	children	219
11	not	1589	54	other	419	97	different	214
12	for	1545	55	i	414	98	get	213
13	be	1467	56	life	408	99	those	211
14	they	1424	57	when	404	100	new	209
15	people	1422	58	would	404	101	live	208
16	have	1310	59	these	397	102	lives	208
17	this	1194	60	television	387	103	fact	205
18	as	1182	61	at	383	104	had	205
19	we	1054	62	only	364	105	its	203
20	their	984	63	way	362	106	her	202
21	with	829	64	should	361	107	may	202
22	s	798	65	he	360	108	men	201
23	can	795	66	also	338	109	job	200
24	or	790	67	most	334	110	system	200
25	on	778	68	even	330	111	person	197
26	more	727	69	fileseparator	326	112	real	194
27	all	722	70	been	324	113	need	192
28	by	680	71	were	300	114	without	192
29	but	652	72	than	290	115	imagination	189
30	what	624	73	t	286	116	religion	188
31	has	592	74	good	283	117	same	188
32	world	582	75	students	278	118	she	188
33	do	579	76	university	278	119	something	188
34	there	568	77	how	276	120	technology	188
35	our	553	78	things	275	121	better	185
36	one	540	79	being	272	122	order	184
37	if	518	80	important	271	123	well	184
38	was	505	81	could	269	124	tv	183
39	who	502	82	his	266	125	much	180
40	will	502	83	human	266	126	does	177
41	money	495	84	women	266	127	into	176
42	some	489	85	no	264	128	any	173
43	from	482	86	however	260	129	still	173

**Continua**

## Conclusão

Rank	Palavras	Freq.	Rank	Palavras	Freq.	Rank	Palavras	Freq.
130	social	172	177	right	123	224	means	96
131	years	172	178	over	120	225	go	95
132	power	171	179	having	119	226	between	94
133	really	168	180	brazil	118	227	little	94
134	up	162	181	each	118	228	today	94
135	become	160	182	day	116	229	every	93
136	course	160	183	own	116	230	themselves	93
137	know	160	184	government	115	231	living	92
138	problems	160	185	my	114	232	considered	91
139	since	160	186	now	114	233	degree	91
140	must	159	187	theoretical	114	234	degrees	91
141	see	158	188	although	112	235	both	90
142	example	156	189	change	112	236	learn	90
143	great	155	190	part	112	237	too	90
144	take	152	191	reality	112	238	keep	89
145	first	148	192	never	111	239	someone	89
146	others	148	193	problem	111	240	doing	88
147	equal	147	194	situation	111	241	future	88
148	want	147	195	find	110	242	process	88
149	another	145	196	ones	110	243	role	87
150	through	145	197	then	109	244	century	86
151	love	142	198	man	108	245	times	86
152	modern	141	199	hand	107	246	around	85
153	use	141	200	practice	107	247	cannot	85
154	used	141	201	prison	107	248	made	85
155	say	140	202	courses	106	249	practical	85
156	possible	139	203	criminals	106	250	control	84
157	after	138	204	why	106	251	dreams	84
158	kind	138	205	able	105	252	bad	83
159	nature	137	206	rights	104	253	enough	83
160	always	136	207	necessary	103	254	look	83
161	theory	136	208	don	102	255	old	83
162	out	135	209	dreaming	101	256	parents	81
163	universities	135	210	give	101	257	consider	80
164	believe	134	211	programs	101	258	countries	80
165	education	134	212	according	100	259	country	80
166	family	134	213	before	100	260	pay	80
167	place	133	214	thing	100	261	end	79
168	information	132	215	market	99	262	matter	79
169	your	132	216	point	99	263	value	79
170	knowledge	127	217	science	99	264	view	79
171	number	127	218	learning	98	265	while	79
172	where	126	219	against	97	266	best	78
173	lot	125	220	did	97	267	everybody	78
174	crime	124	221	help	97	268	idea	78
175	everything	124	222	said	97	269	once	78
176	therefore	124	223	going	96	270	sometimes	78

**APÊNDICE B - As 270 palavras mais frequentes LOCNESS (Freq. bruta).**

Rank	Palavras	Freq.	Rank	Palavras	Freq.	Rank	Palavras	Freq.
1	the	21106	44	many	925	87	make	433
2	to	10753	45	who	910	88	world	432
3	of	10730	46	also	861	89	us	420
4	and	8325	47	because	854	90	does	415
5	a	6853	48	when	843	91	him	408
6	in	6371	49	these	839	92	women	402
7	is	6307	50	been	793	93	money	400
8	that	4934	51	should	771	94	children	399
9	it	3324	52	only	757	95	any	395
10	be	3198	53	were	735	96	candide	365
11	for	3145	54	so	723	97	much	364
12	as	2833	55	other	715	98	use	362
13	this	2806	56	do	708	99	icle	361
14	are	2555	57	what	698	100	xe	359
15	not	2403	58	life	680	101	just	358
16	he	2220	59	no	674	102	own	352
17	they	2100	60	could	635	103	like	346
18	have	2046	61	however	591	104	over	333
19	with	1909	62	had	590	105	made	327
20	on	1797	63	our	584	106	those	326
21	by	1703	64	them	581	107	human	322
22	s	1631	65	about	578	108	must	322
23	people	1627	66	being	572	109	man	320
24	his	1564	67	out	569	110	now	320
25	has	1555	68	you	551	111	death	319
26	was	1555	69	some	547	112	after	317
27	their	1541	70	such	535	113	good	316
28	would	1463	71	t	516	114	two	309
29	but	1298	72	time	497	115	work	308
30	or	1292	73	her	495	116	see	307
31	an	1246	74	into	489	117	fact	306
32	from	1187	75	may	487	118	first	305
33	can	1174	76	up	481	119	therefore	304
34	more	1169	77	than	480	120	new	303
35	which	1137	78	way	477	121	example	302
36	will	1116	79	then	470	122	another	300
37	one	1100	80	even	458	123	through	295
38	there	1085	81	most	457	124	become	294
39	if	1067	82	very	446	125	power	294
40	i	1037	83	she	443	126	years	294
41	at	1025	84	its	439	127	states	292
42	all	946	85	society	438	128	take	291
43	we	935	86	how	435	129	right	290

**Continua**

## Conclusão

Rank	Palavras	Freq.	Rank	Palavras	Freq.	Rank	Palavras	Freq.
130	government	288	177	part	207	224	schools	170
131	person	287	178	having	206	225	something	170
132	system	286	179	party	206	226	optimism	168
133	where	286	180	idea	205	227	find	166
134	still	284	181	same	205	228	never	165
135	feel	281	182	end	204	229	often	165
136	get	276	183	different	201	230	things	165
137	problem	271	184	law	200	231	give	164
138	used	271	185	number	200	232	long	164
139	public	265	186	play	200	233	reason	163
140	well	264	187	boxing	198	234	united	163
141	child	263	188	guilt	198	235	certain	162
142	argument	260	189	order	198	236	come	161
143	my	260	190	act	197	237	place	161
144	quote	257	191	beef	196	238	live	160
145	case	243	192	while	195	239	old	160
146	others	243	193	europe	194	240	yet	160
147	think	239	194	know	194	241	cannot	159
148	each	238	195	national	193	242	point	159
149	why	238	196	voltaire	193	243	possible	159
150	students	237	197	lottery	192	244	support	159
151	against	236	198	parents	192	245	violence	158
152	need	234	199	hugo	189	246	due	157
153	did	233	200	today	189	247	issue	157
154	lives	233	201	every	187	248	sovereignty	157
155	britain	232	202	say	187	249	woman	157
156	caligula	232	203	better	185	250	change	155
157	without	232	204	since	185	251	means	155
158	european	228	205	education	181	252	bad	154
159	able	226	206	himself	181	253	television	154
160	before	225	207	best	179	254	don	152
161	both	225	208	day	179	255	away	151
162	between	224	209	great	179	256	political	151
163	although	222	210	back	178	257	president	151
164	school	222	211	whether	178	258	put	151
165	sex	222	212	less	177	259	rather	151
166	believe	221	213	crime	176	260	de	150
167	men	221	214	themselves	176	261	down	150
168	problems	220	215	always	175	262	help	150
169	state	219	216	suicide	175	263	once	150
170	seen	218	217	computer	174	264	single	150
171	want	217	218	social	174	265	though	150
172	important	214	219	age	173	266	disease	149
173	family	210	220	freedom	173	267	british	148
174	too	209	221	show	173	268	really	148
175	american	207	222	year	173	269	around	147
176	go	207	223	country	170	270	god	147

**APÊNDICE C - Relação dos 40 verbos lexicais mais frequentes no *corpus* de estudo.**

<b>Verbos lexicais mais frequentes no <i>corpus</i> de Br-ICLE</b>	
make	help
like	learn
work	mean
think	go
get	consider
live	keep
need	control
become	look
know	pay
see	end
take	feel
want	put
love	seem
use	face
say	come
believe	present
change	show
find	call
dream	try
give	cause
<b>Total: 40 verbos</b>	

**APÊNDICE D - Lista de partículas adverbiais formando *phrasal verbs*, conforme a frequência (ocorrências brutas e ocorrências por 100 mil).**

<b>Br-ICLE</b>			<b>LOCNESS</b>		
<b>Partícula</b>	<b>Freq. bruta de PAV formando PV</b>	<b>Freq. Normal. de PAV formando PV</b>	<b>Partícula</b>	<b>Freq. bruta de PAV formando PV</b>	<b>Freq. Normal. de PAV formando PV</b>
<b>In</b>	12	7	<b>In</b>	48	15
<b>On</b>	57	35	<b>On</b>	132	40
<b>By</b>	14	9	<b>By</b>	7	2
<b>About</b>	4	2	<b>About</b>	40	12
<b>Without</b>	-	-	<b>Out</b>	305	94
<b>Up</b>	139	86	<b>Up</b>	304	93
<b>Through</b>	7	4	<b>Over</b>	15	5
<b>Out</b>	84	52	<b>Through</b>	9	3
<b>Over</b>	5	3	<b>Without</b>	-	-
<b>Around</b>	6	4	<b>Back</b>	95	29
<b>Together</b>	18	11	<b>Away</b>	99	30
<b>Past</b>	-	-	<b>Down</b>	57	17
<b>Under</b>	1	1	<b>Around</b>	18	6
<b>Back</b>	24	15	<b>Past</b>	-	-
<b>Away</b>	23	14	<b>Under</b>	1	0
<b>Behind</b>	9	6	<b>Off</b>	61	19
<b>Off</b>	15	9	<b>Together</b>	31	10
<b>Down</b>	15	9	<b>Along</b>	4	1
<b>Forward</b>	17	11	<b>Behind</b>	4	1
<b>Along</b>	1	1	<b>Forward</b>	15	5
<b>Apart</b>	4	2	<b>Across</b>	4	1
<b>Ahead</b>	2	1	<b>Apart</b>	11	3
<b>Across</b>	1	1	<b>Round</b>	-	-
<b>Aside</b>	1	1	<b>Ahead</b>	10	3
<b>Forth</b>	-	-	<b>Forth</b>	8	2
<b>Round</b>	2	1	<b>Aside</b>	5	2
<b>Aback</b>	2	1	<b>Aback</b>	-	-
<b>TOTAL</b>	<b>463</b>	<b>287</b>		<b>1288</b>	<b>395</b>

**APÊNDICE E - Produtividade das partículas (combinação com verbos diferentes).**

<b>Br-ICLE</b>		<b>LOCNESS</b>	
<b>Partícula</b>	<b>Número de verbos</b>	<b>Partícula</b>	<b>Número de verbos</b>
Up	35	Up	68
Out	30	Out	80
On	12	Back	38
Down	11	Away	33
Off	11	Down	29
Away	10	In	23
In	8	Off	22
Back	7	On	20
Together	7	Together	17
Forward	4	Around	13
By	3	Over	7
Over	3	By	5
Around	3	Forward	5
Through	3	Through	5
Behind	2	About	5
Apart	2	Apart	5
Aback	1	Ahead	5
About	1	Aside	4
Across	1	Forth	4
Ahead	1	Across	2
Along	1	Along	2
Aside	1	Behind	2
Under	1	Under	1
Round	1	Aback	0
Forth	0	Past	0
Past	0	Round	0
Without	0	Without	0

**APÊNDICE F - Lista de *phrasal verbs* encontrados nos corpora, conforme a frequência.**

Br-ICLE			LOCNESS		
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
end up	23	14,27	go on	53	16,25
give up	22	13,65	carry out	48	14,72
go on	22	13,65	point out	41	12,57
sum up	20	12,41	take away	29	8,89
come back	14	8,68	bring up	29	8,89
point out	14	8,68	bring about	27	8,28
grow up	14	8,68	take on	27	8,28
put forward	13	8,06	end up	26	7,97
find out	13	8,06	grow up	26	7,97
keep on	13	8,06	give up	25	7,67
take away	10	6,20	find out	19	5,83
go by	9	5,58	make up	18	5,52
call out	9	5,58	set up	17	5,21
go out	9	5,58	go back	16	4,91
leave behind	8	4,96	break down	12	3,68
get together	8	4,96	cut off	12	3,68
turn on	8	4,96	come about	11	3,37
come up	7	4,34	bring in	11	3,37
wake up	6	3,72	run up	11	3,37
put together	5	3,10	carry on	11	3,37
stand up	5	3,10	go out	11	3,37
bring about	4	2,48	turn out	11	3,37
go back	4	2,48	bring back	10	3,07
pass by	4	2,48	fit in	10	3,07
turn off	4	2,48	get out	10	3,07
get out	4	2,48	set out	10	3,07
bring up	4	2,48	put forward	9	2,76



Br-ICLE			LOCNESS		
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
go through	4	2,48	back up	9	2,76
look around	3	1,86	come out	9	2,76
run away	3	1,86	start out	9	2,76
go away	3	1,86	give back	8	2,45
sit down	3	1,86	take over	8	2,45
bring out	3	1,86	build up	8	2,45
move out	3	1,86	come up	8	2,45
go up	3	1,86	hold up	8	2,45
open up	3	1,86	come back	7	2,15
move on	3	1,86	draw up	7	2,15
take on	3	1,86	pick up	7	2,15
Taken aback	2	1,24	speed up	7	2,15
go ahead	2	1,24	stand up	7	2,15
fall apart	2	1,24	take up	7	2,15
set apart	2	1,24	bring out	7	2,15
go around	2	1,24	get back	6	1,84
bring back	2	1,24	look back	6	1,84
calm down	2	1,24	break away	6	1,84
fall down	2	1,24	go away	6	1,84
step forward	2	1,24	pay off	6	1,84
show off	2	1,24	start off	6	1,84
come out	2	1,24	turn off	6	1,84
figure out	2	1,24	bring together	6	1,84
leave out	2	1,24	open up	6	1,84
play out	2	1,24	take out	6	1,84
spread out	2	1,24	throw out	6	1,84
turn out	2	1,24	throw away	5	1,53
dress up	2	1,24	turn away	5	1,53
feed up	2	1,24	slow down	5	1,53
pick up	2	1,24	go down	5	1,53

Br-ICLE			LOCNESS		
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
keep up	2	1,24	lay off	5	1,53
look up	2	1,24	come together	5	1,53
make up	2	1,24	face up	5	1,53
walk up	2	1,24	put up	5	1,53
go round	2	1,24	wake up	5	1,53
break trough	2	1,24	pass on	5	1,53
run over	2	1,24	turn on	5	1,53
take over	2	1,24	go about	4	1,23
run on	2	1,24	go ahead	4	1,23
get in	2	1,24	turn around	4	1,23
turn in	2	1,24	lead away	4	1,23
come in	2	1,24	move away	4	1,23
get in	2	1,24	put forth	4	1,23
come across	1	0,62	take off	4	1,23
come along	1	0,62	bring on	4	1,23
put aside	1	0,62	get on	4	1,23
sit around	1	0,62	hold on	4	1,23
wash away	1	0,62	move on	4	1,23
melt away	1	0,62	act out	4	1,23
break away	1	0,62	keep out	4	1,23
throw away	1	0,62	sort out	4	1,23
fade away	1	0,62	come across	3	0,92
walk away	1	0,62	get ahead	3	0,92
put away	1	0,62	come along	3	0,92
give back	1	0,62	tear apart	3	0,92
look back	1	0,62	break apart	3	0,92
change back	1	0,62	set apart	3	0,92
turn back	1	0,62	turn back	3	0,92
fall behind	1	0,62	give away	3	0,92
come by	1	0,62	rip away	3	0,92

Br-ICLE			LOCNESS		
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
bring down	1	0,62	run away	3	0,92
tear down	1	0,62	sign away	3	0,92
hand down	1	0,62	explain away	3	0,92
pin down	1	0,62	leave behind	3	0,92
keep down	1	0,62	go by	3	0,92
step down	1	0,62	get down	3	0,92
turn down	1	0,62	move forward	3	0,92
write down	1	0,62	give in	3	0,92
push forward	1	0,62	come off	3	0,92
move forward	1	0,62	put off	3	0,92
come off	1	0,62	carry through	3	0,92
drop off	1	0,62	go through	3	0,92
go off	1	0,62	catch up	3	0,92
cut off	1	0,62	get up	3	0,92
head off	1	0,62	line up	3	0,92
take off	1	0,62	look up	3	0,92
trigger off	1	0,62	put on	3	0,92
put off	1	0,62	drop out	3	0,92
top off	1	0,62	kick out	3	0,92
draw out	1	0,62	lose out	3	0,92
drop out	1	0,62	make out	3	0,92
hang out	1	0,62	miss out	3	0,92
kick out	1	0,62	play out	3	0,92
watch out	1	0,62	pull out	3	0,92
map out	1	0,62	run out	3	0,92
miss out	1	0,62	seek out	3	0,92
pick out	1	0,62	single out	3	0,92
read out	1	0,62	stand out	3	0,92
shout out	1	0,62	think out	3	0,92
stand out	1	0,62	watch out	3	0,92

Br-ICLE			LOCNESS		
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
stay out	1	0,62	come around	2	0,61
stuck out	1	0,62	rush around	2	0,61
take out	1	0,62	put aside	2	0,61
throw out	1	0,62	send back	2	0,61
walk out	1	0,62	fight back	2	0,61
work out	1	0,62	pay back	2	0,61
come together	1	0,62	strike back	2	0,61
stick together	1	0,62	seize back	2	0,61
blend together	1	0,62	take back	2	0,61
go together	1	0,62	trace back	2	0,61
hold together	1	0,62	run back	2	0,61
go under	1	0,62	tear away	2	0,61
beat up	1	0,62	get away	2	0,61
blow up	1	0,62	stay away	2	0,61
break up	1	0,62	blow away	2	0,61
built up	1	0,62	bring down	2	0,61
catch up	1	0,62	write down	2	0,61
cover up	1	0,62	die down	2	0,61
ease up	1	0,62	lay down	2	0,61
fill up	1	0,62	let down	2	0,61
hold up	1	0,62	pass down	2	0,61
take up	1	0,62	sit down	2	0,61
turn up	1	0,62	come forth	2	0,61
move up	1	0,62	come in	2	0,61
own up	1	0,62	set in	2	0,61
pop up	1	0,62	set in	2	0,61
put up	1	0,62	tune in	2	0,61
round up	1	0,62	go off	2	0,61
show up	1	0,62	shake off	2	0,61
split up	1	0,62	band together	2	0,61

Br-ICLE			LOCNESS		
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
pierce through	1	0,62	join together	2	0,61
think over	1	0,62	link together	2	0,61
bring on	1	0,62	put together	2	0,61
come on	1	0,62	tie together	2	0,61
bring on	1	0,62	turn over	2	0,61
stay on	1	0,62	look over	2	0,61
hang on	1	0,62	clean up	2	0,61
put on	1	0,62	clear up	2	0,61
Tune in	1	0,62	follow up	2	0,61
take in	1	0,62	keep up	2	0,61
give in	1	0,62	show up	2	0,61
kick in	1	0,62	spring up	2	0,61
			whip up	2	0,61
			cling on	2	0,61
			hang on	2	0,61
			back out	2	0,61
			break out	2	0,61
			cry out	2	0,61
			figure out	2	0,61
			fill out	2	0,61
			hit out	2	0,61
			lay out	2	0,61
			leave out	2	0,61
			live out	2	0,61
			move out	2	0,61
			pick out	2	0,61
			sell out	2	0,61
			walk out	2	0,61
			wipe out	2	0,61

LOCNESS					
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
work out	2	0,61	put back	1	0,31
run about	1	0,31	sit back	1	0,31
strew about	1	0,31	stand back	1	0,31
get across	1	0,31	step back	1	0,31
lie ahead	1	0,31	revert back	1	0,31
move ahead	1	0,31	talk back	1	0,31
press ahead	1	0,31	think back	1	0,31
move along	1	0,31	allow back	1	0,31
keep apart	1	0,31	ask back	1	0,31
rip apart	1	0,31	date back	1	0,31
get around	1	0,31	follow back	1	0,31
sit around	1	0,31	push back	1	0,31
run around	1	0,31	shoot back	1	0,31
drive around	1	0,31	stem back	1	0,31
stand around	1	0,31	throw back	1	0,31
sneak around	1	0,31	travel back	1	0,31
go around	1	0,31	carry away	1	0,31
roam around	1	0,31	persuade away	1	0,31
roll around	1	0,31	fade away	1	0,31
lie around	1	0,31	fritter away	1	0,31
cast aside	1	0,31	send away	1	0,31
push aside	1	0,31	plug away	1	0,31
set aside	1	0,31	put away	1	0,31
bounce back	1	0,31	push away	1	0,31
call back	1	0,31	roll away	1	0,31
cut back	1	0,31	shy away	1	0,31
fall back	1	0,31	slave away	1	0,31
feed back	1	0,31	erode away	1	0,31
slip back	1	0,31	steer away	1	0,31
hold back	1	0,31	veer away	1	0,31

LOCNESS					
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
tuck away	1	0,31	push forward	1	0,31
walk away	1	0,31	turn in	1	0,31
wear away	1	0,31	allow in	1	0,31
lag behind	1	0,31	cram in	1	0,31
come by	1	0,31	flow in	1	0,31
pass by	1	0,31	go in	1	0,31
scrape by	1	0,31	lure in	1	0,31
walk by	1	0,31	take in	1	0,31
cut down	1	0,31	sit in	1	0,31
look down	1	0,31	hold in	1	0,31
hold down	1	0,31	mix in	1	0,31
roll down	1	0,31	rain in	1	0,31
fall down	1	0,31	march in	1	0,31
set down	1	0,31	chip in	1	0,31
stand down	1	0,31	call in	1	0,31
step down	1	0,31	walk in	1	0,31
swoop down	1	0,31	step in	1	0,31
take down	1	0,31	back off	1	0,31
tone down	1	0,31	lop off	1	0,31
calm down	1	0,31	let off	1	0,31
keep down	1	0,31	pass off	1	0,31
hand down	1	0,31	push off	1	0,31
lie down	1	0,31	run off	1	0,31
knock down	1	0,31	show off	1	0,31
weight down	1	0,31	shove off	1	0,31
close down	1	0,31	spark off	1	0,31
call forth	1	0,31	shrug off	1	0,31
bring forth	1	0,31	tip off	1	0,31
go forward	1	0,31	ward off	1	0,31
step forward	1	0,31	bind together	1	0,31

LOCNESS					
<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil	<i>Phrasal verb</i>	Frequência bruta	Frequência por 100 mil
combine together	1	0,31	free up	1	0,31
gather together	1	0,31	go up	1	0,31
get together	1	0,31	hit up	1	0,31
group together	1	0,31	hook up	1	0,31
keep together	1	0,31	join up	1	0,31
melt together	1	0,31	lead up	1	0,31
hold together	1	0,31	light up	1	0,31
mix together	1	0,31	mark up	1	0,31
merge together	1	0,31	measure up	1	0,31
go under	1	0,31	meet up	1	0,31
follow through	1	0,31	own up	1	0,31
shine through	1	0,31	pair up	1	0,31
get through	1	0,31	pass up	1	0,31
win over	1	0,31	pop up	1	0,31
carry over	1	0,31	round up	1	0,31
run over	1	0,31	spark up	1	0,31
bring over	1	0,31	step up	1	0,31
beat up	1	0,31	suck up	1	0,31
boil up	1	0,31	sum up	1	0,31
bottle up	1	0,31	table up	1	0,31
buckle up	1	0,31	tie up	1	0,31
break up	1	0,31	turn up	1	0,31
check up	1	0,31	wind up	1	0,31
cover up	1	0,31	wrap up	1	0,31
cut up	1	0,31	flame on	1	0,31
divvy up	1	0,31	linger on	1	0,31
dig up	1	0,31	live on	1	0,31
drag up	1	0,31	look on	1	0,31
fill up	1	0,31	sign on	1	0,31
flare up	1	0,31	spur on	1	0,31



LOCNESS					
<i>Phrasal verb</i>	<b>Frequência bruta</b>	<b>Frequência por 100 mil</b>	<i>Phrasal verb</i>	<b>Frequência bruta</b>	<b>Frequência por 100 mil</b>
stay on	1	0,31	push out	1	0,31
walk on	1	0,31	reach out	1	0,31
bear out	1	0,31	read out	1	0,31
block out	1	0,31	rule out	1	0,31
blow out	1	0,31	rush out	1	0,31
call out	1	0,31	scout out	1	0,31
check out	1	0,31	scream out	1	0,31
chill out	1	0,31	send out	1	0,31
turn out	1	0,31	ship out	1	0,31
contract out	1	0,31	snap out	1	0,31
drive out	1	0,31	snuff out	1	0,31
gouge out	1	0,31	spread out	1	0,31
hand out	1	0,31	stamp out	1	0,31
help out	1	0,31	step out	1	0,31
hold out	1	0,31	stress out	1	0,31
lash out	1	0,31	toss out	1	0,31
pay out	1	0,31	train out	1	0,31
plan out	1	0,31	tune out	1	0,31
price out	1	0,31	weed out	1	0,31
print out	1	0,31			

**ANEXO A - Learner Profile**

**LEARNER PROFILE**

=====

===

Text code : (do not fill in)

**Essay :**

Title :

Approximate length required :        -500 words **0**                +500 words                **0**

Conditions :                                timed                **0**                untimed                **0**

Examination :                                yes                **0**                no                **0**

Reference tools :                                yes                **0**                no                **0**

What reference tools ?

Bilingual dictionary :

English monolingual dictionary :

Grammar :

Other(s) :

=====

===

**Surname :**                                        **First names :**

**Age :**                                        Male                **0**                                Female                **0**

**Nationality :**

Native language :

Father's mother tongue :

Mother's mother tongue :

Language(s) spoken at home : (if more than one, please give the average % use of each)

**Education :**

Primary school - medium of instruction :

Secondary school - medium of instruction :

Current studies :

Current year of study :

Institution :

Medium of instruction :

English only **0**

Other language(s) (specify) **0**

Both **0**

=====

===

Years of English at school :

Years of English at university :

**Stay in an English-speaking country :**

Where ?

When ?

How long ?

=====

===

**Other foreign languages in decreasing order of proficiency :**

=====

===

**I hereby give permission for my essay to be used for research purposes.**

Date : .....

Signature : ...

## ANEXO B - Circunstâncias de produção do LOCNESS

### LOCNESS description

LOCNESS contains

149,574 words of argumentative essays written by American university students

18,826 words of literary-mixed essays written by American university students

95,695 words of argumentative and literary essays written by British university students

60,209 words of British A-level argumentative essays.

### BRITISH ESSAYS: University students

#### I. brsur1.cor - 59,568 words

March 1991

+ 500 words

Exams

1) **15 essays** on 'French Intellectual tradition' - codes 1-15 - 41,439 words

-**literary** (Camus, Sartre, 'Oreste', 'La Chute', 'Caligula', 'Les Justes' etc.)

2) **18 essays** on 'French Society and Institutions' - codes 16-33 - 18,129 words

-**expository - historical** (French higher education, French constitution, unionism in France, presidentialism, demography)

#### II. brsur2.cor - 17,108 words

+500 words

Exams

**24 essays** on 'French Intellectual Tradition'

-**literary** (Most of them: Voltaire: 'Candide' - philosophical optimism, Rousseau - 'Discours sur l'origine de l'inégalité' nr 8 and 20, Montesquieu 'De l'esprit de loi' nr 14)

-nr7 - 17th century scientific revolution

#### III. brsur3.cor - 19,019 words

**33 argumentative essays** on 'A single Europe: A loss of sovereignty for Britain'

500 words

Not exams

Not rigidly timed - +/- 1 hour

No reference tools used

#### IV A levels

**60,209 words of British A-level argumentative essays.**

Transport, boxing, parliamentary system, fox hunting

### AMERICAN ARGUMENTATIVE ESSAYS

**argumentative essays: 149,574 words (usarg.txt)**

**Marquette University (codes: ICLE-US-MRQ-0001.1-46.1)**

**54,285 words**

**46 essays**

March 1995

Untimed essays

+ 500 words

Argumentative

Reference tools used

All students are fully English native speakers. (both parents with English mother tongue)

Age: from 18 to 21 (+ 1 of 30, 1 of 31 and 1 of 40)

Topics:

- controversy in the classroom
- capital punishment
- does affirmative action work?
- yoga
- nuclear power
- values and consequences of school interaction
- pride or segregation
- surrogate motherhood
- can we afford wellness
- prozac: the wonder drug
- homosexuality
- animal testing
- prayer in schools
- praying for a miracle
- sex equality
- teenagers
- aids
- orphanages
- profit: good or evil
- freedom of the press
- sex in schools
- welfare reforms needs a return to family values
- the cost of grass
- abortion
- ethics
- would anyone care for a drink
- cheating in colleges
- O.J Simpson
- suicide

**Indiana University at Indianapolis (codes: ICLE-US-IND-0001.1-28.1)**

**13,454 words**

**28 essays**

March 1995

Timed essays ??

Argumentative

Reference tools used ??

All students are fully English native speakers .

Age: from 22 to 48

Topics:

- Money is the root of all evil
- Crime does not pay
- A man / woman's financial reward should be commensurate with their contribution to the society in which they live
- Feminists have done more harm to the cause of women than good

**Presbyterian College, South Carolina (codes: ICLE-US-PRB-0034.2-39.2)**

**12,447 words**

**1) 6 lengthy (+ 500) essays**

April 1995

Untimed essays

Argumentative (cd. note below)

Reference tools used (library documentation - books, journals)

All students are fully English native speakers.

Age: from 20 to 22

Topics:

- Adolescent suicide
- Water pollution
- Legalization of marijuana
- Homelessness
- The welfare system
- Divorce

NB: these issues are discussed from the point of view of the structural functionalist, conflict, exchange, symbolic interactionist theories

**University of South Carolina**

**I. usscu1.cor - 5,710 words**

**6 essays**

April 1995

Untimed essays (??)

+ 500 words

Argumentative

Reference tools used (??) - 'works cited' section for nr 2?, 4 and 5 (library books, articles)

All students are fully English native speakers except for nr3 .NL not specified, mother's and father's mother tongue = Yorube, nationality: Nigerian, home: 99% English, 1% Yorube)

Age: from 17 to 20 ( + 1 of 34 - nr5 - and 1 of 66 - nr 7)

Topic:

- gender roles in our society (all but nr 1 - 'football')

**II. usscu2.cor - 18,630 words**

**17 essays**

November/December 95

Timed essays?

+ 500 words

Argumentative

Reference tools used? (ref. for nr 11 and 12)

All students are fully English native speakers

Age: 18-19 (most), 28

Topics:

- The Confederate Flag
- Rules and regulations
- Death penalty
- Legalization of marijuana
- Teachers deserve recognition and reward
- Salary caps
- Sex in the Media
- Euthanasia
- Gender roles, feminism, etc.
- US government

### **III. ususc3.cor - 15,815 words**

#### **13 essays**

November/December 1995

Untimed essays

+ 500 words

Argumentative

Reference tools used - all? (books, magazines)

All students are fully English native speakers

Age: 17-19 (no age for nr4)

Topics:

- Premarital sex
- Football
- Drinking age
- Talk shows
- Professors that don't speak English shouldn't teach English speaking students
- Violence on television
- Gun control
- Portrayal of women in fashion magazines
- Recycling
- The wild card and its effect on Baseball
- Journalists should not reveal their sources
- Welfare

### **IV. usscu4.cor - 12,730 words**

#### **17 essays**

November 1995

Untimed essays

Argumentative

+ 500 words (except for nr17)

No reference tools used (except for nr 5, 14 and 15)

All students are fully English native speakers

Age: 17-21, nr1=39, nr12=57, nr16=41

Topics:

- Women in combat
- Rules

- Sink or Swim
- Early are drinking
- Should the Browns stay in Cleveland?
- Curfew
- Government support for the Arts
- Abortion
- Stereotyping the colours pink and blue
- Capital punishment
- The media's right to know
- Emerging women
- Legalization of marijuana
- Bookbanning in America
- Frivolous lawsuits

### **University of Michigan (codes: ICLE-US-MICH-0001.1-45.1)**

#### **43 essays (16,502 words)**

Timed essays

- 500 words

Argumentative

No reference tools used

All students are fully English native speakers (except for - nr23: NL=Engl, father=Spanish, mother=Spanish and Engl, home=Spanish 30-50%; nr24: NL=Engl/French, father=Engl, mother=Fr, home=Fr 50%; nr34: NL=Engl, father=Engl/Sp, mother=Engl, home=Sp 25%)  
Age: 19-23

Topic: Great inventions and discoveries of 20th century and their impact on people's lives (one per interview - computer, television, etc.)

### **AMERICAN LITERARY-MIXED ESSAYS**

**literary-mixed: 18,826 words (usmixed.txt)**

**Presbyterian College, South Carolina**

#### **I. usprb1.cor - 18,826 words**

1) **16 essays** - codes 1 to 17 (no nr 7-lost) (8 students produced 2 essays each) (**9,296 words**)

April 1995

Timed essays (exams)

Mixed: about literature but most are rather argumentative (cf. instruction sheet, 'I will prove', 'I firmly believe')

No reference tools used

7 students are fully English native speakers

1 : mother's mother tongue: Greek

Language spoken at home: English 80 % , Greek 20 %

Age: from 19 to 22

Topic:

-An Aspect of Studying Ethnic American Literature

2) **8 essays** codes 18-25 (**4,436 words**)

February 1995

Timed essays

- 500 words



Literary (NB: 'persuasive essay' on instruction sheet attached to the essays but literary!!!)

Reference tools used

All students are fully English native speakers.

Age: from 18 to 21

Topics:

- Who is Hamlet ?

- What is an appropriate label for Voltaire's Candide ?

3) **32 essays** codes 26-33 ( **5,094 words**) (8 students produced 4 essays each)

April 1995

Timed essays (exams - social psychology tests)

- 500 words

Not really argumentative, answers to 5 exam questions (cf. topics below)

No reference tools used.

All students are fully English native speakers ( exception: one student : Spanish speaking mother but language spoken at home = English)

Age: from 18 to 21

Topic:

-Aspects of Social Psychology: 1 § on homicide, 1 § on racial prejudice, 1 § on formation of sound personal relationships, conformity, etc. - not very incoherent and in some §, reference to what the client should do ('I would advise our client to