



**Universidade do Estado do Rio de Janeiro**  
Centro de Educação e Humanidades  
Instituto de Letras

Maria Izabel de Andrade Almeida

**Prosa argumentativa em língua inglesa: um estudo contrastivo sobre  
advérbios em corpora digitais**

Rio de Janeiro  
2010

Maria Izabel de Andrade Almeida

**Prosa argumentativa em língua inglesa: um estudo contrastivo sobre advérbios em corpora digitais**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Letras, da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Tania M. G. Shepherd

Rio de Janeiro  
2010

CATALOGAÇÃO NA FONTE  
UERJ/REDE SIRIUS/CEHB

A447 Almeida, Maria Izabel de Andrade.  
Prosa argumentativa em língua inglesa: um estudo contrastivo sobre advérbios em corpora digitais / Maria Izabel de Andrade Almeida. – 2010.  
152 f.

Orientadora: Tania Maria Granja Shepherd.  
Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro, Instituto de Letras.

1. Língua inglesa – Advérbios – Teses. 2. Língua inglesa – Estudo e ensino – Falantes de português – Teses. 3. Língua inglesa – Aquisição – Teses. I. Shepherd, Tania Maria Granja. II. Universidade do Estado do Rio de Janeiro. Instituto de Letras. III. Título.

CDU 802.0-27

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação

---

Assinatura

---

Data

Maria Izabel de Andrade Almeida

**Prosa argumentativa em língua inglesa: um estudo contrastivo sobre  
advérbios em corpora digitais**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Letras, da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Aprovada em 30 de março de 2010.

Banca Examinadora:

---

Profa. Dra. Tania M. G. Shepherd (Orientadora)  
Instituto de Letras da UERJ

---

Profa. Dra. Anna Elizabeth Balocco  
Instituto de Letras da UERJ

---

Profa. Dra. Solange Coelho Vereza  
Instituto de Letras da UFF

Rio de Janeiro  
2010

## **DEDICATÓRIA**

Dedico esta dissertação a minha mãe, a oitava maravilha do mundo, cujo exemplo de luta e perseverança em meio às adversidades me motiva a continuar a caminhada e cuja sublime existência faz cada passo valer a pena. A ela, todo o meu amor, respeito e gratidão.

## AGRADECIMENTOS

A Deus por ser o Grande Mestre por trás das grandes obras e por ser a Verdadeira Fonte de bênçãos em minha vida.

A minha família por ser a base forte que me dá sustento.

A Tania Shepherd, por ter acreditado em mim quando eu mesma já não mais acreditava. Pela paciência e dedicação com que me orientou ao longo de minha vida acadêmica e principalmente pela amizade e pelo carinho que demonstrou tantas e tantas vezes os quais seriam impossíveis de descrever em poucas linhas de agradecimento. A ela também devoto meu respeito e minha gratidão.

Aos amigos que conheci na graduação, uns mais distantes, outros mais próximos, mas sempre especiais: Elierme Mantaia, pelos risos e lágrimas compartilhados, pelos *capuccinos* na Livraria da Travessa e, acima de tudo, por me presentear com sua eterna amizade; Patrícia Raiz, pela franqueza das palavras, por aturar meus dias de mau-humor e pela paixão compartilhada pelo cinema e os seriados de tv; Aline Mota, pela doçura das palavras e pelos abraços afetuosos e Maria Glória, pelas experiências de vida que compartilhou com tanto alto-astral.

As amigas mais chegadas que irmãs que a Vida me deu: Daiane, Vânia e Samarina, por compreenderem a minha ausência em alguns momentos e ainda assim me amarem.

Aos mestres: Gisele de Carvalho e Maria Alice Antunes pelas palavras de incentivo durante a difícil transição da graduação para o mestrado; Anna Elisabeth Balocco pela atenção dispensada quando eu apenas começava a vida acadêmica como bolsista do Pró-Iniciar e Ricardo Joseh Lima que, graças aos debates “calorosos” durante as aulas, me despertou a paixão pela Linguística.

A todos os demais mestres que tive oportunidade de conhecer na Uerj ao longo da graduação e da pós-graduação que, de alguma forma, contribuíram para minha formação acadêmica.

Aos funcionários da secretaria da Pós-graduação em Letras, em especial, a Cláudia, pela imensa boa vontade e principalmente pela paciência.

A FAPERJ, por ter financiado este último ano de Mestrado.

## RESUMO

ALMEIDA, Maria Izabel de Andrade. *Prosa argumentativa em língua inglesa: um estudo contrastivo sobre advérbios em corpora digitais*. 2010. 152 f. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010.

Esta pesquisa tem como objetivo principal investigar como aprendizes brasileiros de língua inglesa usam advérbios com terminação em *-ly* no inglês escrito, e comparar ao uso que deles fazem os falantes de inglês como língua materna. Para tanto, o trabalho encontra suporte teórico e metodológico na Linguística de *Corpus* e fundamenta-se na área chamada de pesquisa sobre *corpora* de aprendizes, que se ocupa da coleta e armazenagem de dados linguísticos de sujeitos aprendizes de uma língua estrangeira, para a formação de um *corpus* que possa ser utilizado para fins descritivos e pedagógicos. Esta área objetiva identificar em que aspectos os aprendizes diferem ou se assemelham aos falantes nativos. Os *corpora* empregados na pesquisa são o *corpus* de estudo (Br-ICLE), contendo inglês escrito por brasileiros, compilado de acordo com o projeto ICLE (*International Corpus of Learner English*) e dois *corpora* de referência (LOCNESS e BAWE), contendo inglês escrito por falantes de inglês como língua materna. Os resultados indicam que os alunos brasileiros usam, em demasia, as categorias de advérbios que indicam veracidade, realidade e intensidade, em relação ao uso que deles fazem os falantes nativos, além de usarem esses advérbios de forma distinta. Os resultados sugerem que, além das diferenças apresentadas em termos de frequência (seja pelo sobreuso ou subuso dos advérbios), os aprendizes apresentavam combinações “errôneas”, ou em termos de colocados ou em termos de prosódia semântica. E finalmente a pesquisa revela que a preferência dos aprendizes por advérbios que exprimem veracidade, realidade e intensidade cria a impressão de um discurso muito assertivo. Conclui-se que as diferenças encontradas podem estar ligadas a fatores como o tamanho dos *corpora*, a influência da língua materna dos aprendizes, a internalização dos elementos linguísticos necessários para a produção de um texto em língua estrangeira, a falta de fluência dos aprendizes e o contexto de sala de aula nas universidades.

Palavras-chave: Advérbio. Linguística de *Corpus*. *Corpora* de Aprendizes. Análise Contrastiva.



## ABSTRACT

This research investigates how Brazilian learners of English use adverbs ending in -ly in written English and compares their use to that of speakers of English as a mother tongue. To this end, the work resorts to Corpus Linguistics as both theoretical and methodological support. The research is based on the area called Learner *Corpora* Research, which deals with the collection, storage and analysis of linguistic data produced by learners of a foreign language, which can then be used for descriptive and teaching purposes. This area aims to identify ways in which learners' use of the foreign language is different or similar to that of native speakers. The *data* used in this research are the *corpus* of study (Br-ICLE), containing written English produced by Brazilian learners, built according to the ICLE project (International *Corpus* of Learner English), as well as two reference *corpora* (Locness and BAWE) containing written English produced by speakers of English as a mother tongue. The results indicate that Brazilian learners overuse the categories of adverbs that indicate truth, reality and intensity in comparison to the use made by native speakers, furthermore they use these adverbs in different ways. The results also suggest that, given the differences in frequency (either by overuse or underuse of adverbs), the learners tend to "misuse" combinations in terms of collocates or in terms of semantic prosody. And finally, the research reveals that the preference of learners for adverbs expressing truth, reality and intensity creates the impression of very assertive voices. We conclude that these differences may be related to factors such as the size of the *corpus*, the influence of the learners' mother tongue, the internalization of linguistic elements needed to produce a text in a foreign language or even the lack of fluency of the learners and the classroom context in the universities.

Keywords: Adverbs. *Corpus* Linguistics. Learner *Corpora*. Contrastive Analysis.

## SUMÁRIO

<b>INTRODUÇÃO</b>	<b>10</b>
Justificativa	10
O objeto de estudo	17
<b>1. REVISÃO DA LITERATURA: LINGÜÍSTICA DE CORPUS</b>	<b>19</b>
1.1. Histórico	19
1.2. Linguística de Corpus: fundamentação teórica e considerações sobre métodos de investigação linguística	21
1.3. Linguística de Corpus: abordagem empírica ou racionalista?	25
1.4. Definição de corpus	27
1.4.1. <u>Conteúdo de um corpus</u>	28
1.4.2. <u>Representatividade de um corpus</u>	29
1.4.3. <u>Autenticidade de um corpus</u>	30
1.5. Abordagens ao estudo do corpus	31
1.5.1. <u>Análise quantitativa ou qualitativa?</u>	31
1.5.2. <u>Abordagem baseada em corpus ou abordagem direcionada pelo corpus?</u>	33
<b>2. REVISÃO DA LITERATURA: ADVÉRBIOS</b>	<b>36</b>
2.1. As gramáticas de Língua Inglesa	36
2.1.1. <u>A gramática sistêmico-funcional</u>	40
2.2. Os advérbios	40
2.2.1. <u>Metafunção experiencial</u>	41
2.2.2. <u>Metafunção interpessoal</u>	43
2.2.3. <u>Metafunção textual</u>	47
<b>3. METODOLOGIA</b>	<b>50</b>
3.1. Materiais: os corpora	50
3.1.1. <u>O corpus de estudo: o que é e como está composto.</u>	50
3.1.2. <u>Os corpora de referência: o que são e como estão compostos</u>	52
3.2. Métodos	53
3.2.1. <u>O ferramental de análise: o programa WordSmith Tools</u>	53
3.2.2. <u>Trabalhando com proporções</u>	55
3.3. Os advérbios analisados	56
3.4. Métodos de análise	61
3.5. Conclusão	63
<b>4. ANÁLISE DOS DADOS</b>	<b>64</b>
4.1. Listando os advérbios da pesquisa	64
4.2. Analisando os advérbios-chave	65
4.3. Really	68
4.3.1. <u>Função</u>	68
4.3.2. <u>Perfil léxico-gramatical</u>	71
4.3.3. <u>Conclusões sobre o advérbio really</u>	75
4.4. Simply	78
4.4.1. <u>Função</u>	78
4.4.2. <u>Perfil léxico-gramatical</u>	79
4.4.3. <u>Conclusões sobre o advérbio simply</u>	83

<b>4.5. Actually</b> .....	<b>84</b>
4.5.1. <u>Função</u> .....	84
4.5.2. <u>Perfil léxico-gramatical</u> .....	86
4.5.3. <u>Conclusões sobre o advérbio actually</u> .....	90
<b>4.6. Really X Actually</b> .....	<b>91</b>
4.6.1. <u>Conclusões sobre really X actually</u> .....	93
<b>4.7. Consequently</b> .....	<b>93</b>
4.7.1. <u>Função</u> .....	93
4.7.2. <u>Perfil léxico-gramatical</u> .....	94
4.7.3. <u>Conclusões sobre o advérbio consequently</u> .....	98
<b>4.8. Especially</b> .....	<b>99</b>
4.8.1. <u>Função</u> .....	99
4.8.2. <u>Perfil léxico-gramatical</u> .....	100
4.8.3. <u>Conclusões sobre o advérbio especially</u> .....	104
<b>4.9. Conclusões sobre o capítulo de análise</b> .....	<b>104</b>
<b>5. CONSIDERAÇÕES FINAIS</b> .....	<b>106</b>
5.1. <u>Percurso da dissertação e limitações</u> .....	106
5.2. <u>Aplicações pedagógicas e futuras pesquisas</u> .....	110
<b>REFERÊNCIAS</b> .....	<b>113</b>
<b>ANEXO 1:</b> descrição dos corpora da pesquisa .....	116
<b>ANEXO 2:</b> lista completa dos colocados e agrupamentos lexicais dos advérbios da pesquisa no corpus BAWE. ....	132

## INTRODUÇÃO

### Justificativa

Esta dissertação de mestrado pretende investigar o uso de formas adverbiais terminadas em *-ly* em três *corpora* digitalizados<sup>1</sup>, formados de textos argumentativos em língua inglesa, em termos de quantidade, extensão e função. A produção escrita a ser analisada advém de três grupos, um deles formado exclusivamente de aprendizes de inglês, falantes de português brasileiro como língua materna, o segundo composto por pré-universitários e universitários cuja língua materna é o inglês e por último, a escrita produzida por graduandos e pós-graduados, usuários de inglês como primeira língua. Os *corpora* a serem usados são respectivamente o *Brazilian International Corpus of Learner English* (doravante Br-ICLE<sup>2</sup>), o *Louvain Corpus of Native English Essays* (doravante LOCNESS<sup>3</sup>) e por fim o *corpus The British Academic Written English* (doravante BAWE<sup>4</sup>)

O estudo dá continuação ao meu trabalho como Bolsista de Iniciação Científica do CNPq<sup>5</sup> de agosto de 2006 a julho de 2007. Durante esse período estudei os verbos modais mais frequentes na escrita de alunos brasileiros, aprendizes de inglês, e falantes nativos<sup>6</sup>. Naquele trabalho, apesar de ambos os grupos utilizarem mais frequentemente os modais *can* e *would*, notou-se que esses mesmos modais eram empregados de formas distintas em textos acadêmicos por cada um dos grupos estudados. Observou-se que, enquanto os aprendizes utilizam o modal *can* de forma deôntica, os falantes nativos utilizam de forma epistêmica. Variação semelhante também foi notada quanto ao uso de *would*. Se por um lado, os aprendizes utilizavam *would* como recurso para designar vontade, os falantes nativos utilizam o modal como recurso para probabilizar.

<sup>1</sup> Br-ICLE, LOCNESS e BAWE não são softwares, mas bases de dados de textos coletados e armazenados em CD-ROM, sendo disponíveis gratuitamente como o Br-ICLE ou para compra como o LOCNESS.

<sup>2</sup> *Corpus* Internacional de Aprendizes Brasileiros de Inglês.

<sup>3</sup> *Corpus* de Ensaios de Inglês como Língua materna da Universidade de Louvain

<sup>4</sup> Escrita Acadêmica do Inglês Britânico

<sup>5</sup> Projetos de IC e de Mestrado, financiado pela FAPERJ, são ligados ao GELC (Grupo de Estudos em Linguística de *Corpus*), grupo de pesquisa cadastrado no CNPq.

<sup>6</sup> A palavra “nativo” foi adotada apenas para se referir aos falantes de inglês como língua materna, sem nenhum julgamento de valor, ou seja, isto não quer dizer que nativos sejam melhores e não nativos sejam piores usuários de uma língua.

O presente trabalho, bem como meu estudo de IC, se encaixam na área de investigação chamada “*learner corpus research*”<sup>7</sup>, uma área híbrida que Granger (1998) já definiu como conjugando quatro disciplinas: a Linguística de *Corpus* (a ser descrita mais adiante), a Teoria Linguística, Aquisição da Linguagem e Ensino de Língua Estrangeira.

Cada uma dessas disciplinas contribui para a otimização e a exploração dos dados contidos em um *corpus*. Por exemplo, o pesquisador precisa ser um especialista em Linguística de *Corpus*, precisa ter familiaridade com a criação de um *corpus*, com sua anotação, com a extração de dados e com a sua análise, assim como também precisa saber lidar com dados estatísticos do *corpus*. O pesquisador precisa estar atento, todavia, para o fato de que os métodos e ferramentas disponíveis para compilação, anotação e análise de *corpora* foram desenvolvidos com base em dados de falantes nativos e portanto, não são totalmente adaptados para processarem a linguagem do aprendiz (Granger 2009). Isso se explica porque há uma taxa de variação ou erros de toda natureza em *corpora* de aprendizes (grafia, concordância verbal, erros de léxico, etc.) que, por sua vez, afeta toda sorte de cálculo, (Granger 2009), medidas essas a serem discutidas no desenvolver desta dissertação.

O conhecimento sobre teoria linguística também é necessário para criar bases para uma análise linguística dos dados. Por exemplo, o analista deve decidir se vai usar categorias gramaticais tradicionais para etiquetar um *corpus* de aprendiz (substantivo, verbo, etc) ou utilizar uma abordagem funcional para analisar a linguagem em uso (Meyer 2002:6 *apud* Granger 2009).

Compreender o processo de aquisição de uma segunda língua também é importante para a interpretação dos dados. Fatores sociais, cognitivos e psicológicos, que desempenham um papel fundamental no ensino de línguas são extensivamente estudados na área de aquisição de uma segunda língua e a familiarização com seus achados permite ao analista uma correta interpretação dos seus resultados (Granger 2009).

Por fim, um bom conhecimento sobre o ensino de uma língua estrangeira é essencial, se a pesquisa linguística tem fins pedagógicos. Isso é particularmente importante no estudo do impacto de fatores contextuais já que esses vão determinar se e qual extensão dos resultados obtidos na pesquisa com *corpus* de aprendiz podem ser integrados ao ensino (Granger 2009).

Para entendermos a interdisciplinaridade da análise de *corpora* de aprendizes, podemos apresentar o diagrama de Granger (2009:15). A pesquisadora coloca a Pesquisa com *corpus* de

---

<sup>7</sup> Pesquisa com *Corpus* de Aprendiz.

aprendiz partindo de várias áreas, mas ao mesmo tempo fornecendo dados as mesmas áreas, como visto abaixo:

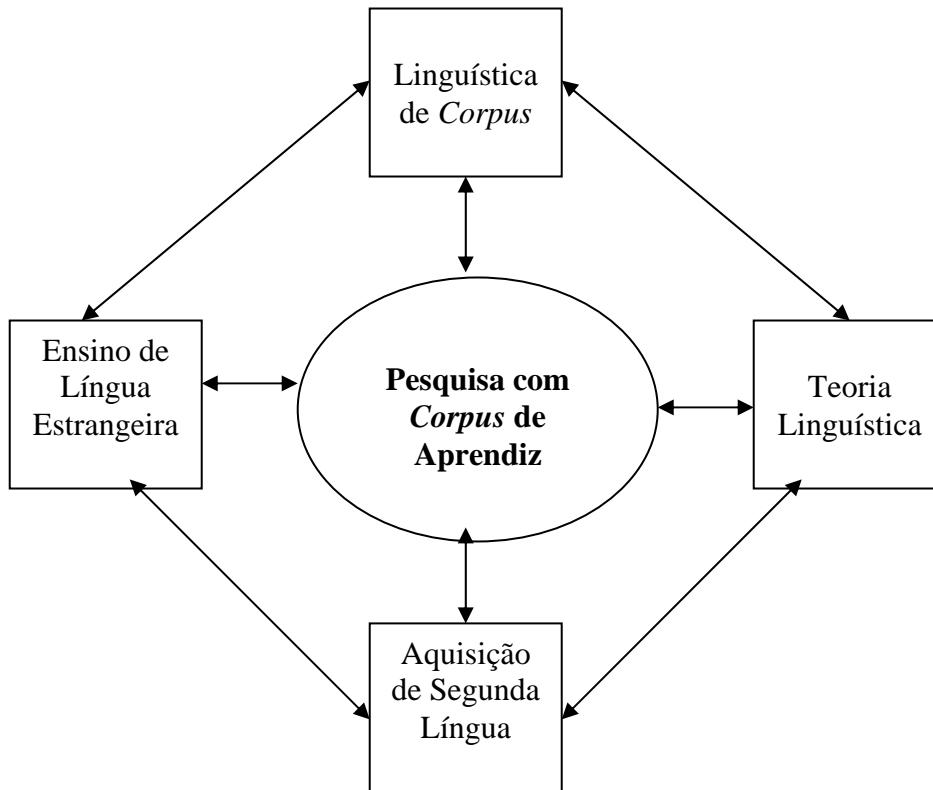


Figura 1:Componentes da pesquisa com *corpus* de aprendiz (adaptado de Granger 2009:15)

A compilação de *corpora* digitais vem se concentrando em *corpora* de textos naturais, escritos ou falados, produzidos ou por escritores publicados (jornalistas, romancistas, historiadores, etc.) ou aqueles produzidos e transcritos também em ocasiões informais. Desde os anos 90, entretanto, há crescente atenção voltada para textos produzidos por aprendizes de línguas.

O estudo de *corpora* de aprendizes é na sua essência um estudo contrastivo: ou se contrasta a produção do aprendiz com a produção do 'não aprendiz', ou se contrastam produções de vários grupos de aprendizes. Mais recentemente, há também estudos diacrônicos de linguagem de aprendiz, mas estes ainda são incipientes<sup>8</sup>.

<sup>8</sup> Shepherd, 2009, comunicação pessoal

O estudo da linguagem de aprendizes, em um passado não tão remoto, era feito através de uma abordagem chamada Análise Contrastiva cuja premissa subjacente era que quanto mais próximas duas línguas fossem em termos de semelhanças estruturais, mais fácil seria o processo de aprendizagem de uma delas como língua estrangeira. Diferenças, por outro lado, induziriam a uma transferência ou interferência negativas. Isso implicava dizer que os professores otimizavam seus esforços para se concentrarem em áreas da gramática da língua materna do aprendiz que mostrariam diferenças marcadas na gramática da língua estrangeira.

Ou seja, a Análise Contrastiva investigava a produção de aprendizes baseada na comparação entre um número restrito de exemplos de alguns poucos aprendizes e exemplos “corretos” da língua fornecidos pela intuição do próprio analista/pesquisador. A Análise Contrastiva sempre visava a extrair o que se denominava **erro**, sem qualquer preocupação de elencar acertos. Todavia, com o advento dos computadores e a possibilidade de compilação de grandes *corpora* eletrônicos, as investigações da produção de aprendizes aumentaram em número, extensão e foco. Com isso surge uma nova área conhecida como Estudos Contrastivos de Interlíngua (CIA) (Granger *apud* Barlow 2005:342) cujas comparações passaram a ser entre aquilo que é ou não nativo, em termos do que é usado ou não; do que é usado em excesso, e do que é subusado. Estudos que levam em consideração a comparação entre nativos e não nativos revelam naturezas da interlíngua cujo foco são os aspectos não nativos da fala e da escrita dos aprendizes e cujos erros aparentes depois do contraste se dão em estágios diferentes na linguagem do aprendiz. Em contrapartida, comparações entre diferentes produções de não nativos podem realçar aspectos do uso e do desenvolvimento da linguagem compartilhada por aprendizes com diferentes níveis de conhecimento da língua. Nesse caso, é possível verificar até que ponto essa diferenças podem estar relacionadas com influências da língua materna (Granger *apud* Barlow 2005:342). Ou seja, pode-se dizer também que nos Estudos Contrastivos de Interlíngua há também espaço para os acertos.

A utilização de *corpora* de aprendizes é, portanto, de grande relevância para o ensino de línguas. Do ponto de vista pedagógico, o interesse em comparar aprendizes e nativos é obvio, uma vez que eles ajudam os professores a identificarem o léxico, a gramática e as características do discurso que diferenciam a produção de aprendizes da norma alvo e são, portanto, uma rica fonte de dados para fins pedagógicos (Granger 2008). O acesso e a exploração desses *corpora* possibilita alguns questionamentos da seguinte ordem:

“a) quais características linguísticas da língua alvo são empregadas com mais (sobretudo) ou menos (substituto) frequência em comparação com falantes nativos?

b) qual é a extensão da influência da língua nativa (transferência) na produção dos aprendizes?

c) em que áreas os aprendizes tendem a usar estratégias para evitar utilizar recursos linguísticos, deixando dessa forma, de explorar a fundo o potencial da língua alvo?

d) em que áreas os aprendizes tendem a demonstrar desempenho nativo ou não nativo?

e) quais são as áreas nas quais os aprendizes de um dado país parecem necessitar de mais ajuda para desenvolver a sua produção na língua alvo? (Leech *apud* Berber Sardinha 2004:266)”.

A observação da linguagem tal qual manifestada em instâncias naturais, coletadas de acordo com princípios metodológicos bem estruturados, que privilegiem níveis distintos de proficiência na língua alvo, pode também ajudar a entender como se dá o processo de aprendizagem da língua e os sistemas intermediários (também chamado de interlíngua) que os aprendizes adquirem ao longo desse processo.

Desta forma, esta dissertação tem como fundamentação teórica e metodológica a Linguística de *Corpus*, área da Linguística que privilegia a análise de dados empíricos para a descrição da linguagem. Essa área do conhecimento “ocupa-se da coleta e da 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” (Berber Sardinha, 2004: 3).

A Linguística de *Corpus*, voltada à investigação de *corpora* de aprendizes, privilegia, em suma, a linguagem dos alunos tal qual manifestada em seus ambientes naturais, ou seja, no emprego, no entretenimento, na escola (na sala de aula, por exemplo, com exceção do laboratório e de situações de teste) e no mundo real em geral (Berber Sardinha 2004:269).

O Linguista de *Corpus* objetiva a descrição de uma determinada língua ou variante linguística, tal como ela é utilizada por um determinado grupo de enunciadores<sup>9</sup> em um evento comunicativo específico a partir da análise de um *corpus* digitalizado e lido por computador, especialmente compilado.

---

<sup>9</sup> os termos enunciadores e co-enunciadores são usados nesta dissertação como sinônimos de falante/ouvinte e de emissor/receptor e não têm relação com a semântica da enunciação.



Os critérios para compilação de *corpora* digitalizados, para qualquer variedade de linguagem (linguagem de ‘nativos’ ou aprendizes) devem seguir critérios semelhantes que tem a ver com o objeto pesquisado e as perguntas de pesquisa.(Nesselhauf, 2004 e Barlow, 2005). Acrescido a isso, deve-se se considerar o nível de escolaridade dos aprendizes, a sua língua materna, a forma de aquisição da segunda língua e os critérios e condições adotados na produção dos textos do *corpus*. Dessa forma, conforme observa Nesselhauf, os tipos de textos que compõem os *corpora* de aprendizes atuais normalmente são ensaios ou textos acadêmicos, produzidos ou sob condições de exame ou como tarefa de casa. Isto porque quanto mais fechadas são as condições de produção de um *corpus*, ou seja, quanto mais específica a linguagem, maior a padronização e conseqüentemente menor a variação do léxico, da gramática, do discurso, entre outros elementos (McEnery e Wilson 1996).

Portanto, não coincidentemente, a maioria dos estudos realizados com *corpora* de aprendizes tem como base os *subcorpora* do *International Corpus of Learner English* (doravante ICLE<sup>10</sup>), da Universidade Católica de Louvain na Bélgica, cujo foco é a análise de vários aspectos de textos argumentativos de aprendizes de inglês em nível avançado. A partir do ICLE, Tankó (2004) estudou o uso de conectivos adverbiais em ensaios argumentativos de universitários húngaros. Também a partir do ICLE, Römer (2004) comparou o uso de modais auxiliares, sua distribuição, significado e contexto, entre o *corpus* do inglês britânico falado e os livros didáticos alemães. Wei-yu Chen (2006) comparou o uso de conjuntivos adverbiais usados na escrita acadêmica por aprendizes taiwaneses com o uso em periódicos da TESOL. Neff et al (2006) estudaram o uso de verbos modais e verbos de citação como expressão de atitude em textos acadêmicos produzidos por falantes de inglês como língua estrangeira de cinco nacionalidades (espanhol, holandês, francês e alemão) em contraste com a escrita de universitários americanos. McEnery e Kifle (2000) examinaram o uso de modalidade epistêmica em ensaios argumentativos de aprendizes de inglês. Neff et al (2004) contrastaram textos argumentativos de universitários espanhóis com a escrita de universitários americanos e escritores profissionais (críticos de jornais) para identificar possíveis ‘transferências’ da língua materna para a segunda língua. Ringbom (1998) explorou a frequência de vocabulário do ICLE a partir de sete *subcorpora* de inglês de nacionalidades diferentes (francês, espanhol, finlandês,

---

<sup>10</sup> *Corpus* Internacional de Aprendizes de Inglês. Contém textos argumentativos de aprendizes de inglês representando quatorze nacionalidades diferentes.

finlandês-sueco, sueco, holandês e alemão) comparando com o *corpus* LOCNESS. Bolton, Nelson e Hung (2004) contrastaram o uso de conectivos na escrita acadêmica de universitários de Hong Kong e da Grã Bretanha, também usando o LOCNESS. Aijmer (2002) comparou a frequência de alguns verbos modais na escrita de inglês como primeira língua em relação aos universitários suecos, aprendizes de inglês, em nível avançado. Guo (2006) examinou o uso de modais na fraseologia de aprendizes chineses em contraste com o *corpus* LOCNESS, composto de *subcorpora* contendo *essays* argumentativos e literários escritos por universitários americanos, vestibulandos ingleses e universitários ingleses, compilados pela própria Universidade Católica de Louvain e usado muitas vezes como *corpus* de referencia.

Em termos de aplicabilidade dos resultados dos trabalhos sobre linguagem de aprendiz a partir de *corpora* digitalizados, Gilquin, et al (2007), avaliaram a importância dos *corpora* de aprendizes de inglês, seja como língua estrangeira ou segunda língua, na confecção de materiais pedagógicos próprios para aprendizes. Essa aplicabilidade ainda é discutível, dada à escassez de trabalhos desse tipo, pois o impacto de estudos de *corpora* de aprendizes ainda é pequeno na produção de material pedagógico<sup>11</sup>.

Em termos do que realmente é ensinado, numerosos estudos mostram que a linguagem apresentada em livros didáticos de inglês ainda é baseada na intuição mais do que em evidências reais de uso. Burns (2001 apud O’Keeffe 2007:21) aponta que em termos da linguagem oral, os diálogos inventados apresentados em livros didáticos raramente refletem a imprevisibilidade e o dinamismo da conversação, ou as características e estruturas do discurso falado natural, e argumenta que alunos que lidam com diálogos montados têm menos oportunidades de ampliar seus repertórios linguísticos de modo que os preparem para interações fora da sala de aula. Um outro exemplo da falta de sintonia de livros didáticos com o mundo real é o estudo de Holmes (1998:40 apud O’Keeffe 2007:21). O autor observou a modalidade epistêmica em livros de inglês como segunda língua comparada com dados de *corpus* e encontrou em muitos livros, demasiada atenção dispensada aos verbos modais ao invés de outras alternativas linguísticas.

A produção de trabalhos sobre linguagem de aprendiz a partir de *corpus* digitalizado é portanto, bastante profícua. Entretanto, Nesselhauf (2004:136) afirma que mais estudos devem ser realizados, principalmente em certas áreas da gramática, do léxico e do discurso, e que mais

---

<sup>11</sup> As exceções figuram entre os dicionários de aprendizes, que utilizam os achados nos *corpora* para compor e /ou melhorar esses materiais, como o Longman Language Activator (1993), The Longman Dictionary of Contemporary English (1995), The Longman Essential Activator e o Cambridge International Dictionary of English (1995), entre outros. (Nesselhauf op. cit.)

do que focar em palavras isoladas, o ideal seria focar na função desempenhada por essas palavras. Isto é exatamente o que a presente dissertação se propõe a fazer que é o mapeamento do uso de advérbios em *-ly*, a sua frequência e distribuição na escrita de aprendizes.

## O objeto de estudo

A presente dissertação examina aquilo que já foi categorizado como a classe de palavras mais insatisfatória (Bloor e Bloor, 1995:22), o advérbio. Esses autores rotulam o advérbio como uma categoria “*ragbag*” ou “*dustbin*”, porque é a categoria que inclui uma amostra amorfa de itens lingüísticos que não podem ser colocados em qualquer outro compartimento, sob a pena de se categorizar o item de forma errônea.

Os advérbios são classes de palavras que marcam circunstâncias mas também têm a função primordial de criar espaço para o enunciador fazer um comentário ou passar julgamento e estabelecer relações entre duas orações, desta forma o advérbio pode ser um marcador de modalização<sup>12</sup>. Dentro da Gramática Funcional, recursos que nos ajudam a passar julgamento e comentar fazem parte da Macrofunção Interpessoal. Eles nos permitem negar, proclamar, probabilizar, atribuir.

Uma outra razão para se estudar o advérbio nesta dissertação tem a ver com a afirmação de Ringbom (1998), que estudou o inglês de aprendizes da União Européia., através do *corpus* ICLE, a autora observou que esses aprendizes fazem pouco ou nenhum uso de recursos que lhes permitem se inserir no texto ou modalizar suas vozes, devido, talvez, a uma possível limitação vocabular.

Ora, textos essencialmente argumentativos, como são os textos do *corpus* ICLE, deveriam fornecer a possibilidade de um enunciador lançar mão de recursos para convencer seu leitor. Saber usar marcadores de atitude, como modais ou advérbios, parece ser uma área nevrálgica para aprendizes de uma língua estrangeira, a qual merece ser estudada contextualmente.

A presente dissertação analisa, desta forma, esse grupo “*ragbag*” de classe de palavras que são os advérbios a partir da produção escrita em língua inglesa de três grupos, um deles

---

<sup>12</sup> ver capítulo 3 seção 3.2 para uma descrição completa sobre os advérbios.

formado exclusivamente de alunos falantes de português brasileiro como língua materna, outra composta por pré universitários e universitários, e a última composta de escrita produzida por graduandos e pós-graduados, usuários de inglês como primeira língua. A pesquisa ficou circunscrita ao uso de advérbios enquanto recursos léxico-gramaticais, inclusive para a expressão de atitude, com vistas ao mapeamento dos seus recursos em termos de frequência e em termos de função.

As perguntas básicas que orientam esta dissertação são:

- a) que tipo de advérbios os aprendizes brasileiros de inglês como língua estrangeira investigados usam em ensaios argumentativos (Br-ICLE)?
- b) quais os mais frequentes, como são posicionados na frase e com que função são usados?
- c) como o seu uso se compara ao uso de advérbios em textos semelhantes em inglês como primeira língua, em termos de função e estrutura?
- d) como os aprendizes brasileiros de língua inglesa constroem seu posicionamento<sup>13</sup> no discurso através do uso de advérbios?

Portanto, esta dissertação se divide em 5 partes. A primeira parte enfoca o que seja a Linguística de *Corpus*, a segunda aborda os advérbios de língua inglesa conforme apresentada em gramáticas funcionais, a terceira traça as bases metodológicas adotadas na pesquisa e a quarta apresenta as análises dos dados obtidos. Por fim são feitas algumas considerações sobre as aplicações pedagógicas dos estudos com *corpus*.

---

<sup>13</sup> Nesta dissertação entende-se como posicionamento aquilo que Biber et al (xxxx) define como stance, ou seja, uma grande categoria que engloba desde comentários até itens epistêmicos e deônticos.

## 1. REVISÃO DA LITERATURA: LINGÜÍSTICA DE CORPUS

O trabalho aqui proposto tem como ponto de partida a Lingüística de *Corpus* e como base lingüística o estudo do uso e da função de advérbios terminados em *-ly* em *corpora* de textos escritos. Portanto este capítulo enfoca o que seja Lingüística de *Corpus*, fazendo um breve histórico da área e levantando os dois tipos de pesquisa que são feitas a partir de extração de dados com o auxílio do computador. Além disso, este capítulo visa definir o que seja um *corpus*, qual é seu conteúdo, representatividade e autenticidade bem como as abordagens utilizadas nos estudos com *corpus*.

### 1.1. Histórico

A presente seção enfoca, ainda que brevemente, os primórdios e o desenvolvimento do que se convencionou chamar de Lingüística de *Corpus*.

O uso de *corpora* não é recente e tão pouco sua existência está condicionada à invenção do computador, ainda que nos últimos anos seja difícil falar de *Corpus* sem relacioná-lo com a tecnologia computacional. Como diz Berber Sardinha:

Havia *corpora* antes do computador, já que o sentido original da palavra *Corpus* é corpo, conjunto de documentos (conforme dicionário Aurélio). Na Grécia Antiga, Alexandre, o Grande definiu o *Corpus* Helenístico. Na Antiguidade e na Idade Média, produziam-se *corpora* de citações da Bíblia (Berber Sardinha 2004:3).

O *Oxford English Dictionary* é um exemplo do uso de *Corpus* de linguagem em uso, compilação essa que precedeu o advento do computador (Scott e Tribble, 2006:4). Segundo esses autores o dicionário foi compilado no século XIX, usando-se um grande número de fichas

contendo exemplos autênticos da linguagem em uma das mais impressionantes iniciativas de pesquisa já realizadas, [...] (Scott & Tribble 2006:4 tradução nossa)<sup>14</sup>”.

Alguns estudos sobre aquisição da linguagem, feitos aproximadamente entre 1876 e 1929, foram baseados em registros de diários, compilados pelos próprios pais, a respeito da fala das crianças, e que inspiraram pesquisadores como Preyer (1889) e Stern (1924). Käding (1897) montou um *Corpus* alemão para mostrar a distribuição da frequência de letras no idioma. Na área da pedagogia, Fries e Traver (1940) e Bongers (1947) são exemplos de lingüistas que usaram *Corpus* para analisar uma língua estrangeira. Em estudos comparativos, Eaton (1940) comparou a frequência do significado das palavras em holandês, francês, alemão e italiano. Todos estes estudos, citados em Berber Sardinha 2004, influenciam até hoje pesquisas recentes

Cabe acrescentar também, que “foi um *Corpus* não-computadorizado que deu feição aos *corpora* atuais, o SEU (*Survey of English Usage*) foi compilado por Randolf Quirk e sua equipe, em Londres, a partir de 1959” (Berber Sardinha, 2004). O SEU foi um *Corpus* organizado manualmente com o objetivo de armazenar um milhão de palavras, analisadas e classificadas gramaticalmente. Este *Corpus* serviu como referência para outros *corpora*, como o Brown, que será tratado mais adiante, e como base para a famosa gramática *Comprehensive Grammar of the English Language*, de Quirk et al. escrita em 1972.

A popularização dos estudos na área de Lingüística de *Corpus* está intrinsecamente ligada ao desenvolvimento de novas tecnologias, conforme aponta Tognini Bonelli (2001). Segundo a autora, o computador alterou a natureza da investigação lingüística uma vez que permitiu a análise, a seleção e a sistematização de *corpora* contendo milhares de palavras, em curto espaço de tempo, tarefa que era praticamente inviável há trinta anos atrás, quando os computadores pessoais não eram acessíveis. Na verdade, a confiabilidade das análises manuais de grandes *corpora* era justamente uma das críticas feitas a esses estudos, de acordo com Berber Sardinha (2004).

Pode-se dizer que a história da Lingüística de *Corpus* está, de certa forma, ligada à disponibilidade de *corpora* eletrônicos, mais notadamente de língua inglesa. A Grã-Bretanha é considerada um dos centros de pesquisa mais desenvolvidos. É lá que várias universidades (Birmingham, Lancaster, Liverpool e Londres) dedicam-se à pesquisa baseada em *Corpus* para

---

<sup>14</sup>Tradução do original em inglês: “In the late 19th century, the Oxford Dictionary was compiled in one of the most impressive research initiatives ever undertaken, using enormous number of slips collected containing authentic examples of language in use...”

descrição dos mais variados aspectos da linguagem. Nos Estados Unidos, os estudos sobre *Corpus* têm mais o viés do processamento das linguagens naturais para utilização em inteligência artificial. Entretanto, há nos Estados Unidos dois nomes mundialmente conhecidos e de maior relevância para a Linguística de *Corpus*: Mark Davies, devido aa disponibilização online de vários ‘bancos’ de textos eletrônicos e Douglas Biber, cujo trabalho será tratado em seção posterior.

Berber Sardinha (2004) cita ainda três *corpora* que nos servem como marcos de referência, Brown, BNC e Bank of English. O Brown foi o pioneiro dentre os *corpora* digitais; o BNC foi o primeiro a atingir 100 milhões de palavras; o Bank of English, diferentemente dos demais que são *corpora* de amostragem, planejados e fechados, é um *Corpus* monitor, orgânico e em crescente expansão. Em 2004, Berber Sardinha (2004: 6) afirmava que no Brasil, a Linguística de *Corpus* ainda estava em estágio inicial e acrescentava que a pesquisa se dava mais naqueles centros voltados ao Processamento de Linguagem Natural (ou Linguística Computacional). Entretanto, em cinco anos, o Banco do Português, compilado por esse pesquisador, já alcançou 1 bilhão de palavras e já há inúmeras teses e dissertações que têm como ponto de partida a Linguística de *Corpus*.

## **1.2. Linguística de *Corpus*: Fundamentação teórica e considerações sobre métodos de investigação linguística**

McEnery e Wilson (1996:1<sup>15</sup> tradução nossa) afirmam que “a Linguística de *Corpus* é talvez mais bem descrita, em termos simples, como o estudo da linguagem baseado em exemplos da ‘vida real’”. Dois anos depois, Kennedy (1998:8<sup>16</sup> tradução nossa) acrescentaria que “a Linguística de *Corpus*, como toda linguística, está preocupada primeiramente com a descrição e explicação da natureza, estrutura e uso da linguagem e línguas e com questões particulares, tais como aquisição da linguagem, variação e mudança”.

Ainda na tentativa de conceituar esse campo de estudo, vale dizer também que “o objetivo da linguística de *Corpus* é a análise e a descrição da linguagem em uso, realizada em

---

<sup>15</sup> Tradução do original em inglês : “*Corpus* Linguistics is perhaps best described for the moment in simple terms as the study of language based on examples of ‘real life’ language use.”

<sup>16</sup> Tradução do original em inglês: “*Corpus* linguistics, like all linguistics, is concerned primarily with the description and explanation of the nature, structure and use of language and languages and with particular matters such as language acquisition, variation and change.”

texto(s) (Tognini-Bonelli, 2001:2<sup>17</sup> tradução nossa)”. Ou seja, o lingüista de *Corpus* objetiva a descrição de uma determinada língua ou variante lingüística, tal como ela é utilizada por um determinado grupo de enunciadores em um evento comunicativo específico a partir da análise de um *Corpus* especialmente compilado para o objetivo do pesquisador. Como diz Berber Sardinha (2004):

A Lingüística de *Corpus* ocupa-se da coleta e da exploração de *corpora*, ou conjuntos de dados lingüísticos textuais coletados criteriosamente com o propósito de servirem para pesquisa de uma língua ou variedade lingüística. Como tal, dedica-se a exploração da linguagem por meio de evidências empíricas, extraídas por computador. (Berber Sardinha 2004:3)

Há, portanto, um consenso sobre o que seja a Linguística de *Corpus*. Entretanto, a definição do que seja o objeto da Linguística de *Corpus* ainda está sendo refinada. Mais recentemente, Adolphs (2006) adotou o termo “análise de textos eletrônicos”, para se referir ao escopo da Linguística de *Corpus*, por se tratar de um termo mais preciso. Segundo a autora:

O termo análise de textos eletrônicos foi adotado para refletir as diferentes prioridades das fontes dos dados e dos processos de pesquisa quando comparamos a lingüística de *Corpus* como uma tradição com outras áreas de pesquisa da linguagem que fazem uso do auxílio computacional. Como tal, o termo análise de texto eletrônico foi escolhido pelo seu caráter inclusivo e significado amplo, que está relacionado com a análise de qualquer texto ou coleção de texto digitalizado (Adolphs 2006:2 tradução nossa)<sup>18</sup>.

A Linguística de *Corpus* investiga exemplos reais de linguagem em uso e, desta forma, ela tem como ponto de partida a *performance* e não a competência (ou o uso idealizado) da mesma. Nos trabalhos que usam a Linguística de *Corpus* como arcabouço, deixam-se de lado os exemplos inventados para se trabalhar com instâncias lingüísticas naturais (Knowles, 1997).

Entretanto, existe um debate quanto ao estatuto da Linguística de *Corpus*, no que diz respeito a ela ser considerada uma metodologia ou apenas um ramo da Linguística.

---

<sup>17</sup> Tradução do original em inglês: “[...] the aim of *Corpus* linguistics can be seen as the analysis and the description of language as use, as realized in text(s).”

<sup>18</sup> Tradução do original em inglês: “The term electronic text analysis had been adopted to reflect the different priorities in terms of data sources and research processes when we compare *Corpus* linguistics as a tradition with other areas of computer-aided language research. As such, the term electronic text analysis has been chosen for its inclusive and broad meaning that relates to the analysis of any digitized text or text collection.”



Scott & Tribble (2006:3 <sup>19</sup> tradução nossa), por exemplo, assumem que a Linguística de *Corpus* é pura e simplesmente uma metodologia, ou seja, “[...] são métodos que usam *corpora* de textos, sejam escritos ou falados, como exemplos genuínos da linguagem em uso.”

A mesma linha de pensamento é seguida por McEnery e Wilson (1996) e Kennedy (1998). Os primeiros refutam a idéia de que a Linguística de *Corpus* seja apenas um ramo da Linguística e declaram que ela deve ser vista como uma metodologia também. O segundo diz que a Linguística de *Corpus* funciona como uma representação metodológica e seria um erro assumi-la como uma teoria da linguagem em competição com outras teorias da linguagem, ou como um ramo novo ou separado da linguística.

Linguística de *Corpus* não é um ramo da linguística tal qual a sintaxe, a semântica e a sociolinguística. Todas essas disciplinas se concentram em descrever/explicar alguns aspectos da linguagem em uso.[...] Sintaxe, semântica e pragmática são apenas três exemplos de áreas da linguística que usam uma abordagem baseada em *Corpus*. A Linguística de *Corpus* é uma metodologia que deve ser usada em qualquer área da linguística, mas não deve delimitar uma área específica (McEnery e Wilson 1996:2 tradução nossa)<sup>20</sup>

A Linguística de *Corpus* é baseada em conjuntos de textos como domínio de estudo, como fonte de evidência para a descrição e argumentação linguística. Ela também vem para representar metodologias para a descrição linguística cuja quantificação da distribuição dos itens linguísticos faz parte da pesquisa (Kennedy 1998:7 tradução nossa)<sup>21</sup>.

Por outro lado, Adolphs (2006) aponta que a utilização de textos eletrônicos na Linguística de *Corpus* levou a uma mudança na forma como as pesquisas foram e são conduzidas, sem que para isso elas estejam necessariamente atreladas a uma metodologia.

A finalidade da análise de textos eletrônicos nesse contexto é de se expandir com o desenvolvimento de novas tecnologias para recuperar, analisar e representar textos eletrônicos. Deve-se notar nesse contexto que a manipulação e a análise de textos eletrônicos não está confinada aos métodos desenvolvidos e usados dentro da área da Linguística de *Corpus*. Os tipos de métodos que são tradicionalmente usados pela Linguística de *Corpus* para facilitar a pesquisa

<sup>19</sup>Tradução do original em inglês: “...methods using *corpora* of texts, whether written or spoken, that is to say genuine examples of language in use.”

<sup>20</sup> Tradução do original em inglês: “*Corpus* linguistics is not a branch of linguistics in the same sense as syntax, semantics, sociolinguistics, and so on. All of these disciplines concentrate on describing/explaining some aspect of language use. Syntax, semantics and pragmatics are just three examples of areas of linguistics enquiry that have used a *Corpus*-based approach. *Corpus* linguistics is a methodology that may be used in almost any area of linguistics, but it does not truly delimit an area of linguistics itself”.

<sup>21</sup>Tradução do original em inglês: “*Corpus* linguistics is based on bodies of texts as the domain of study as the source of evidence for linguistic description and argumentation. It has also come to embody methodologies for linguistic description in which quantification of the distribution of linguistics items is part of the research activity.”

lexicográfica encontraram caminho para uma série de áreas, particularmente aquelas associadas com o campo mais amplo da linguística aplicada. (Adolphs, 2006:9 tradução nossa)<sup>22</sup>

Pode-se acrescentar ainda, que a Linguística de *Corpus* está muito além de uma metodologia (Tognini-Bonelli, 2001)

Enquanto uma metodologia pode ser definida como o uso de um dado conjunto de regras ou áreas de conhecimento em uma determinada situação, a linguística de *Corpus* está em posição de definir suas próprias regras e áreas de conhecimento antes delas serem aplicadas. (Tognini-Bonelli 2001:1 tradução nossa)<sup>23</sup>

Entretanto, Berber Sardinha (2004) diz que aceitar a Linguística de *Corpus* como metodologia ou não, depende daquilo que se entende por metodologia.

Entendendo metodologia como instrumental então é possível aplicar o instrumental da Linguística de *Corpus* livremente e manter a orientação teórica da disciplina original. Desse modo, teríamos, por exemplo, a sintaxe baseada em *Corpus* versus a sintaxe tradicional [...]. o que mudaria entre essas vertentes seria o instrumental; os dados, a orientação, os pressupostos teóricos, as implicações dos resultados e tudo o mais permaneceriam (Berber Sardinha, 2004:36)

Qualquer que seja o estatuto da Linguística de *Corpus*, o trabalho com *Corpus* envolve vários grupos de linguistas/pesquisadores de *Corpus*, que Kennedy (1998) divide em quatro. O primeiro consiste em compiladores de *corpora*. Eles estão preocupados com o *design* e a compilação dos *corpora*, com a coleta dos textos e com a sua preparação para posterior análise. O segundo, desenvolve ferramentas para análise de *corpora*. O terceiro é composto por linguistas cujo objetivo é utilizar *Corpus* computadorizado para descrever o léxico e a gramática das línguas, a partir de uma visão probabilística, onde o que importa não é apenas o quê, quando, onde e por quem algo foi dito ou escrito, mas sobretudo, como as formas linguísticas foram usadas. O quarto grupo se ocupa da exploração de *Corpus* para aplicação em diversas áreas de estudo, tais como ensino e aprendizagem de línguas, processamento de linguagem natural,

---

<sup>22</sup>Tradução do original inglês: “The scope of electronic text analysis in this context is expanding with the development of new techniques to retrieve, analyse and represent electronic texts. It should be noticed in this context that electronic text manipulation and analysis is not confined to the methods developed and used within the area of *Corpus* linguistics. The types of methods that have traditionally been used by *Corpus* linguists to facilitate research in lexicography have found their way into a range of areas, particularly those associated with the broader field of applied linguistics.”

<sup>23</sup>Tradução do original em inglês: “While a methodology can be defined as the use of a given set of rules or pieces of knowledge in a certain situation [...] *Corpus* linguistics is in a position to define its own sets of rules and pieces of knowledge before they are applied.”

reconhecimento da fala e tradução. O *Corpus* de aprendizes BR-ICLE, utilizado nesta dissertação, encaixa-se no terceiro e quarto grupos apresentados.

Em suma, mesmo que a Linguística de *Corpus* pudesse ser definida pelo uso que as pessoas fazem dos *corpora* “seria um erro assumir que a Linguística de *Corpus* é apenas um meio mais rápido e simples para se descrever como a linguagem é processada (Kennedy 1998:9 tradução nossa)<sup>24</sup>”.

### 1.3. Linguística de *Corpus*: Abordagem empírica ou racionalista?

Em linguística, existem visões que se opõem no campo de estudos da linguagem. De um lado, temos as obras de Chomsky (1957 e 1965) que são representativas da linha racionalista, e do outro, temos trabalhos de Halliday (1985 e 1991), seguidos pelos de Sinclair (1989 et seq.), que são representantes da linha empírica.

Essas duas vertentes colidem principalmente na questão que envolve a observação dos dados, ou seja, ou se escolhe basear as observações em dados produzidos por introspecção do analista ou em dados naturais. Dessa forma, entende-se que em uma abordagem racionalista há maior preocupação com a forma pela qual a mente processa a linguagem, enquanto que na abordagem empírica, a preocupação está em observar a ocorrência de dados naturais. Como diz Tsui:

Antes da existência dos *corpora*, as descrições lingüísticas de baseavam muito na intuição e introspecção do falante nativo. Elas descreviam o que as pessoas *sabiam* sobre a língua, ou o que elas *achavam* que a língua fosse, mais do que *como* a língua era usada. A facilidade de acesso a grandes amostras de textos naturais nos possibilitou testar a robustez das descrições lingüísticas que eram baseadas na introspecção e na elicitación, para obter uma compreensão nova sobre a estrutura e o uso da língua. Isso nos ajudou a ter um melhor entendimento sobre como a língua é realmente usada, mais do que como a língua é percebida (Tsui 2004:39 tradução nossa)<sup>25</sup>.

<sup>24</sup>Tradução do original inglês: “it would be a mistake to assume that *Corpus* linguistics is *simply* a faster way of describing how a language works”

<sup>25</sup>Tradução do original em inglês: “Before the existence of *corpora*, linguistic descriptions relied very much on native-speaker intuition and introspection. They describe what people *know* about language or what they perceive language to be, rather than *how* language is used. The easy accessibility of huge bodies of naturally occurring texts on the computer has made it possible for us to test the robustness of linguistic descriptions which were based on introspection and elicitation, and to gain new insights into language structure and use. It has helped us to gain a better understanding of language is *actually* used rather than how language is perceived to be used”.

Dentro da visão racionalista da língua, Chomsky sugeria que um *Corpus* não poderia ser uma ferramenta útil para um lingüista uma vez que o *Corpus* deveria descrever a língua através da competência e não da *performance*. Além disso, um *Corpus* não poderia representar uma língua em sua totalidade; então, não poderia ser representativo dessa língua. “Chomsky argumentava que *corpora* eram inadequados como base para descrever regras gramaticais na medida que “algumas frases não ocorrerão porque são óbvias, outras porque são falsas e outras porque são grosseiras (Kennedy 1998:23 <sup>26</sup>)”.

Ainda dentro dessa visão, McEnery e Wilson (1996:05) afirmam que Chomsky argumentava que era a competência, mais do que o desempenho, que o lingüista estava tentando modelar; a competência explicaria e caracterizaria o conhecimento da linguagem do falante. Isso quer dizer que para os chomskyanos, o papel principal dos lingüistas seria descrever e detalhar um modelo de competência lingüística. Por outro lado, “a análise de um *Corpus* pode revelar, e freqüentemente revela, fatos a respeito de uma língua que nunca se pensou em procurar (Kennedy 1998:9) <sup>27</sup>”, ou seja, que não seriam possíveis apenas através da mera introspecção do pesquisador.

Pode-se dizer, então, que existe uma evidente oposição entre uma abordagem racionalista e abordagens baseadas em *Corpus*. A primeira vê a linguagem como possibilidade e enfatiza, desta forma, a descrição de quais grupamentos sintáticos são possíveis (permissíveis) dado o conhecimento que um falante ideal possui de sua língua. O estudo a partir de *Corpus* privilegia a **probabilidade** de sistemas linguísticos ocorrerem dentro da língua em relação a determinados contextos em que os falantes os empregam. De acordo com essa noção, portanto, apesar de várias estruturas léxico-gramaticais poderem ocorrer em um determinado contexto, elas não ocorrem com a mesma freqüência ou distribuição em gêneros distintos. Parte-se então do pressuposto que, para se estudar o infinito da língua, cria-se uma amostra e a essa amostra dá-se o nome de *Corpus* de estudo.

Entretanto, segundo aponta Tognini Bonelli (2001), é possível realizar uma pesquisa linguística unindo essas duas correntes:

---

<sup>26</sup>Tradução do original em inglês: “Chomsky is reported to have argued that *corpora* was inadequate as a basis for describing grammatical rules on the ground that ‘some sentences won’t occur because they are obvious, others because they are false, still others because they are impolite”

<sup>27</sup>Tradução do original em inglês: “Analysis of a *Corpus* by means of standard *Corpus* linguistics research software can and frequently does reveal facts about language which we might never previously have thought of seeking.”

O trabalho com *Corpus* pode ser visto como uma *abordagem empírica* que, como em todo processo científico, tem como ponto de partida dados autênticos. O procedimento para descrever os dados, no entanto, é *indutivo* porque são as formulações de natureza teórica ou cultural que advêm de observações de exemplos reais. (Tognini-Bonelli 2001:2 tradução nossa)<sup>28</sup>

A mesma linha de pensamento é abarcada por Adolphs (2006:6<sup>29</sup> tradução nossa) quando diz que a “descrição da língua e as teorias resultantes podem ser baseadas tanto na observação da linguagem natural quanto na intuição do falante”. Ou seja, a observação dos dados leva à formulação de hipóteses, que por sua vez, leva à generalização baseada em evidências que finalmente permitirá uma formulação teórica.

No caso desta dissertação, que investiga advérbios, entende-se que esses aparecerão com maior ou menor frequência levando-se em conta alguns fatores tais como, o tipo de texto, a temática do texto ou o nível de proficiência linguística dos sujeitos. Além disso, determinados advérbios poderão estabelecer padrões, ou seja poderão ocorrer lado a lado com determinados grupos de palavras (colocados). Para viabilizar a pesquisa, é necessário que o material de estudo, ou seja, os *corpora*, estejam de acordo com aquilo que se pretende investigar. Dessa forma, eles devem atender a algumas especificações no que diz respeito ao seu conteúdo, representatividade e autenticidade, mas primeiro, é necessário definir o que seja um *Corpus*.

#### 1.4. Definição de *Corpus*

Qualquer coleção composta de mais de um texto é chamada *Corpus*, cujo termo é simplesmente a forma em latim para corpo. Todavia, quando o termo é usado em contextos atuais ele assume conotações mais específicas.

Berber Sardinha (2004) afirma que “A existência de uma coletânea de dados lingüísticos naturais, legíveis por computador é central para a Lingüística de *Corpus*”. Esses dados lingüísticos constituem o *Corpus* que, segundo Tognini-Bonelli (2001), é a representação da linguagem natural (autêntica) selecionada e organizada de acordo com critérios explícitos e com

<sup>28</sup>Tradução do original em inglês: “*Corpus* work can be seen as an *empirical approach* in that, like all types of scientific enquiry, the starting point is actual authentic data. The procedure to describe the data that makes use of a *Corpus* is therefore *inductive* in that it is statements of a theoretical nature about language or the culture which are arrived at from observations of the actual instances.”

<sup>29</sup>Tradução do original em inglês: “Description of language and resulting theories can be based either on the observation of naturally occurring discourse in use or on the intuition of a person who uses particular language.”

o objetivo de ser usada como amostra da língua. Pode-se dizer também que “um *Corpus* constitui uma base empírica, não só para identificar elementos e padrões estruturais que constroem os sistemas que usamos na língua, mas também para mapear nosso uso desses sistemas (Kennedy 1998:4 <sup>30</sup> tradução nossa)”.

Uma definição considerada mais completa é apresentada por Sanchez e Cantos (apud Berber Sardinha 2004:18). Segundo esses autores, um *Corpus* é:

um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise.

Esta citação parece ser mais completa porque leva em consideração alguns pontos relevantes quanto à característica de um *Corpus*, tais como: a origem do *Corpus*, já que os dados devem ser autênticos, ou seja, produzidos pela comunicação genuína dos seus falantes **sem** o propósito de serem alvo de pesquisa lingüística; o propósito do *Corpus*, cuja finalidade é de ser um objeto de estudo lingüístico; o conteúdo do *Corpus*, que deve ser criteriosamente escolhido, respeitando-se as condições de naturalidade e autenticidade; a formatação do *Corpus*, cujos dados, devem ser legíveis por computador; a representatividade do *Corpus*, que deve ser uma amostra representativa de uma língua ou variedade lingüística; e a extensão do *Corpus*, que deve ser vasto o suficiente para ser representativo.

Questões como o conteúdo, a representatividade e a autenticidade de um *Corpus*, serão abordadas a seguir.

#### 1.4.1. Conteúdo de um corpus

Um *Corpus* adequado aos objetivos da pesquisa deve ser selecionado a partir de critérios como o tipo de texto a ser selecionado, a quantidade e a extensão dos textos. Biber (apud Tognini

---

<sup>30</sup>Tradução do original em inglês: “A *Corpus* constitutes an empirical basis not only for identifying the elements and structural patterns which make up systems we use in a language, but also for mapping out our use of these systems.”

Bonelli 2004:59) afirma que já que um *Corpus* precisa ser representativo para ser usado como base para se fazerem generalizações da língua, o critério primordial deveria ser a definição da população alvo que visa representar. Deve ser o mais específico possível principalmente no que diz respeito à composição e na medida do possível, aos critérios de escolha.

Ao se definir o *Corpus* que será utilizado em qualquer estudo, cabe, no entanto, decidir como o *Corpus* será explorado, se quantitativa ou qualitativamente, e qual abordagem de análise será adotada, aspectos estes que serão tratados na seção a seguir.

#### 1.4.2. Representatividade de um corpus

“A característica mais facilmente associada à representatividade é justamente a extensão do *Corpus*, o que significa, em termos simples, que para ter representatividade o *Corpus* deve ser o maior possível. (Berber Sardinha 2004:22)”. Considerando-se a linguagem como um sistema probabilístico, quanto maior for o *Corpus*, maior serão as chances de que um determinado elemento lingüístico ocorra.

Sendo a extensão do *Corpus* um critério para sua representatividade, cabe acrescentar aqui uma classificação, sugerida por Berber Sardinha, quanto ao tamanho do *Corpus*.

<b>Tamanho em palavras</b>	<b>Classificação</b>
Menos de 80 mil	Pequeno
80 a 250 mil	Pequeno-médio
250 mil a 1 milhão	Médio
1 milhão a 10 milhões	Médio-grande
10 milhões ou mais	Grande

Tabela 1: Classificação relativa ao tamanho do *Corpus* (Berber Sardinha 2004:26)

Segundo Sardinha, portanto, há *corpora* pequenos, pequeno-médios, médios, médio-grandes e grandes. Entretanto, como argumentam McEnery e Wilson, um *Corpus* é representativo do quê? Para responder a essa questão deve se ter em mente que o *Corpus* precisa ser uma amostra representativa daquilo que se pretende investigar.

Em linguística, estamos freqüentemente interessados em uma variedade completa da língua mais do que em um único texto ou autor. Ao construir um *Corpus* de uma variedade linguística estamos interessados em uma amostra que seja a mais representativa possível da variedade sob investigação, que ela possa nos dar o perfil mais exato possível das tendências dessa variedade, incluindo suas proporções. (McEnery & Wilson 1996:21 tradução nossa<sup>31</sup>)

Ainda sobre representatividade, Leech (apud Kennedy 1998:62 tradução nossa<sup>32</sup>) sugere que “um *Corpus* é representativo na medida em que as descobertas baseadas nele possam ser generalizadas para língua como um todo ou para parte específica dela. Uma amostra de textos pode ser representativa de todos os gêneros possíveis ou mesmo de um gênero particular ou de determinados assuntos ou tópicos”.

Uma outra pergunta relevante que antecede a compilação é para quem um *Corpus* é representativo? Ao desenvolver um estudo, o pesquisador deve ter em mente que o *Corpus* cumpre um propósito e que tal propósito está destinado a uma comunidade específica. “Um *Corpus* constitui uma referência padrão para a variedade linguística que representa. Isso pressupõe sua ampla utilidade para outros pesquisadores (McEnery & Wilson 1996:24 tradução nossa<sup>33</sup>)”.

### 1.4.3. Autenticidade de um corpus

<sup>31</sup>Tradução do original em inglês “In linguistics, we are more often interested in a whole variety of a language, rather than in an individual text or author. In building a *Corpus* of a language variety, we are interested in a sample which is maximally representative of the variety under examination, that is, which provides us with as accurate a picture as possible of the tendencies of that variety, including their proportions.”

<sup>32</sup>Tradução do original em inglês: “a *Corpus* is representative in the sense that findings based on an analysis of it can be generalized to the language as a whole or a specific part of it. [...] A sample of texts can be thoroughly representative of all possible genres or even of a particular genre or subject field or topic”

<sup>33</sup> Tradução do original em inglês: “A *Corpus* constitutes a standard reference for the language variety which it represents. This presupposes its wide availability to other researches”.



Para um *Corpus* ser considerado autêntico não basta apenas que ele seja a representação de instâncias linguísticas reais. O critério de coleta dos textos precisa estar adequado aos propósitos da investigação. Além disso, a adequação do *Corpus* está ligada à validade de seu conteúdo. Ou seja, *corpora* de língua inglesa como o *Brown*, BNC ou o *Bank of English*, citados na seção 2.1 deste trabalho, podem ser considerados como representantes de uma língua como um todo ou como parte, seja ela falada ou escrita. Há outros *corpora*, como os *corpora* de aprendizes, objeto de estudo desta dissertação, os quais podem ser considerados autênticos. Sabe-se entretanto, que há uma grande controvérsia com relação à autenticidade desses *corpora*.

Inúmeros linguistas, como por exemplo Tan (2005), dizem que há dúvidas se o *Corpus* composto de linguagem de aprendiz é realmente autêntico. A noção de linguagem autêntica foi sempre associada ao uso `real` e este termo se refere ao idioma falado no contexto real de nativos. Entretanto, há um outro grupo de linguistas, que inclui os linguistas que estudam a produção de escrita e fala de aprendizes (Granger 2004), que acreditam que a linguagem produzida em sala de aula ou em circunstâncias de ensino, espelha a realidade desse grupo de usuários e é portanto, genuína.

Uma vez tendo o pesquisador definido o *Corpus* que será usado, cabe, no entanto, decidir como este *Corpus* será explorado e qual abordagem de análise pode ser adotada, aspectos estes tratados na próxima seção.

## **1.5. Abordagens ao estudo do *Corpus***

Os dados contidos no *Corpus* podem ser trabalhados por meio de uma abordagem qualitativa ou quantitativa. Além disso, a finalidade do *Corpus* irá definir se a abordagem é ‘baseada em *Corpus*’ ou ‘direcionada pelo *Corpus*’. Essa seção do trabalho visa fundamentar e definir sob quais perspectivas a análise de textos digitalizados pode ser feita.

### **1.5.1. Análise quantitativa ou qualitativa?**

Normalmente, os trabalhos sobre pesquisa através de textos se dividem na seguinte dicotomia: pesquisa qualitativa oposta à pesquisa quantitativa. Na análise qualitativa não há preocupação com a frequência de ocorrências dos elementos linguísticos encontrados no *Corpus*. Já na análise quantitativa, os elementos são classificados, contados e algumas vezes, são construídos modelos estatísticos para tentar explicar o que foi observado. Na análise qualitativa os dados são usados apenas como base para identificar e descrever aspectos da linguagem em uso e assim fornecer exemplos genuínos de um determinado fenômeno. McEnery e Wilson argumentam que:

A análise qualitativa oferece uma perspectiva rica e detalhada dos dados. Fenômenos raros recebem, ou pelo menos deveriam receber, a mesma atenção que os mais frequentes fenômenos e porque o objetivo é uma descrição detalhada mais do que uma quantificação, uma pequena variação nos dados é levada em consideração: ambigüidade e valor semântico de algumas palavras. (McEnery & Wilson 1996:62 tradução nossa <sup>34</sup>)

A principal desvantagem da análise qualitativa é que os seus dados não podem ser estendidos a populações mais amplas com o mesmo grau de certeza que a análise quantitativa. Ou seja, “a análise quantitativa permite a descoberta de quais fenômenos são mais prováveis de serem reflexos genuínos do comportamento de uma língua ou variedade e quais são meramente obras do acaso (McEnery e Wilson 1996:63 tradução nossa) <sup>35</sup>”.

Por outro lado, pode-se dizer que “os dados que surgem da análise quantitativa são necessariamente menos ricos do que aqueles da análise qualitativa. A análise qualitativa pode fornecer maior riqueza e precisão enquanto que a análise quantitativa pode fornecer resultados mais confiáveis e generalizáveis (McEnery & Wilson 1996:63 tradução nossa) <sup>36</sup>”.

Entendemos, porém, que a análise quanti-qualitativa ou quali-quantitativa é a abordagem adotada no presente trabalho, na medida em que ambas as abordagens se complementam, ou

---

<sup>34</sup> Tradução do original em inglês: “Qualitative form of analysis offer a rich and detailed perspective of the data. In qualitative analyses, rare phenomena receive, or at least ought to receive, the same attention as more frequent phenomena, and, because the aim is complete detailed description rather than quantification, delicate variation in the data is foregrounded: ambiguity, semantic value of some words.”

<sup>35</sup> Tradução do original em inglês: “Quantitative analysis enables one to discover which phenomena are likely to be genuine reflections of the behavior of a language or variety and which are merely chance occurrences.”

<sup>36</sup> Tradução do original em inglês: “...the picture of the data which emerges from quantitative analysis is necessarily less rich than that obtained from qualitative analysis. Qualitative analysis can provide greater richness and precision whereas quantitative analysis can provide statistically reliable and generalisable results.”

seja, os dados qualitativos são usados para explicar os achados quantitativos e os dados quantitativos são usados para contextualizar os dados qualitativos.

### 1.5.2. Abordagem baseada em corpus ou abordagem direcionada pelo corpus?

Um *Corpus* pode ser usado de maneiras distintas com a finalidade de validar, exemplificar ou construir uma teoria linguística. Tognini-Bonelli (2001) aponta pelo menos duas dessas maneiras: a abordagem **baseada em *Corpus*** e a abordagem **direcionada pelo *Corpus***. Toda essa seção é baseada no trabalho da autora em questão.

A abordagem baseada em *Corpus* se refere a uma metodologia que se aproveita do *Corpus* principalmente para expor, testar, exemplificar teorias e descrições que foram formuladas antes dos *corpora* se tornarem disponíveis para estudar a linguagem.

Há inúmeros exemplos de análise baseada em *Corpus*. Como exemplo, podemos citar a gramática de Biber et al. (1999) os quais mostram, entre muitos outros fatos linguísticos, que os substantivos, as nominalizações, os sufixos derivacionais e os advérbios de ligação são mais comuns na prosa acadêmica enquanto que os verbos, a omissão do *that* e as contrações são mais raros. Em outro estudo baseado em *Corpus*, Wei-yu Chen (2006) oferece uma contribuição quantitativa e qualitativa acerca do uso de adverbiais conjuntivos na escrita de aprendizes *taiwaneses* em nível avançado de inglês como língua estrangeira. No aspecto quantitativo, as produções dos aprendizes são comparadas a artigos acadêmicos publicados cujo objetivo é generalizar o padrão de uso dos dois grupos. No aspecto qualitativo, a pesquisa foca no uso de determinados adverbiais conjuntivos pelo grupo de aprendizes para ilustrar que estes podem usar alguns conectivos erroneamente, mesmo estando em nível avançado de inglês. Em todos esses estudos, as categorias usadas são parte de uma gramática já estabelecida.

Já na abordagem direcionada pelo *Corpus* o compromisso é com as evidências produzidas pelo *Corpus*, para uma posterior construção de uma teoria, ou de descrição de um achado. Pode-se assumir que se trata de uma abordagem contextual e funcional. O processo metodológico é simples: a observação leva a hipótese, que leva à generalização, que leva à unificação de

fundamentos teóricos. Em outras palavras, os dados obtidos dos *corpora* são usados para a formulação de descrições que podem ter conteúdo gramatical, lexical ou léxico-gramatical.

Na abordagem direcionada pelo *Corpus*, pode-se extrair, por exemplo, palavras individuais, palavras-chave, concordâncias de palavras individuais ou chave, além de ser possível fazer-se etiquetagem de todo o *Corpus* a partir das classes gramaticais, partes sintáticas ou áreas semânticas existentes. Cada uma dessas operações é possível graças a um ferramental de análise, os programas “garimpadores” (text-mining) de texto, e programas etiquetadores. Com relação ao *Corpus* de aprendiz, as abordagens preferidas são as direcionadas pelo *Corpus*, porque os *corpora* de aprendiz contêm ‘erros’ de toda a natureza, que os indispõem a serem tratados com etiquetadores desenhados originariamente para a linguagem em uso de ‘nativos’.

São exatamente os erros que podem ser investigados por uma abordagem direcionada pelo *Corpus*. Nesselhauf (2004) enfatiza a importância do uso de *corpora* de aprendizes como base para análises direcionadas pelo *Corpus* na medida em que este permite, por exemplo, observar os erros ou diferenças entre a linguagem do aprendiz e do nativo. Desta forma, o estudo de Römer (2004) pode ser classificado como ‘direcionado pelo *Corpus*’. A autora resumiu os resultados de uma investigação cujo foco foi a análise direcionada pelo *Corpus* para analisar nove verbos modais em uma gramática de inglês falado e em um livro didático alemão. Römer sugere que esse tipo de abordagem é muito útil no ensino/aprendizagem tanto para professores como para editoras, uma vez que permite que o ensino da língua esteja de acordo com a linguagem em uso, embora ainda muito deva ser feito para se atingir o objetivo de tornar a língua aprendida a mais autêntica e a mais próxima dos falantes nativos.

Nada impede, entretanto, que se faça uma junção das duas abordagens, conquanto isso ainda seja mais raro. Como exemplo citamos Tankó (2004) que uniu as duas abordagens (baseada em *corpus* e direcionada pelo *corpus*) em sua pesquisa ao estudar o uso de conectivos em ensaios argumentativos de aprendizes húngaros e ao propor uma análise direcionada pelo *Corpus* no ensino de conectivos em sala de aula.

Portanto, neste capítulo vimos os princípios fundadores da Linguística de *Corpus*, bem como aquilo que se entende sobre *corpus* e algumas abordagens que podem ser feitas no estudo com *corpus*.

O próximo capítulo levantará algumas considerações sobre como algumas gramáticas estudam uma língua e como o advérbio, elemento a ser pesquisado nesta dissertação, é classificado de acordo com a abordagem sistêmico-funcional.

## **2. REVISÃO DA LITERATURA: ADVÉRBIOS**

O presente capítulo faz uma revisão da literatura existente sobre o tópico principal deste estudo, o advérbio. Para tal, a primeira seção deste capítulo levanta aspectos sobre o teor de algumas gramáticas de língua inglesa, no que diz respeito à forma como estudam os fatos lingüísticos. A segunda seção aborda o advérbio propriamente dito, de acordo com duas gramáticas de orientação funcional.

### **2.1. As Gramáticas de Língua Inglesa**

Quando se pensa em ensino de língua estrangeira uma dúvida parece surgir: é fato que uma língua deva ser ensinada tendo como base algum ponto de referência ou norma alvo? Se for, qual seria essa referência ou norma alvo? Esta seção parte destes questionamentos, para discutir o papel desempenhado pelas gramáticas de língua.

A gramática de qualquer língua pode ser estudada por diferentes perspectivas. Uma diferença básica figura entre a teoria e a descrição. Estudos com orientação teórica se baseiam na descoberta de princípios abstratos em relação a um modelo de competência linguística. Se for levada em consideração essa abordagem prescritiva, há de se supor que haja apenas uma única forma correta da língua que deveria ser acatada por todos. É importante reconhecer, entretanto, que nem todo mundo usa a “linguagem correta” e, dessa forma, aquilo que não se encaixa no que é correto, simplesmente é considerado como errado.

A gramática prescritiva, portanto, estabelece uma série de regras que são baseadas naquilo que é considerado correto ou incorreto em variedades padrão, além de fazer avaliações daquilo que é certo ou errado a partir de julgamentos de aceitabilidade e adequação, que freqüentemente entram em desacordo (c.f. Quirk, 1976).

Já os estudos descritivos tentam criar um panorama mais abrangente dos fenômenos gramaticais em uma língua específica. . Esses estudos estão comprometidos em descrever como a linguagem é usada, mais do que em prescrever como ela deveria ser usada. Se a linguagem

estabelece comunicação , então é provável que os interlocutores estejam seguindo regras, mesmo embora não sejam elas regras aprovadas universalmente. Ou seja, se a abordagem a gramática de uma língua for considerada descritiva, não se poderá classificar algumas formas da linguagem como incorretas. Entretanto é importante que se descrevam as variações, que podem estar de acordo com o usuário da língua, sua origem, classe social, idade e sexo, e também com o uso, no que concerne a questões como a relação entre enunciador e co-enunciador, linguagem falada ou escrita, e o objetivo pretendido.

Ainda segundo Quirk (op.cit.), na gramática vista como descritiva, entende-se que o vocabulário e/ou as escolhas gramaticais que fazemos são influenciados por uma série de fatores, tais como o nosso propósito comunicativo, o contexto, com quem estamos interagindo ou se escrevemos ou falamos. Para o autor, a maioria das gramáticas está mais preocupada com considerações estruturais e em descrever a forma ou o significado das construções e categorias gramaticais do que em como elas são usadas no discurso falado ou escrito e isso acontece porque a gramática é vista como o estudo das estruturas de uma língua.

Uma das principais funções da linguagem é a comunicação com outras pessoas. É fato, entretanto, que seja estabelecido um divisor de águas entre aquilo que é uma boa comunicação ou não. Referindo-se especificamente à língua inglesa, Leech diz que:

...a língua não deveria ser avaliada de acordo com o tipo de regras gramaticais que segue, mas em termos de se a mensagem foi transmitida de forma eficaz. [...] é possível que se escreva ou fale de acordo com regras gramaticais do inglês padrão e ainda assim produzir uma linguagem opaca ou de difícil compreensão. Isso pode ser descrito como um 'estilo ruim'(Leech et al 1982: 12 tradução nossa)<sup>37</sup>.

Ver a língua como um instrumento de comunicação é ver a gramática, não como uma série de regras, mas como um recurso comunicativo, assim como o faz Halliday em sua Gramática Funcional.

---

<sup>37</sup>Tradução do original em inglês“ ...language should not be evaluated according to what type of grammatical rules it follows, but according to whether it conveys its message effectively.[...]It is quite possible, for example, to speak or write according to the grammatical rules of standard English, and yet produce language which is unclear or difficult to follow. This can be described as “bad style”.’(Leech et al 1982: 12)<sup>37</sup>”.

Todas as línguas organizam-se em tipos fundamentais de significado ou componentes, ou seja, um sistema lingüístico que permite a criação do texto: o ideacional ou reflexivo, manifestando o propósito de compreender o ambiente, a criar significados a partir de um contexto; o interpessoal ou ativo, manifestando o propósito de interagir com outros no ambiente; e o textual (texto escrito, falado ou as figuras) combinado a outros componentes, cria textos coerentes dentro de um contexto e representam as escolhas. O contexto é o que motiva a criação do texto. A língua muda de acordo com as diferentes situações, as chamadas, metafunções<sup>38</sup>.

A Gramática Funcional considera que a linguagem é o resultado das três macrofunções e que, por conseguinte, é um recurso comunicativo. Entretanto, mesmo a Gramática Funcional não utiliza exemplos a partir de frequência de uso. E nem mesmo extrai a totalidade de seus exemplos da linguagem em uso. Na última versão da *Introduction to Functional Grammar* de 2004, há exemplos extraídos de fontes tão diversas quanto David Copperfield (Halliday e Matthiessen, 2004: 78) e o *corpus* UTS/Macquarie (idem:179), incluindo-se na referida gramática muitos exemplos inventados.

Uma gramática que se diferencia do conceito tradicional de gramática é a *Longman Grammar of Spoken and Written English* (doravante LGSWE), de Biber et al (1999) que descreve o uso real de componentes gramaticais, a partir de sua frequência de uso em diferentes variedades (ou registros) do inglês tais como a conversação, a ficção, a linguagem jornalística e a prosa acadêmica.

A LGSWE adota uma abordagem baseada em *corpus*, o que significa que as descrições gramaticais são baseadas em padrões de estrutura e uso encontrados com determinada frequência em uma vasta quantidade de textos escritos e falados, armazenados eletronicamente e manipulados via computador. No total, o *corpus* usado no LGSWE contém cerca de 40 milhões de palavras, com textos representados em diferentes registros. Essas descrições baseadas na frequência de uso mostram que estrutura e uso não são aspectos independentes da língua. A análise de ambos é necessária para entender como a gramática (neste caso a gramática da língua inglesa) funciona nas interações do dia-a-dia de falantes e escritores.

A LGSWE não é isenta de problemas. Todo o *corpus* escrito e compilado advém de escrita publicada e o *corpus* oral advém de trocas entre “nativos”. Portanto, as estatísticas fornecidas refletem a o uso freqüente e a norma culta.

Por outro lado, em meio à discussão sobre as gramáticas de língua inglesa dentro de um estudo sobre escrita de aprendiz surge uma outra questão: que variedade da língua deveria ser

---

<sup>38</sup> Carvalho, nota de aula, 2009



usada como norma? No caso do inglês, em particular, essa é uma questão difícil de se responder uma vez que o inglês não é visto mais apenas como uma variedade de primeira e segunda língua, mas é visto como língua franca, com diferentes características ao redor do mundo (Nesselhauf 2005:37).

Alguns autores sugerem um inglês simplificado como norma (Kasper 1976; Quirk, 1982 *apud* Nesselhauf 2005). Outros sugerem que desde que todas as variedades de inglês como primeira ou segunda língua têm várias características em comum, o ensino deveria ser baseado nessas características. Recentemente, entretanto, vêm crescendo as opções pelo ensino de inglês como língua franca ou como uma língua internacional, ou seja, o inglês falado especialmente em ambientes internacionais principalmente entre falantes não nativos. Tal variedade, que ainda não foi descrita e cuja existência ainda não é clara, é vista como sendo uma amálgama de características do inglês produzido por nativos e por falantes de inglês como segunda língua ou como língua estrangeira.

Embora ainda algumas questões não estejam claras quanto ao modelo a ser adotado no ensino de uma língua estrangeira, nota-se que as normas descritivas mais do que prescritivas já são levadas em consideração e que também há uma crescente consciência de que o objetivo da grande maioria dos aprendizes é chegar ao nível de proficiência do nativo (Nesselhauf 2005).

Dessa forma, e na falta de um “modelo”, um aspecto que pode ser levado em consideração, no que diz respeito a ser proficiente em uma língua, é a habilidade de o aprendiz usar diferentes recursos, recursos esses que sejam de uso frequente na língua ou que façam parte de um repertório frequente em determinadas trocas verbais. A única maneira de avaliar se esses recursos são ou não são frequentemente usados por uma população semelhante e compara-los através de *corpus* construído a partir de condições semelhantes: usuários com características semelhantes, textos com propósitos semelhantes, contextos semelhantes. E é isso que pode ser feito através da Lingüística de *Corpus*.

A presente dissertação estuda um desses recursos, os advérbios, usados com frequência na escrita dentro do ambiente acadêmico (Biber et al, *op cit*) em toda a gama de funções que desempenha, através de uma abordagem funcional.

### 2.1.1. A gramática sistêmico-funcional

A gramática sistêmico-funcional desenvolvida por Michael Halliday e seus colegas tem um grande impacto no ensino de línguas. É diferente de todos os modelos de gramáticas anteriores uma vez que interpreta a linguagem como uma rede de escolhas significativas interrelacionadas e busca fornecer uma clara relação entre funções e sistemas gramaticais. (Halliday 2004).

Sob ponto de vista funcional, o texto é uma harmoniosa coleção de significados apropriados para um determinado contexto (Butt et al 2000). A ampla compreensão do texto é normalmente impossível sem se referir ao contexto em que ocorre. E o contexto pode ser considerado por meio de duas perspectivas: o contexto de cultura e o contexto de situação. O primeiro se refere ao amplo ambiente sociocultural que inclui a ideologia, as convenções sociais e as instituições; a última se refere a situações específicas dentro de um ambiente sociocultural (Droga e Humphrey 2002).

Com relação ao contexto de situação, todas as diferenças de situação entre textos pode ser explicadas por três aspectos de contexto, como o campo, a relação e o modo. O campo se refere àquilo que é falado ou escrito; a relação se refere às relações estabelecidas entre falante-ouvinte, e entre escritor-leitor; o modo se refere ao canal de comunicação (Butt op cit). Esses três aspectos refletem três metafunções da linguagem:

A metafunção ideacional ou experiencial, que é influenciada pelo campo, permite que as pessoas usem a linguagem para representar experiências;

A metafunção interpessoal, influenciada pela relação, permite que pessoas usem a linguagem para estabelecer relações sociais;

A metafunção textual, que é influenciada pelo modo, permite que as pessoas usem a linguagem para construir contextos lógicos e coerentes. Na abordagem sistêmico-funcional estas três metafunções co-existem simultaneamente.

## 2.2. Os Advérbios

Esta seção aborda de forma resumida, conceitos sobre os advérbios de língua inglesa sob a ótica da gramática funcional. Todos os exemplos apresentados, bem como as definições e termos utilizados, foram retirados das obras de Halliday e Matthiessen (2004) e Droga e Humphrey (2002) e serão adotados na discussão dos dados.

O advérbio, objeto de nosso estudo, pode desempenhar uma função específica, dependendo do texto e do contexto, dentro de qualquer das três metafunções.

<b>Metafunção</b>	<b>Tipo de adjunto</b>
Experiencial	Adjunto de circunstância
Interpessoal	Adjunto modal
Textual	Adjunto conjuntivo

Tabela 2: adaptado de Halliday e Matthiessen 2004:125

Por exemplo, na metafunção experiencial, o advérbio é visto enquanto circunstância, como **adjunto de circunstância**; na metafunção Interpessoal, o advérbio é visto enquanto sinalizador de comentários como **adjunto modal** e na Textual, o advérbio é visto enquanto recurso para unir orações como **adjunto conjuntivo**. Cada uma dessas funções exercidas pelos advérbios será explanada nas seções subsequentes.

### 2.2.1 Metafunção experiencial

A metafunção experiencial está relacionada com aquilo que acontece no mundo. Ela retrata os participantes de um processo, o processo propriamente dito e as circunstâncias através das quais o processo acontece.

As circunstâncias adicionam sentido às orações ao descreverem o contexto dentro dos quais os processos ocorrem. Uma circunstância é normalmente compreendida por um sintagma nominal ou um grupo adverbial (Droga e Humphrey 2002:31). Os tipos de circunstâncias, suas

subdivisões, bem como a pergunta-teste que se faz para escolher entre uma categoria ou outra encontram-se na tabela abaixo.

<b>Tipos de Circunstâncias</b>	<b>subcategorias</b>	<b>Subcategorias teste</b>
<b>Extensão</b>	Temporal	Quanto tempo?
	Espacial	Quão distante?
<b>Localização</b>	Temporal	Quando?
	Espacial	Onde?
<b>Modo</b>	Meio	Por quais meios?
	Qualidade	Como?
	Comparação	Como o quê?
<b>Causa</b>	Razão	Por quê?
	Propósito	Para qual propósito?
	Benefício	A benefício de quem?
<b>Contingência</b>	Condição	Sob quais condições?
	Concessão	A despeito de que?
	Falta	Falta o quê?
<b>Acompanhamento</b>	Comitativo	Com quem/o quê?
	aditivo	E quem/o que mais?
<b>Papel</b>	Guisa	Como o quê?
	Produto	Em que?
<b>Assunto</b>		Sobre o quê?
<b>Ângulo</b>		Quem disse?

Tabela 3: Adaptado de Martin et al (1997:104 *apud* Droga 2002:16)

*Ex: But you mustn't take it **personally**. [UTS/Macquarie Corpus]*

*Ex: Yeah, but it didn't **aggressively** market them. [UTS/Macquarie Corpus]*

*Ex: You've coped **beautifully** tonight, you've coped so well compared to how I would have coped.*

### 2.2.2. Metafunção Interpessoal

A metafunção interpessoal expressa o julgamento do falante e do escritor ou a atitude destes com relação ao conteúdo da mensagem. Dentro desta metafunção podem ser encontrados os adjuntos modais, que podem ser classificados em **adjunto de modo** e **adjunto de comentário**.

**a) adjuntos de modo:** são assim chamados porque estão associados com os sentidos construídos pelo sistema de modo que abrange a **modalidade**, a **temporalidade** e a **intensidade**.

Os adjuntos de modalidade estão inseridos dentro do campo da modalidade que se refere ao julgamento do falante ou o pedido de julgamento por parte do ouvinte diante da proposição. O que o sistema de modalidade faz é construir a região da incerteza que rodeia o sim e o não.

Em uma proposição, o sentido de negativo e positivo é asseverar e negar, positivo é *it is so* e negativo é *it isn't so*. Em termos de modalidade existem dois tipos de possibilidades intermediárias: os graus de probabilidade: *possibly/probably/certainly* e os graus de habitualidade: *sometimes/usually/always*.

As primeiras são equivalentes a *either yes* ou *either no*, ou seja, *maybe yes* ou *maybe no*. As segundas são equivalentes a *both yes and no*, ou seja, *sometimes yes*, *sometimes no*, com diferentes graus de frequência anexadas. São a essas escalas de probabilidade e habitualidade que o termo modalidade está atado e desta forma tem-se o processo de **modalização**.

Ao comparar *That's certainly John* com *That's John*, nota-se que a primeira oração expressa menos certeza do que a segunda. Assim como ao comparar *It always rains in Summer* com *It rains in Summer* nota-se que a primeira oração é menos variável que a segunda. Em outras palavras, você só diz que você está certo quando você de fato não está (Halliday 2004:147).

Em uma proposta, o sentido negativo e positivo é prescrever e proscrever: positivo é *do it*, negativo é *don't do it*. Aqui também, em termos de modalidade, há dois tipos de possibilidade intermediária, neste caso, dependendo da função do discurso (uma ordem ou um oferecimento): em uma ordem, o ponto intermediário representa graus de obrigação: *allowed to/supposed*

*to/required to: John's supposed to know that, Mary will help*; em um oferecimento, ele representa graus de inclinação: *willing to/anxious to/determined to: I'm anxious to help them*. A essas escalas de obrigação e inclinação dá-se o nome de **modulação**.

	<b>Média</b>	<b>Alta</b>	<b>Baixa</b>
Probabilidade	<i>Probably</i>	<i>Certainly, definitely, no way (no how)</i>	<i>Possibly, perhaps, maybe, hardly</i>
Habitualidade	<i>Usually</i>	<i>Always, never,</i>	<i>Sometimes, occasionally, seldom, rarely</i>

Tabela 4: advérbios que funcionam como adjuntos de modalidade (Halliday 2004:128)

Os adjuntos de temporalidade estão relacionados com o tempo interpessoal (dêitico). Eles podem ser o próprio tempo, que pode ser próximo ou remoto, passado ou futuro, relacionado com o falante no agora; ou podem ser uma expectativa, positiva ou negativa com respeito ao tempo em questão:

*Ex: They've been typed up and they were the responses to the various things which you had **already** raised in relation to the questions that we were asking in that short survey, so I thought you'd **probably** want to keep a record of that; and **already** when I look over this, I was absolutely astounded (UTS/Macquaire Corpus) (Halliday 2004:127).*

	Remoto	Próximo
Futuro	<i>eventually</i>	<i>soon</i>
Não-futuro (passado/presente)	<i>once</i>	<i>just</i>

Tabela 5: advérbios que funcionam como adjuntos de temporalidade (Halliday 2004:128)

Os adjuntos de intensidade se dividem em duas classes: a) expectativa, cujo grau pode ser total, alto ou baixo - *I entirely disagree* ou *I don't entirely disagree*. Esses adjuntos, em especial

os “totais” estão associados a interpessoalidade; b) contra-expectativa ou limitam ou excedem aquilo que se espera. O significado é ou *nothing else than, went no further than* ou ainda *went as far as* (Halliday 2004:127).

*Ex: They have **actually** calculated all the consequences of what they are doing with their tanks and planes in Berlin? [LOB\_B]*

*Ex: I **actually** didn't have a lot of chicken; I had **probably** more vegetables. [UTS/Macquarie Corpus]*

Grau	Total	<i>Totally, utterly, entirely, completely</i>
	Alto	<i>Quite, almost, nearly</i>
	Baixo	<i>Scarcely, hardly</i>
Contra-expectativa	Excede	<i>Even, <b>actually, really</b>, in fact</i>
	Limita	<i>Just, <b>simply</b>, merely, only</i>

Tabela 6: advérbios que funcionam como adjuntos de intensidade (Halliday 2004:129)

**b) adjuntos de comentário:** não há uma linha muito clara entre os adjuntos de comentário e os adjuntos de modo. A diferença é que adjuntos de comentário expressam a atitude do falante seja na proposição como um todo ou numa função particular do discurso. Em outras palavras, a carga do comentário pode ser ou ideacional ou interpessoal.

O tipo ideacional (proposicional) ocorre apenas em orações declarativas. Neste tipo, o comentário pode ser asseverativo (*it's so*) ou qualificativo (*this is what I think about it*).

*Ex: I think Dr Chatterji is **wisely** implying that reverence is open, faith is blind; reverence permits freedom, faith demands obedience (KOHL\_G) (Halliday 2004:131).*

O tipo interpessoal (discurso funcional) pode ocorrer tanto em orações declarativas quanto em orações interrogativas, mas com uma mudança de orientação: em orações declarativas, eles expressam o ângulo de quem fala enquanto que nas interrogativas, busca o ângulo do ouvinte:

Ex: **Frankly**, were you surprised at the outcome?

Ex: **Apparently**, he`s got a wife and a couple of kids [UTS/Macquarie Corpus]

Ex: **Specifically**, what have you been working on this evening Bruno?  
[UTS/Macquarie Corpus]

Tipo					Exemplos de advérbios	
Proposicional	No total	asseverativo	natural		<i>naturally, of course</i>	
			óbvio		<i>obviously, clearly</i>	
			certo		<i>indubitably, no doubt</i>	
		qualificativo	predição	previsível		<i>unsurprisingly, predictably</i>
				surpresa		<i>surprisingly, unexpectedly</i>
			presunção	rumor		<i>evidently, supposedly</i>
				argumento		<i>arguably</i>
				achismo		<i>presumably</i>
			vontade	vontade:sorte		<i>luckily, fortunately</i>
	vontade:esperança			<i>hopefully</i>		
	sem vontade			<i>sadly, unfortunately</i>		
	No sujeito		sabedoria	positiva		<i>wisely, cleverly</i>
		negativa		<i>foolishly, stupidly</i>		
		moralidade	positiva		<i>rightly, correctly</i>	
			negativa		<i>wrongly, unjustifiably</i>	
Discurso-funcional		desqualificado	persuasivo	asseverar	<i>truly, honestly</i>	
				concessão	<i>admittedly, certainly</i>	



	qualificado	fato		<i>actually, really, in fact</i>
		validade		<i>generally, ordinarily</i>
		engajamento pessoal	honestidade	<i>frankly, honestly</i>
			sigilo	<i>confidentially</i>
			individualidade	<i>personally</i>
			precisão	<i>truly, strictly</i>
			hesitação	<i>tentatively</i>

Tabela 7: advérbios funcionando como adjuntos de comentário (adaptada de Halliday 2004:130)

### 3.2.3. Metafunção Textual

A metafunção textual cria relevância no contexto. Dentro desta metafunção encontra-se os **adjuntos conjuntivos**, que estabelecem uma relação contextualizada com alguma parte do texto, normalmente, aquela a que sucede. A base semântica desta função contextualizada é a relação lógico-semântica da expansão, ou seja, a segunda oração expande a primeira seja elaborando-a, estendendo-a ou intensificando-a. Todas estas relações são construídas através da coesão. A tabela 8 abaixo representa os itens que funcionam como adjuntos conjuntivos.

*Ex: However, Mr. Daylal has made a representation to the commisioner...[KHOL\_A]*

*Ex: It was not, however, to be a precedent, he said. [KHOL\_A]*

*Ex: This did not happen, however. [KHOL\_D]*

Tipo de Expansão	subtipos	Itens
------------------	----------	-------

Elaboração	aposição	exposição		<i>in other words, that is, i mean, to put in another way</i>
		exemplificação		<i>for example, for instance, thus</i>
	clarificação	correção		<i>or rather, at least, to be more precise</i>
		distração		<i>by the way, incidentally</i>
		dispersão		<i>in any case, anyway</i>
		particularização		<i>in particular, <b>especially</b></i>
		resumo		<i>as i was saying, to resume,</i>
		sumarização		<i>in short, to sum up, briefly</i>
verificação		<i><b>actually</b>, as a matter of fact, in fact</i>		
Extensão	adição	positivo		<i>and, also, moreover</i>
		negativo		<i>nor</i>
		adversativo		<i>but, yet, however,</i>
	variação	alteração		<i>on the contrary, instead</i>
		subtração		<i>apart from that, except for that</i>
alternativo		<i>alternatively</i>		
Intensificação	Espaço-temporal	simples	seguinte	<i>then, next, afterwards</i>
			simultâneo	<i>just then, at the same time</i>
			precedente	<i>before that, previously</i>
			conclusivo	<i>in the end, finally</i>
		complexo	imediató	<i>at once, straightaway</i>
			interrompido	<i>soon, after a while</i>
			repetitivo	<i>next time, on another occasion</i>
			específico	<i>next day, an hour late</i>
			durativo	<i>meanwhile, all that time</i>
			terminal	<i>until then, up to that point</i>
			pontual	<i>at this moment</i>
		simples interno	seguinte	<i>next, secondly,</i>
			simultâneo	<i>at this point, here, now</i>
			precedente	<i>hitherto, up to now</i>
conclusivo			<i>lastly, finally</i>	
	modo	comparação	positiva	<i>likewise, similarly</i>
			negativa	<i>in a different way</i>
		meios		<i>thus, thereby, by such means</i>
	Causa-condição	geral		<i>so, then, <b>consequently</b>, hence, because of that</i>
		específico	resultado	<i>in consequence, as a result</i>
			razão	<i>for that reason</i>
			propósito	<i>for that purpose</i>

			condicional positivo	<i>then, in that case, in that event</i>
			condicional negativo	<i>otherwise, if not</i>
			concessivo	<i>yet, still, though, despite this, however</i>
	Assunto	positivo		<i>here, there, as to that, in that respect</i>
		negativo		<i>in other respects, elsewhere</i>

Tabela 8: itens que funcionam como adjuntos conjuntivos (adaptada de Halliday 2004:543)

Em resumo, “a teoria da gramática sistêmico-funcional é a resposta à ‘miopia’ da corrente que vê a linguagem como um conjunto de regras<sup>39</sup>”. Usa a perspectiva da linguagem enquanto um recurso ou um sistema ao invés de olhar fragmentos. Por essa razão, as categorias a serem adotadas nesta dissertação são as categorias funcionais porque são elásticas, são dependentes do texto e são baseadas em funções e não em lugares em que a palavra ocupa na frase.

Portanto, tendo estabelecido o pano de fundo da presente dissertação, o próximo capítulo, sobre métodos de pesquisa, terá como objetivo explicitar as bases metodológicas adotadas no processo de investigação dos advérbios formados por sufixo *-ly*.

<sup>39</sup> (Shepherd 2010 comunicação oral)

### 3. METODOLOGIA

Neste capítulo metodológico será discutida a escolha dos *corpora* de estudo e referência, bem como o emprego das ferramentas de análise. Também será abordado o tratamento que foi dispensado ao tópico da pesquisa, os advérbios em *-ly*. São descritos os métodos de entrada nos *corpora*, e os parâmetros para a análise de qualquer palavra de busca propostos por O’Keefe, que são usados para os advérbios em *-ly* nesta dissertação. Somente após estas considerações será apontada a abordagem adotada na dissertação, seja ou a partir de listas de palavras, de colocados, de coligados, de clusters, ou agrupamentos lexicais, ou da criação de um perfil léxico-gramatical.

#### 3.1. Materiais: os *Corpora*

Nesta dissertação foram usados três *corpora*,: um de estudo (denominado Br-ICLE) e dois de referência (denominados LOCNESS e BAWE). O propósito de utilizar esses *corpora* é contrastar aspectos da linguagem utilizados por aprendizes brasileiros de inglês com aqueles utilizados por falantes de inglês como língua materna, observando, portanto, suas semelhanças e diferenças. Dessa forma, o presente trabalho utiliza a abordagem direcionada pelo *corpus* uma vez que o objetivo aqui é observar um dos fenômenos linguísticos que ocorrem nos *corpora*, ou seja, o uso de advérbios. A partir dessa observação, pretende-se traçar o perfil das populações investigadas. .

##### 3.1.1. O *corpus* de estudo<sup>40</sup> : o que é e como está composto

---

<sup>40</sup> A descrição dos textos contidos no *corpus* encontra-se no apêndice desta dissertação.

O *corpus* de estudo, denominado Br-ICLE (Brazilian *International Corpus of Learner English*), é composto por ensaios argumentativos em língua inglesa escritos por universitários brasileiros. Contém 127 textos, com o total de 65.304 palavras, sendo considerado, dessa forma, um *corpus* pequeno, de acordo com a classificação de Sardinha (*ver capítulo 3 seção 3.4.2 desta dissertação*). O *corpus* Br-ICLE faz parte de um projeto maior chamado ICLE (*International Corpus of Learner English*), coordenado pela Universidade Católica de Louvain na Bélgica, na pessoa de Sylviane Granger, especialista em *corpora* de aprendizes de inglês (*learner corpora*). O projeto conta com a participação de grupos de pesquisadores de quatorze países: França, Alemanha, Holanda, Espanha, Suécia, Finlândia, Polônia, República Tcheca, Bulgária, Rússia, Itália, Israel, Japão e China. Cada país envolvido no projeto é responsável pela formação, organização e informatização de um *subcorpus* local.

O Centro de Lingüística de *Corpus* de Louvain (*The Louvain Centre for English Corpus Linguistics*), por ser o primeiro no gênero, tem desempenhado um papel pioneiro na implementação de *corpora* informatizados de inglês de aprendizes de vários países do mundo. O banco de textos do ICLE, resultado de oito anos de trabalho colaborativo envolvendo universidades de todo o mundo, é composto por redações em língua inglesa, feitas por alunos de nível avançado de 14 nacionalidades diferentes: Cada *subcorpus* nacional possui cerca de 200 mil palavras, que embora não pareça muito, se comparado com os *corpora* de língua de materna, é muito significativo em termos de pesquisas na área de aquisição de segunda língua. O banco contém, até o momento, mais de dois milhões de palavras e representa o inglês de aprendizes estrangeiros na sua modalidade escrita. Um dos objetivos da análise desses *corpora* é descrever e comparar as tendências observadas em cada uma das nacionalidades envolvidas no projeto.

O ICLE é um *corpus* compilado segundo dois conjuntos de critérios, um compartilhado e um variável. O primeiro é comum a todas as variedades e possui valores fixos (controlados) em relação aos aprendizes no que diz respeito a idade (cerca de 20 anos), contexto educacional (EFL – inglês como língua estrangeira), nível de proficiência (avançado); quanto aos textos (escrito), a tecnicidade (não-técnico) e gênero (ensaios argumentativos). Quanto ao segundo critério, as variáveis são: sexo, língua materna, região, conhecimento de outras línguas estrangeiras, experiência prática, tópico e ambiente da tarefa, e variam de acordo com a variedade nacional. Além disso, os textos possuem extensão fixa, cerca de 500 palavras, mas são todos integrais, isto é, nenhum foi mutilado para caber no limite de palavras (Berber Sardinha, 2004).

A importância do ICLE reside na possibilidade de tornar a investigação de *corpora* de aprendizes um projeto viável. Em termos de estudos digitais sobre a escrita em língua inglesa produzida por aprendizes de várias nacionalidades, o Projeto ICLE está em franco desenvolvimento. Através do projeto, são estudadas as variedades de inglês escrito como o inglês-espanhol (Sp-ICLE), o inglês-francês (F-ICLE), e o próprio Br-ICLE, a ser explicado abaixo.

O *corpus* Br-ICLE é subordinado ao ICLE e, desta forma, segue os padrões e fundamentos de compilação de *corpora* eletrônicos, temáticas e gênero para a produção escrita em língua inglesa estabelecidos pelo ICLE. Portanto, são compiladas composições argumentativas obedecendo a um número restrito de temas, com um número mínimo de palavras, as quais são escritas por universitários, todos estudantes de língua inglesa.

No Brasil, este sub-projeto é intitulado “A Coleta de Sub*Corpus* de Aprendizes de Inglês no Brasil” tradução nossa<sup>41</sup>, cujo grupo oficialmente responsável pela sua montagem é coordenado pelo Prof. Dr. Tony Berber Sardinha, da PUC-SP, responsável também pelo grupo GELC/Cnpq, Grupo de Estudos em Linguística de *Corpus*. Para localizar o *subcorpus* de português brasileiro, o Br-ICLE, é só acessar <http://lael.pucsp.br/corpora/bricle>, mas este *site* está aberto somente aos participantes do projeto aqui no Brasil. Vale dizer que o Br-ICLE ainda está em fase de coleta, por isso o número total de palavras.

#### 4.1.2. Os *corpora* de referência<sup>42</sup>: o que são e como estão compostos

O *corpus* de referência LOCNESS é composto por ensaios argumentativos em inglês compilados em três *subcorpora* diferentes. O primeiro contém 113 textos ,com um total de 60.129 palavras, provenientes de redações de pré-universitários ingleses, cursando os chamados A-LEVELS. . O segundo *subcorpus* contém 90 textos de universitários ingleses com cerca de 95.482 palavras no total. O terceiro *subcorpus* contém 205 textos de universitários americanos com 168.303 palavras. No total, o *corpus* LOCNESS, considerado médio, contém 408 textos e

<sup>41</sup> *The collection of a Sub-Corpus of Learner English in Brazil*

<sup>42</sup> Os dados descritos nesta dissertação a respeito do LOCNESS e do BAWE fazem parte da documentação dos *corpora*. A descrição dos textos contidos nos *corpora* encontra-se no apêndice desta dissertação.

totaliza 323.914 palavras. O *corpus* LOCNESS é vendido comercialmente pela Universidade de Louvain e tem como objetivo ser usado como *corpus* de referência nos estudos individuais dos *corpora* do ICLE. O *corpus* contém todos os problemas que um *corpus* oriundo de universitários, falantes de uma língua específica como uma língua materna normalmente conteria.

O outro *corpus* de referência, denominado BAWE (*The British Academic Written English*<sup>43</sup>), considerado médio-grande, é formado por ensaios argumentativos de universitários ingleses, em nível de graduação e pós-graduação, já corrigidos (prontos para publicação), das diferentes faculdades das Universidades participantes. Os textos foram separados em categorias de acordo com a faculdade, o departamento, o módulo do curso e o ano de estudo. Ele contém 2.761 textos com um total de 6.742.318 palavras. O *corpus* completo, contendo 4.000 textos de alunos, pode ser baixado ou acessado gratuitamente através de um concordanciador *online*.

O *corpus* BAWE é resultado de um projeto de pesquisa intitulado “*Uma investigação dos gêneros textuais na educação superior Britânica*” (tradução nossa<sup>44</sup>), realizado no período entre 2004 e 2007 nas universidades de Oxford Brookes, Reading e Warwick. O Centro Internacional para Estudos da Língua Inglesa (tradução nossa<sup>45</sup>) do Instituto de Educação Westminster recebeu fundos para desenvolver um *corpus* eletrônico da escrita proficiente de alunos.

O principal objetivo do *corpus* é permitir identificar as características da escrita de alunos e compará-la em diferentes disciplinas e em diferentes níveis de estudo. O *corpus* Bawe pode também ser usado, como é o caso da presente dissertação, como um *corpus* de referência, visto que os textos estão corrigidos. Portanto, as amostras do *corpus* podem ser eventualmente utilizadas em materiais de ensino.

## 3.2 Métodos

### 3.2.1 O Ferramental de Análise: o programa WordSmith Tools

---

<sup>43</sup> A Escrita Acadêmica do Inglês Britânico

<sup>44</sup> Tradução do original: *An Investigation of Genres of Assessed Writing in British Higher Education*

<sup>45</sup> Tradução do original: *The International Centre for English Language Studies*

Wordsmith Tools é uma compilação de programas que disponibilizam vários recursos essenciais na análise de diferentes aspectos linguísticos, como a composição lexical, a temática de textos selecionados e a organização retórica e composicional de gêneros discursivos. O WordSmith Tools, foi idealizado por Mike Scott e pode ser baixado em versão demo através do [www.lexically.net](http://www.lexically.net). Nesta pesquisa foi utilizada a versão 3.0 em sua forma integral.

Dentre os vários recursos que o Wordsmith Tools disponibiliza, serão apresentados, ainda que brevemente, apenas três programas que foram usados nesta pesquisa como o *Wordlist* (lista de palavras), o *KeyWords* (lista de palavras-chave) e o *Concord* (lista de concordância).

O conceito básico do *Wordlist* é a criação de listas de palavras. O programa é pré-definido para produzir listas de palavras organizadas por ordem alfabética e por frequência (da palavra mais freqüente para a menos freqüente). Anexada a essas duas listas surge uma janela contendo os dados estatísticos relacionados aos dados das listas de palavras, no que diz respeito à densidade lexical, ao número de itens (*tokens*) e formas (*types*), à quantidade de parágrafos e sentenças. Por meio desta ferramenta é possível observar a partir de diferentes *corpora*, diferentes variedades linguísticas em diferentes contextos de uso.

Como exemplo, segundo O’Keeffe (2007), de acordo com o *corpus* LIBEL, as palavras mais frequentes da língua escrita em inglês são os artigos *a* e *the*, que indicam alta frequência de frases nominais; a preposição *of* sugerindo pósmodificação de frases nominais; *that*, apontando para multifuncionalidade, como subordinador e como pronome relativo; preposição *to*, *for* e *in* sugerindo frases preposicionais. Por outro lado, há a falta dos pronomes interativos *I* e *you*; o único que aparece na lista dos dez mais frequentes é o pronome *it* que é referencial em oposição à interatividade dos marcadores discursivos como *yeah*, *like* e *now*.

A partir da ferramenta *Wordlist*, também é possível fazer um levantamento de *clusters* (grupos lexicais compostos de 2, 3, n palavras, também chamado de *bundles* ou *n-gramas*), que permite observar como as palavras se agrupam, além de descrever o vocabulário de uma língua.

Por meio do *KeyWords*, é possível fazer uma seleção de itens de uma lista de palavras (ou mais) por meio da comparação de suas frequências com uma lista de referência. O resultado do contraste é uma lista de palavras-chave, ou palavras cujas frequências são estatisticamente diferentes no *corpus* de estudo e no *corpus* de referência. Uma palavra é considerada chave se ela ocorrer tantas vezes quanto for especificada pelo pesquisador como frequência mínima e quando a sua frequência no texto de estudo em comparação ao texto de referência é tal que a



probabilidade estatística como computada por um procedimento padrão é menor ou igual aquele especificado pelo pesquisador.

A ferramenta *Concord* produz concordâncias ou listas de ocorrências de um item determinado (palavra de busca) acompanhado do seu cotexto, ou seja, com uma parte do texto que fica ao seu redor. Além disso, ela produz também uma lista de colocados (*collocates*), ou seja, a lista de palavras que ocorrem em diferentes posições ao redor da palavra de busca. Pode-se também analisar a lista de agrupamentos lexicais que são as seqüências fixas de palavras recorrentes na concordância e a lista de padrões de colocados dispostos de acordo com a frequência. Essa ferramenta disponibiliza ainda um gráfico no qual as ocorrências das palavras de busca são marcadas por meio de um traço (|). Uma linha de concordância com a palavra *across* como palavra de busca teria a seguinte forma:

The Post Office is | across | the road from the Mall.  
She is trying to get | across | the lake.

### 3.2.2. Trabalhando com proporções

Nesta dissertação serão comparadas as frequências de advérbios em *-ly* em três *corpora* que, por sua composição diferem consideravelmente em termos de tamanho. A contagem simples da frequência de palavras teria sido uma abordagem útil na quantificação desses dados linguísticos, já que é freqüentemente usada nas pesquisas com base em *corpus*. Entretanto, ela constitui uma desvantagem na medida em que os *corpora* que se pretende comparar são diferentes no tamanho.

Existem inúmeras maneiras de se calcular se a diferença entre números inteiros obtidos na frequência de itens lexicais de *corpora* diversos é ou não é maior ou menor. O cálculo da proporção é a forma mais comum: através de um cálculo simples: R (razão) = n° de ocorrências de um determinado elemento linguístico / o n° total de palavras do *corpus*. O resultado normalmente é expresso em decimais. McEnery & Wilson (1996:68 tradução nossa)<sup>46</sup> dizem que

---

6.<sup>46</sup>Tradução do original em inglês: “The most common way of representing proportions in empirical linguistics are percentages.”

“a maneira mais comum de se representar proporções na linguística empírica é através de porcentagens”.

### 3.3.Os advérbios analisados

O foco de análise dessa dissertação são os advérbios derivados por sufixo, mais especificamente aqueles terminados em *-ly*, que de acordo com Biber et al (1999:540), representam 55% dos advérbios utilizados na prosa acadêmica, enquanto que um pouco mais de 30% dos advérbios são simples.

Em inglês arcaico, os advérbios eram derivados de adjetivos pelo acréscimo de *-e* ou *-lice*. Como resultado de processos fonológicos, o sufixo *-e* desapareceu e assim os adjetivos e os advérbios passaram a ter a mesma forma e o sufixo *-lice* evoluiu para o atual *-ly*. Alguns advérbios ainda têm a mesma forma correspondente dos adjetivos, por exemplo, *hard*, *long*, *early*, *daily kindly*. Em outros casos é possível ter ambas as formas, com ou sem *-ly*, embora às vezes, possa haver diferença de significado (Leech et al.1982:144). Além disso, alguns advérbios simples vêm de uma origem composta embora os significados das partes que os compõem não sejam mais transparentes, por exemplo, *already*, *indeed* (Biber et al. 1999:539). É provável que o advérbio *only* também se enquadre neste perfil, porém já que a sua derivação pelo sufixo *-ly* não é clara, optou-se por excluir este advérbio do campo de análise apesar de ele ter sido o de maior frequência em todos os três *corpora*, tanto o de estudo, quanto os de referência.

Para obter uma lista destes advérbios a partir dos *corpora*, foi necessário primeiro fazer uma lista de palavras em cada um dos *corpora*, com o auxílio do programa *WordSmith Tools versão 3* e dos seguintes comandos:

No *Controler*, (primeira janela que aparece ao se inicializar o programa) clicou-se em *Tools* e em seguida em *WordList*. Na janela *WordList* clicou-se em *File* e depois em *Start*. Na janela *Choose Texts*, clicou-se na pasta que continha o *corpus* de aprendizes e depois em *Ok*. Por fim, na janela *Getting Started*, clicou-se em *Make a WordList Now*. O programa inicia o levantamento da lista de palavras e ao final do processo, o programa apresenta três janelas com três listas de palavras organizadas em termos de estatística, frequência e em ordem alfabética. Na janela organizada em ordem alfabética, clicou-se em *Re-Sort* e em seguida em *Reverse Sort*. Este

comando possibilitou organizar as palavras em ordem alfabética de acordo com as suas terminações, conforme ilustrada na figura 2 a seguir.

N	Word	Freq	%	Lemmas
5642	WHY	49	0,08	
5643	ANARCHY	1		
5644	HIERARCHY	1		
5645	GEOGRAPHY	1		
5646	PHILOSOPHY	9	0,01	
5647	PHYLOSOPHY	1		
5648	HEALTHY	5		
5649	WEALTHY	7	0,01	
5690	UNHEALTHY	1		
5651	WORTHY	4		
5652	TRUSTWORTHY	1		
5653	SILKY	2		
5654	FLY	1		
5655	SCIENTIFICALLY	1		
5656	REALY	1		
5657	OFFICIALY	1		
5658	INFORMALY	1		
5659	EQUALY	1		
5660	PROBABLY	20	0,03	
5661	INEXTRICABLY	1		
5662	REMARKABLY	1		
5663	PRESUMABLY	1		
5664	UNQUESTIONABLY	1		
5665	REASONABLY	1		
5666	INESCAPABLY	1		
5667	CONSIDERABLY	2		
5668	PREFERABLY	1		
5669	INDUBITABLY	1		
5670	COMFORTABLY	1		
5671	REGRETTABLY	1		
5672	INCREDIBLY	1		
5673	POSSIBLY	6		
5674	IRRESISTIBLY	1		
5675	BADLY	4		
5676	SADLY	1		
5677	GLADLY	1		

Figura 2: lista de palavras do Br-ICLE

A etapa seguinte foi eliminar as palavras que não se encaixavam no perfil de advérbios, como por exemplo, a palavra *family* (que embora termine em *-ly* não é um advérbio)<sup>47</sup>. Terminado este processo, ainda na janela do *WordList*, clicou-se novamente em *Re-sort* e desmarcou-se a opção *Reverse Sort*. O resultado desse processo é apresentado na figura 3 a seguir:

<sup>47</sup> Como alternativa para esse método quase que manual, os corpora poderiam ter sido etiquetados, entretanto, essa tarefa só poderia ser bem sucedida com os corpora Br-ICLE e LOCNESS devido ao tamanho destes. A mesma tarefa com o corpus BAWE seria quase impossível já que os etiquetadores disponíveis hoje suportam somente arquivos com o tamanho de 2Mb.

N	Word	Freq.	%	Lemmas
1	ABSOLUTELY	1		
2	ACCORDINGLY	2		
3	ACTIVELY	1		
4	ACTUALLY	24	0,04	
5	ADDITIONALLY	1		
6	ADMITTEDLY	1		
7	AGRESSEVILY	1		
8	AMAZINGLY	1		
9	APPARENTLY	4		
10	APPEARENTLY	1		
11	APPROXIMATELY	1		
12	BADLY	4		
13	BARELY	3		
14	BASICALLY	3		
15	BLINDLY	1		
16	BRIEFLY	1		
17	BROADLY	2		
18	CAREFULLY	6		
19	CERTAINLY	20	0,03	
20	CERTAINLY	1		
21	CLEARLY	10	0,02	
22	CLOSELY	1		
23	COINCIDENTALLY	1		
24	COMFORTABLY	1		
25	COMMONLY	4		
26	COMPLETALLY	1		
27	COMPLETELY	21	0,03	
28	COMPLETELY	2		
29	COMSEQUENTLY	1		
30	CONSEQUENTLY	24	0,04	
31	CONSIDERABLY	2		
32	CONSTANTLY	6		
33	CONTENTEDLY	1		
34	CONTINUOUSLY	1		
35	CORRECTLY	1		
36	CRITICALLY	4		

Figura 3: Lista de advérbios em ordem alfabética terminados em *-ly* (amostra do *corpus* Br-ICLE).

Concluída esta etapa, foi necessário utilizar a ferramenta *KeyWords*, para que assim, pudessem ser levantadas as listas de advérbios-chave nos *corpora*. Para se obter uma lista de advérbios-chave foi preciso realizar os seguintes comandos:

No *Controler* (a primeira janela que aparece ao inicializar o programa), clicou-se em *Tools* e depois em *KeyWords*. Na janela *KeyWords*, clicou-se em *File* e depois em *Start*. Na janela *Choose WordList*, selecionou-se os arquivos que continham a lista de advérbios produzidas pelo *WordList*: na janela à esquerda, a lista do *corpus* de estudo e na janela à direita, a lista do *corpus* de referência.. Por fim, clicou-se me OK. Com isso foi possível observar quais advérbios são mais ou menos frequentes, quais são sobreusados ou subusados.

Todavia, por causa da natureza do presente trabalho, uma dissertacao de mestrado, a análise se concentrou apenas nos 5 primeiros compartilhados pelos *corpora*, tendo como base, aqueles encontrados no Br-ICLE, ou seja, *really*, *simply*, *actually*, *consequently* e *especially*. Logo, a partir da lista obtida e da classificação dos advérbios de acordo com as gramáticas (*ver capítulo 3 desta dissertação*), foi estabelecido qual seria o tratamento dispensado aos advérbios

encontrados, a partir de uma abordagem direcionada pelo *corpus* (*data-driven*), isto é, olhando como os advérbios são usados, com que frequência são usados e comparando um *corpus* com outro, seja em termos de colocados, de padrões observáveis ou pela criação de um perfil léxico-gramatical.

Por fim, o programa *Concord*, possibilitou visualizar os advérbios encontrados dentro do contexto em que ocorrem. Para ter essa visualização, na tela do *Controler*, clicou-se em *Tools* e em seguida em *Concord*; na janela do *Concord*, clicou-se em *Settings* e depois em *Choose Texts* para selecionar os arquivos contendo os *corpora* de estudo e de referência e em seguida clicou-se em *OK*. Depois, foi necessário voltar em *Concord*, *Settings* e finalmente *Search Word*, onde foi digitado o elemento de busca, por exemplo, o advérbio *really* na caixa de diálogo *Search Word or Phrase*. Esse comando possibilita listar apenas as ocorrências deste advérbio. Finalmente clicou-se em *Go Now* para iniciar a criação das linhas de concordância onde o elemento de busca fica centralizado. A figura 4 a seguir mostra como é apresentada as linhas de concordância do advérbio *really* no *corpus* Br-ICLE. Além das linhas de concordância, com a ferramenta *Concord* também foi possível observar quais foram os padrões de colocados (*collocates*) e agrupamentos lexicais (*clusters*) recorrentes nos *corpora* dentro do universo da palavra de busca. Como a presente pesquisa analisou as ocorrências de colocados imediatamente à esquerda (doravante 1L) e à direita (doravante 1R) do elemento de busca, foi necessário alterar as configurações da ferramenta *Concord* para que só fossem mostradas as ocorrências de colocados 1L e 1R. Para isso, na tela do *Concord*, clicou-se em *Settings* e em seguida, no campo *Collocates*, foi marcado o horizonte de busca dos colocados, ou seja, 1L e 1R. Finalmente clicou-se em *Ok* para validar a configuração. Para exibir a lista de colocados, clicou-se no ícone *show collocates* e em seguida, para exibir a lista de agrupamentos lexicais, clicou-se no ícone *clusters*. Para esta dissertação foram analisados os agrupamentos lexicais formados por três palavras. O resultado de cada comando está ilustrado nas figuras 5 e 6 a seguir:

N	Concordance	Set	Tag	Word No.	File	%
1	... really denotes that man are not equal. People really fought against it, mainly during the m	i		142	a1109.txt	47
2	... the streets and using of alternative fuels are really effective. Concerning this fact, it wo	i		454	a1084.txt	75
3	... should be revised. An individual item that really deserves to be revised are the laws r	i		628	a1059.txt	83
4	... strength of this link. In many cases, this is a really hard chore, because lots of women h	i		304	a1059.txt	41
5	... mes to turn it off. Not to mention those who really get addicted to TV shows, soup oper	i		320	a1015.txt	57
6	... veloped countries. As crime has become really hard to control, it feels like people ha	i		253	a1080.txt	49
7	... ty that the heir of the private property would really be the son of the owner. This happen	i		245	a1061.txt	33
8	... has its own collaboration to mankind. Men really are physically stronger than women.	i		207	a1062.txt	51
9	... ologies, which shows that some people are really "more equal than others". Somethin	c		358	a1036.txt	61
10	... iers, wars or social differences, real world is really cold and cruel for those who are part	i		496	a1036.txt	86
11	... oughts and ideas, which can be something really dangerous. In this sense, perhaps rel	i		346	a1015.txt	62
12	... ight seem a little crazy from my part, but I really believe that even the world in which w	i		441	a1041.txt	81
13	... ve been fighting for. However, I wonder if we really want this equality. We want the sam	c		181	a1081.txt	33
14	... e minds of those who know what TV is. It is really sad to accept that people let their l	i		725	a1018.txt	96
15	... tional role. More than ever, what universities really need is to learn teaching and practi	i		262	a1025.txt	54
16	... d a whole nation to believe that the war was really necessary. Secondly, government clai	i		130	a1065.txt	15
17	... ple's lives, and in some places this system really works. As was demonstrated befor	i		419	a1033.txt	67
18	... uropeans, university degrees insofar as are really theoretical are of great value.	i		567	a1106.txt	99
19	... men involved in the crime are free. This fact really marks Brazilian law ineffectuality to d	p		308	a1093.txt	63
20	... does not focus on the real world. I think it is really important to offer students the entire	i		491	a1085.txt	83
21	... de their kids by themselves. In fact, what is really harmful for kids is lack of attention fro	i		322	a1059.txt	43
22	... ty after some of us, called feminists, fought really hard for our rights. Indeed, it seems t	i		41	a1014.txt	9
23	... erism, egoism, and many other "isms" are really making people incapable of making t	i		50	a1036.txt	9
24	... The kids are growing up learning that what really matters in life is: sex, violence and th	i		255	a1013.txt	59
25	... Otherwise, there were some acts that were really made by people who knew exactly w	i		229	a1086.txt	35
26	... ncreasing in Brazil? Does crime pay here? I really think that it pays, and it is so for thre	p		98	a1093.txt	21
27	... r are able to set people's rights, then do we really have to consider them as equals, sin	c		160	a1020.txt	28
28	... indispensable in the modern world, can be really harmful. Furthermore science techn	c		176	a1052.txt	33
29	... ilitation should come hand in hand but what really happens is that they come in disson	c		564	a1045.txt	87
30	... lves a few questions: is a university degree really an essential requirement without whi	c		161	a1047.txt	20
31	... nd fast-food. This sci-fi kind of world can be really helpful at times, but how far it will get	c		90	a1022.txt	16
32	... e who can and those who can't. And what really makes me be sick and tired of all t	c		402	a1008.txt	78
33	... to the law and to human rights, all men are really equal and have the same rights. Non	c		193	a1088.txt	36
34	... d so many options of TV programs that it is really difficult sometimes to turn it off. Not t	c		308	a1015.txt	55
35	... ven if some of the most combative feminists really contributed to create an atmosphere	c		525	a1021.txt	85
36	... lity of men and women". Does this equality really exist? That is exactly what women h	c		165	a1081.txt	30

Figura 4: Exemplo de lista de concordância do advérbio *really* (amostra do *corpus* Br-ICLE).

N	WORD	TOTAL	LEFT	RIGHT	L1	*	R1
1	REALLY	52	0	0	0	52	0
2	TO	8	1	7	1	0	7
3	ARE	7	6	1	6	0	1
4	IS	7	5	2	5	0	2

Figura 5: lista de colocados (amostra do *corpus* Br-ICLE)

N	cluster	Freq.
1	it is really	3

Figura 6: Lista de agrupamentos lexicais (amostra do *corpus* Br-ICLE)

### 3.4. Métodos de Análise

Conforme já foi descrito, com a ferramenta *Concord* é possível observar as linhas de concordância de um *corpus* e identificar muitos exemplos de **padrões** de uso de um determinado elemento linguístico

As pesquisas recentes baseadas em *corpus* (Scott e Tribble, 2006) sugerem que os padrões nos quais as palavras ocorrem, tanto para a sua direita ou para a esquerda, são de fato importantes para a sua interpretação. Padrões são importantes também para comparar o significado e a função de itens lexicais aparentemente similares. É uma prática comum em dicionários a identificação de significados apresentando sinônimos. Entretanto, itens que parecem sinônimos podem ser usados com diferentes funções pragmáticas em diferentes contextos. Tognini-Bonelli (*apud* Vandenberg & Aijmer 2007) mostra que embora *largely* e *broadly* apareçam como tendo o mesmo significado, eles se diferenciam por seus padrões de co-seleção e seus perfis colocacionais. Isso quer dizer que, observar padrões nos permite compreender que alguns advérbios se “casam” bem com algumas palavras mais do que com outras<sup>48</sup>. Desta forma, a análise sintagmática de qualquer palavra determina o seu significado. Por exemplo, alguns advérbios co-ocorrem tipicamente com marcadores discursivos como em: *But surely from the*

<sup>48</sup> De acordo com Saussure (*apud* Vandenberg & Aijmer 2007:10) o sentido das palavras só pode ser descrito em termos de oposição e diferença entre os seus termos vizinhos em um sistema linguístico

*point of view of the farmer it's it's all to do with the hard ecu and and the hard facts of of driving tractors across large field, isn't it? (ICE-GB:S1B-037/21)(Aijmer 2007).*

Tais padrões determinam a interpretação do advérbio de formas distintas. Por exemplo, a ocorrência freqüente de *certainly* seguida de um *but-clause* “empresta” ao advérbio o sentido de concessão : *I certainly believe that the police system in the United Kingdom and in England and Wales should be nationally organized but I also believe that uh many of its services should be locally delivered<, > (ICE-GB:S1B-033/6)(Aijmer 2007).* No nível sintagmático, é levada em consideração a influência do contexto de uma forma mais abrangente, ou seja, de que forma são estabelecidas as relações de lógica entre os termos que os antecedem e como eles ocorrem ou co-ocorrem com outros advérbios ou marcadores discursivos.

Uma outra forma de análise possível, sugerida por O’Keeffe (2007), e também empregada nesta dissertacao, é através da criação de um **perfil léxico-gramatical** da palavra de busca a partir das suas linhas de concordância. O perfil léxico-gramatical descreve contextos típicos em termos de:

**a) Colocados:** que palavras ocorrem mais freqüentemente com a palavra de busca e com que relevância no ambiente lexical estudado;

**b) Agrupamentos lexicais:** se a palavra faz parte de um determinado grupo de palavras e quais são os tipos que ocorrem;

**c) Padrões sintáticos** (chamado por alguns autores de ‘coligação’) : se existe algum padrão sintático que restrinja essa palavra. Por exemplo, existem preposições que acompanham a palavra? Qual é a posição da palavra na frase? Início, meio ou fim? Existe alguma restrição quanto ao tempo usado, no caso de ser um verbo?

**d) Padrões semânticos:** existe alguma restrição? Por exemplo, a palavra é aplicada somente a seres animados? Como é usada? Nunca é usada como intensificador? Restringe o quê?

**e) Prosódia semântica:** termo usado por Louw (1993 *apud* O’Keeffe 2007) e significa simplesmente que as palavras, uma vez tendo colocados típicos ( por exemplo, *blond* que é um colocado típico de *hair*, mas não de *car*), tendem a ocorrer em ambientes próprios, de forma que seus significados, especialmente seus significados conotativos e atitudinais, parecem abranger muitas outras palavras. e algumas palavras tendem a ocorrer em contextos negativos ou positivos.

Stubbs (1995 *apud* O’Keeffe 2007) mostra que mais de 90% dos colocados de *cause* são negativos, por exemplo, *accident, cancer, commotion, crisis* e *delay*. Em termos de prosódia



semântica positiva, a palavra *provide* é colocado de *care, food, help, jobs, relief* e *support*. Antes do advento da análise computadorizada da linguagem, esse fenômeno não havia sido codificado em termos de linguagem em uso.

### 3.5. Conclusão

Neste capítulo foram mostrados os materiais e métodos de abordagem ao estudo de advérbios terminados em *-ly* em língua inglesa, com vistas a comparar os resultados em *corpora* formados de textos de aprendizes brasileiros de língua inglesa, textos não editados escritos por vestibulandos e alunos de graduação falantes de língua inglesa como língua materna, além de textos editados produzidos por alunos de graduação e pós-graduação falantes de inglês como língua materna.

O próximo capítulo, análise de dados, apresenta uma análise contrastiva da **frequência** desses advérbios nos *corpora* e **como** esses advérbios são usados em cada um deles bem como os **padrões**, sejam de colocados, de coligados e de agrupamentos lexicais ou ainda em termos de prosódia semântica.

#### 4. ANÁLISE DOS DADOS

Neste capítulo, tem-se por objetivo a apresentação e discussão dos resultados obtidos por meio da análise de ocorrências de advérbios formados pelo sufixo *-ly*, extraídas dos *corpora* Br-ICLE, LOCNESS e BAWE, conforme explicado no capítulo quatro sobre Metodologia. A análise, entretanto, ficou restrita apenas aos 5 primeiros, tendo como base o *corpus* Br-ICLE, ou seja, *really*, *simply*, *actually*, *consequently* e *especially*. Isto se deu por conta do escopo e da extensão do presente trabalho. Cada seção que trata dos advérbios individualmente está dividida em duas partes: a primeira, está relacionada com a função de cada advérbio e a segunda traça o seu perfil léxico-gramatical. É importante ressaltar que todas as classificações aqui propostas, são baseadas nas evidências de colocações e coligações advindas dos *corpora* e na interpretação desta pesquisadora para a função que os referidos advérbios parecem desempenhar no contexto das redações estudadas.

##### 4.1. Listando os advérbios da pesquisa

Na tabela 9 abaixo, veem-se três colunas, cada uma delas contendo o resultado da extração de advérbios em *-ly* pelo componente *Wordlist* do programa *Wordsmith Tools*. Os cinco primeiros advérbios que constam da primeira coluna foram negritados nas colunas subsequentes. Desta forma, mostra-se que os advérbios *really*, *simply*, *actually*, *consequently* e *especially* são os cinco mais usados pela população brasileira estudada, os quais, por sua vez, aparecem em ordem diferente nos *corpora* de referência.

ADVÉRBIOS	Br-ICLE %	ADVÉRBIOS	LOCNESS%	ADVÉRBIO S	BAWE%
<b>really</b>	<b>0.07</b>	<i>really</i>	<b>0.04</b>	likely	0.03
<b>simply</b>	<b>0.03</b>	<i>especially</i>	<b>0.03</b>	<i>especially</i>	<b>0.02</b>
<b>actually</b>	<b>0.03</b>	<i>simply</i>	<b>0.03</b>	particularly	0.02
<b>consequently</b>	<b>0.03</b>	<i>actually</i>	<b>0.02</b>	clearly	0.02
<b>especially</b>	<b>0.03</b>	probably	0.02	<i>simply</i>	<b>0.02</b>

completely	0.03	likely	0.02	generally	0.01
certainly	0.03	finally	0.02	highly	0.01
probably	0.03	usually	0.02	usually	0.01
finally	0.03	clearly	0.02	finally	0.01
easily	0.03	eventually	0.01	<b>actually</b>	<b>0.01</b>
unfortunately	0.02	early	0.01	relatively	0.01
usually	0.02	easily	0.01	directly	0.01
daily	0.02	certainly	0.01	easily	0.01
mainly	0.02	completely	0.01	mainly	0.01
totally	0.02	merely	0.01	probably	0.01
definitely	0.02	highly	0.01	<b>consequently</b>	<b>0.01</b>
clearly	0.01	possibly	0.01	slightly	0.01
perfectly	0.01	recently	0.01	approximately	0.01
secondly	0.01	unfortunately	0.01	completely	0.01
gradually	0.01	obviously	0.01	<b>really</b>	<b>0.01</b>

Tabela 9: Lista dos 20 advérbios mais usados pelos *corpora*.

#### 4.2. Analisando os advérbios-chave

Conforme apresentado no capítulo 4 desta dissertação, palavras-chave são aquelas palavras cujas frequências são extraordinariamente altas em comparação a alguma norma. As tabelas 10, 11 e 12 a seguir dão uma amostra dos advérbios-chave dos *corpora*. Uma lista de advérbios-chave foi estabelecida com o objetivo de caracterizar o *corpus* de estudo, que é automaticamente menor que os de referência. Os *corpora* maiores fornecem dados com base para comparação.

Como se pode verificar na tabela 10, que compara o *corpus* Br-ICLE com o LOCNESS, em termos de advérbios-chave, os itens em preto são todos os advérbios-chave do Br-ICLE, em relação ao LOCNESS. Isso quer dizer que, dado o tamanho do Br-ICLE, e usando o *corpus* LOCNESS como “norma”, todos os advérbios em preto são usados em demasia pelos aprendizes. Se estes advérbios forem agrupados em campos semânticos, nota-se que há uma preponderância de advérbios que tem funções semelhantes, como *only*, *especially*, *simply*, *exactly* (que particularizam); *really*, *actually*, *certainly* (que falam sobre a veracidade de algo); *usually*,

*generally* e *daily* ( que indicam frequência); *finally* e *secondly* (que enumeram); *completely* e *totally* (que indicam grau de intensidade). Nota-se também uma voz assertiva dos aprendizes pela escolha dos advérbios como *certainly*, *clearly*, *exactly*, *extremely*, *completely*, *totally*, em contraste com *probably* e *unfortunately*, que marcam uma probabilidade e uma opinião, respectivamente. Os advérbios em vermelhos são os advérbios-chave negativos para o LOCNESS, isto é, aqueles advérbios que ocorrem muito infrequentemente de acordo com uma norma, ou seja que são pouco usados comparativamente.

<b>N</b>	<b>WORD</b>	<b>FREQ. 1</b>	<b>BRICY.LST</b> <b>%</b>	<b>FREQ. 2</b>	<b>LOCLY.LST</b> <b>%</b>	<b>KEYNESS</b>
1	only	140	0,22	0	459,3	
2	really	53	0,07	0	172,7	
3	actually	24	0,03	0	58,9	
4	consequently	22	0,03	0	86,3	
5	especially	22	0,03	0	70,6	
6	certainly	19	0,03	0	74,6	
7	usually	17	0,03	0	66,7	
8	completely	17	0,03	0	66,7	
9	simply	17	0,03	0	66,7	
10	unfortunately	16	0,03	0	62,8	
11	probably	16	0,03	0	62,8	
12	daily	15	0,03	0	58,9	
13	finally	15	0,03	0	58,9	
14	easily	15	0,03	0	58,9	
15	mainly	14	0,03	0	54,9	
16	secondly	9	0,02	0	35,3	
17	gradually	9	0,02	0	35,3	
18	totally	8	0,02	0	31,4	
19	clearly	8	0,02	0	31,4	
20	perfectly	8	0,02	0	31,4	
21	exactly	8	0,02	0	31,4	
22	generally	7	0,01	0	27,5	
23	specially	7	0,01	0	27,5	
24	extremely	7	0,01	0	27,5	

25	probably	0	82	0.03	24.8
26	actually	0	91	0.03	27.6
27	simply	0	108	0.03	32.7
28	especially	0	112	0.03	33.9
29	really	0	148	0.05	44.8
30	only	0	756	0.23	229.3

Tabela 10: Lista de advérbios-chave usados pelo Br-ICLE e pelo LOCNESS.

A outras comparações possíveis em termos de advérbios-chave são entre os *corpora* Br-ICLE e BAWE, e entre LOCNESS e BAWE, dados estes que estão nas tabelas 11 e 12 abaixo. Ao se observarem e compararem essas tabelas, pode-se inferir erroneamente de que o *corpus* Br-ICLE contém muito mais itens em comum com o *corpus* BAWE, do que o *corpus* LOCNESS. Entretanto, na tabela 12, encontram-se evidências de que o *corpus* LOCNESS tem formas errôneas (ver itens 5, 19 e 23). Além disso, percebe-se que tanto o LOCNESS quanto o Br-ICLE não usam marcadamente as palavras *previously* e *formely* (em vermelho) quando comparados com o BAWE. Essas palavras sugerem que os textos no BAWE devem estar ancorados em narrativas, uma vez que *formerly* e *previously* marcam relações temporais.

N	WORD	FREQ. 1	BRICY.LST %	FREQ. 2	BAWLY.LST %	KEYNESS
1	really	53	0,07	2.342	0,01	107,8
2	consequently	22	0,03	2.421	0,01	27,8
3	unfortunately	16	0,02	857	39	0
4	specially	7	0,01	177	26,5	0
5	previously	6	0,01	13.758	0,06	30,8
6	formerly	2	11.186	0,05	36,7	0

Tabela 11: Lista de advérbios-chave usados pelo Br-ICLE e pelo BAWE.

N	WORD	FREQ. 1	LOCLY.LST %	FREQ. 2	BAWLY.LST %	KEYNESS
1	really	148	0,05	2.342	0,01	211,6

2	only	756	0,23	37.025	0,16	96,8
3	morally	38	0,01	448	71,7	0
4	sexually	22	195	51,7	0	
5	truely	6	0	51,5	0	
6	unfortunately	43	0,01	857	47	0
7	personally	27	363	45,3	0	
8	eventually	60	0,02	1.552	44,9	0
9	probably	82	0,03	2.767	0,01	36,3
10	deadly	11	63	33,9	0	
11	financially	19	243	33,4	0	
12	greatly	52	0,02	1.490	32,4	0
13	supposedly	18	249	29,4	0	
14	seriously	26	517	28,5	0	
15	actually	91	0,03	3.552	0,02	27,5
16	simply	108	0,03	4.475	0,02	27
17	honestly	8	42	25,8	0	
18	basically	21	381	25,8	0	
19	unfortunatly	3	0	25,7	0	
20	naïvely	3	0	25,7	0	
21	lucidly	3	0	25,7	0	
22	hopefully	20	358	24,9	0,000001	
23	quantatively	4	3	24,8	0,000001	
24	previously	25	13.758	0,06	229,3	0
25	formerly	7	11.186	0,05	252,1	0

Tabela 12: Lista de advérbios-chave usados pelo LOCNESS e pelo BAWE.

Após ter analisado as listas de advérbios-chave nos *corpora*, as próximas seções terão como objetivo analisar os cinco advérbios mais frequentes em termos absolutos no *corpus* de aprendiz: *really*, *simply*, *actually*, *consequently* e *especially*.

### 4.3.Really

#### 4.3.1. Função

Segundo as gramáticas funcionais adotadas nesta dissertação *really* é considerado índice ou indicador de modalidade. Dentro desta perspectiva, o advérbio é descrito por Halliday e Matthiessen (2004:82,129-130) e na gramática funcional pedagógica de Droga e Humphrey (2002:73,98) como **adjunto de modo** com função intensificadora, quando puder ser substituído por *even*, *actually* e *in fact*. O advérbio é também descrito como um **advérbio modal** com a função de persuasão (Halliday e Matthiessen, op.cit.), quando puder ser substituído pela expressão *I assure you* e tendo como sinônimos *honestly*, *believe me*, *seriously*. Além disso, pode ser classificado ainda como um **adjunto de comentário** quando for um marcador de veracidade, podendo ser, nestes casos, substituído por *actually*. O gráfico 1 abaixo apresenta a distribuição das funções desempenhadas pelo advérbio *really* conforme observado nos *corpora*.

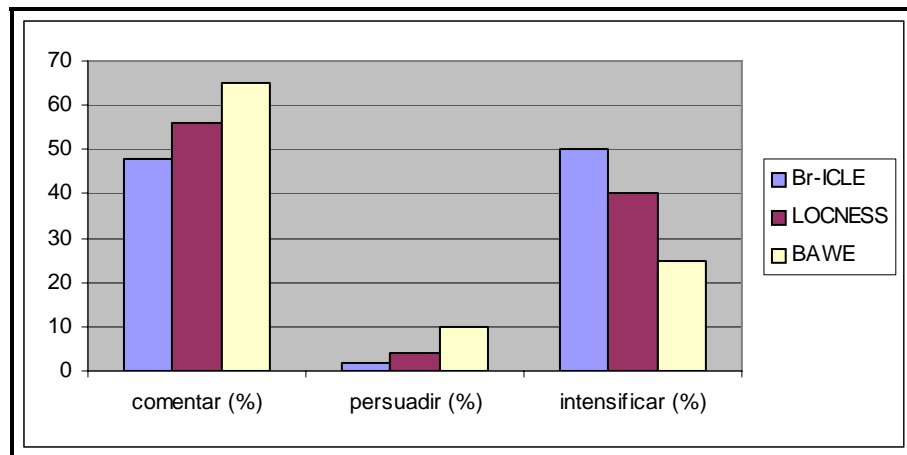


Gráfico 1: distribuição das funções de *really* nos *corpora*.

Pela observação do gráfico 1 acima, as duas funções mais recorrentes para o *really* são de comentar e intensificar, seguindo por último, da função de persuadir. Entretanto, a distribuição dessas funções nos *corpora* de estudo e de referência é reversa.

Quando *really* é usado para comentar, pode assumir dois papéis: em orações declarativas, apresenta um ponto de vista dentro de um contexto, sem necessariamente haver a intenção de persuadir o leitor, ou seja, o discurso parece estar centrado naquele que fala; já nas interrogativas, há o diálogo com o leitor, ou seja, o foco está em que lê (Halliday e Matthiessen, 2004). A seguir encontram-se exemplos de *really* extraídos dos *corpora* com função de comentário:

1) *However, most of these laws are not respected by Brazilians who generally break them or have no idea that they exist, which means that they are neither followed nor **really** put into practice by the government. (Br-ICLE: a1036.txt)*

2) *Can we **really** afford to play for time by dragging our feet? (LOCNESS:h\br67.txt)*

3) *Survey research can never **really** go into such depths as to completely change conceptions of a particular group of people, or uncover major illegal activity for example, previously unknown about. (BAWE:0140c.txt )*

Quando *really* é usado para intensificar, o autor parece assumir a responsabilidade direta sobre o que enuncia, assim como a sua veracidade (Halliday e Matthiessen 2004), sem necessariamente estar atado a alguma evidência que lhe dê sustentação. Abaixo seguem alguns exemplos de *really* com função intensificadora:

4) *As crime has become **really** hard to control, it feels like people have never been so insecure about their own safety as they have been these days. (Br-ICLE:a1080.txt).*

5) *These teachers **really** care and try to work with their students (LOCNESS:us133.txt)*

6) *The war in Vietnam **really** highlighted and confirmed King's new ideas. (BAWE:0310a.txt)*

Por outro lado, quando *really* é persuasivo é possível notar que, assim como no comentário, também existe a **marcação da voz** de quem escreve, porém o argumento é sustentado com a apresentação de fatos, conforme apontam os exemplos a seguir:

7) *Does crime pay here? I **really** think that it pays, and it is so for three main reasons: people are always prepared to buy smuggled goods; criminals never go to jail; and crime is a very profitable activity. (Br-ICLE: a1093.txt).*

8) *On economic matters, I could not **really** see what the debate was about on whether to have a central European Bank and how, as far as I am concerned the Bundesbank does a very good job of it at the moment. (LOCNESS:h\br81.txt).*

9) *An event at time t1 in the first dimension of time can be simultaneous with God's existence in the second dimension of time, as can event t2. t1 and t2 are not*



*simultaneous. This is really an explication of Stump and Kretzmann's (1981) ET-simultaneity. (BAWE:0303b.txt).*

#### 4.3.2. Perfil léxico-gramatical.

##### a) Colocados

A tabela 13 abaixo apresenta os colocados<sup>49</sup> 1L e 1R mais frequentes do advérbio *really* nos *corpora*. Os itens em azul são aqueles compartilhados pelos grupos comparados. Como se observa, foram encontrados apenas 3 colocados mais frequentes do advérbio *really* no Br-ICLE. Embora o *corpus* Br-ICLE apresente colocados em comum com os *corpora* de referência, nota-se que estes últimos apresentam uma gama maior de variedades de colocados.

	Br- N ICLE	TOTAL	L1	R1	LOCNESS	TOTAL	L1	R1	BAWE	TOTAL	L1	R1
1	is	7	5	2	not	26	22	4	not	90	85	5
2	are	7	6	1	is	14	12	2	is	83	75	8
3	people	2	2	0	to	13	2	11	to	80	14	66
4					a	12	3	9	the	69	4	65
5					it	12	10	2	it	45	32	13
6					the	9	0	9	in	35	0	35
7					are	7	6	1	only	33	28	5
8					they	7	4	3	was	33	31	2
9					never	6	6	0	a	32	10	22
10					people	6	3	3	what	27	16	11

Tabela 13: lista de colocados do advérbio *really* nos *corpora*

<sup>49</sup> Com relação ao *corpus* BAWE, são apresentados apenas um recorte dos mais frequentes. A lista completa de colocados e agrupamentos lexicais dos advérbios encontra-se no apêndice desta dissertação.

## b) Agrupamentos lexicais

O *corpus* Br-ICLE apresentou apenas um agrupamento lexical formado pelo pronome referencial *it* e pelo verbo *is* que pode estar associado a um processo relacional ou material. Este agrupamento só foi compartilhado com o *corpus* BAWÉ (item em vermelho). Os *corpora* de referência, embora contemplem textos de nativos, também se mostram distintos, seja pelas escolhas de combinações de palavras dos agrupamentos seja pelo número de ocorrências.

N	Br-ICLE	Total	LOCNESS	Total	BAWE	Total
1	it is really	3	is not really	7	is not really	23
2			did not really	4	it is really	16
3			does not really	3	does not really	13
4			he is really	3	are not really	11
5			is it really	3	only really be	11
6			no one really	3	not really a	10
7					can only really	9
8					it really is	8
9					it is not	6
10					was not really	6

Tabela 14: lista de agrupamentos lexicais do advérbio *really* nos *corpora*

## c) Padrões sintáticos

Conforme apresentado na tabela 13 acima, no *corpus* Br-ICLE, os colocados 1L e 1R do advérbio *really* formam padrões sintáticos, por exemplo, o advérbio é precedido por **verbos** *is* e *are* que podem estar associados a processos materiais ou relacionais e por um **substantivo** *people*.

10) *It is really sad to accept that people let their lives be invaded and their families mesmerized by non-sense programs that will add nothing but disorder to their lives. (Br-ICLE:a1018.txt)*

11) *People really fought against it, mainly during the most recent era, since the idea of democracy and equality was spread around the world. (Br-ICLE:a1109.txt)*

No *corpus* LOCNESS, o advérbio *really* é precedido pelo **modificador** *not*, pelo **advérbio** *never*, pelos **verbos** *is* e *are* (que podem representar processos materiais ou relacionais) e pelos **pronomes** *it*, *they* e *he*. É sucedido pelos **artigos** *a* e *the* (que sugerem um sintagma nominal), por **preposição** *in* (que sugere um sintagma preposicional) e pelo pronome relativo *what*.

12) *Hugo has too many hang ups about himself and he always acts in his own interests and not really for the good of the party. (LOCNESS:h\br15.txt)*

13) *His claim is never really clear because it is not his own. (LOCNESS:us26.txt)*

14) *In Les Mains Sales Hugo is really acting in his own interests. (LOCNESS:h\br10.txt)*

15) *Also, if the death penalty is really supposed to bring about justice to the victims of the murder and to society, it really does not accomplish that. (LOCNESS:us16.txt).*

16) *At twelve years old the mother could be near seventy and is this really the age to bring up a teenager?(LOCNESS:Is\a75.txt)*

No BAWE, o advérbio é precedido pelo **modificador** *not*, pelos **verbos** *is* e *was* (que também podem estar associados a processos materiais e relacionais), pelo **advérbio** *only* e pelos **pronomes** *it* e *what* e é sucedido pelos **artigos** *a* e *the* que sugerem um sintagma nominal e pela **preposição** *in* que sugere um sintagma preposicional.

17) *The technocratic perspective on the EU's functioning stresses that experts in certain fields across all countries gather together as epistemic communities and, since they seem to know best what is needed, they may not really need a mandate from citizens. (BAWE:0195d.txt)*

18) *Firstly, because the concept of decommodification is based on formal wage labour, women can only really be included in Esping-Anderson's analysis if they are in the labour market. (BAWE:0075I.txt)*

19) *In this section, we will examine the differences and similarities between Plutarch and his Egyptian predecessors more thoroughly. In doing so we will also assess just how damaging **it really** is to place Plutarch's version of the Osiris myth among other native Egyptian myths. (BAWE:6211f.txt)*

20) *It's **really** a waste of time. (BAWE:0166a.txt)*

21) *So, in what was nominally a democracy, power was **really in** the hands of the first citizen. (BAWE:6066b.txt)*

Além de padrões sintáticos em termos de categorias gramaticais, também foram observados padrões de acordo com a posição do advérbio nas orações dos *corpora*, conforme quantificado na tabela 15 abaixo:

<i>REALLY</i>	início (%)	meio (%)	fim (%)
Br-ICLE	2	98	0
LOCNESS	0	100	0
BAWE	0,5	99	0,5

Tabela 15: posição do advérbio *really* nos *corpora*

Pela tabela, conclui-se que a posição mais recorrente do advérbio é no meio das orações (exemplo 22). Todas as ocorrências em início de orações encontradas no Br-ICLE marcam um discurso persuasivo (exemplo 23). Entretanto, no BAWE, o advérbio foi encontrado em início ou final de orações apenas como parte de uma citação (exemplo 24) ou de uma enumeração (exemplo 25).

22) *Can we **really** afford to play for time by dragging our feet? The issue of Political Union is still more disputed. (LOCNESS:h\br67.txt)*

23) *[...] they see cartoons full of deaths and they imagine them in that situation and want to be like the hero... **Really**, it's possible to conclude that the television has the whole domination in the people in the actual world.(Br-ICLE:a1007.txt)*

24) *My dear sir, this is a very great country, a very great country. Its system of equity is a very great system, a very great system. **Really, really!**" Dickens seems to be mocking the Court of Chancery through Mr. Kenge's exclamations; (BAWE:0312h.txt)*

25) *A few examples are found with expressions of disagreements such as no, not a lot, **not really**. (BAWE:6173a.txt)*

#### d) Padrões semânticos

Como apresentado na seção anterior, o advérbio *really* pode expressar um comentário, pode intensificar um processo realizado por um verbo ou pode intensificar um adjetivo além de ser usado como um elemento persuasivo. No *corpus* de aprendiz, a função mais recorrente é do advérbio enquanto intensificador. No *corpus* de nativo, é do advérbio enquanto recurso para fazer comentário.

#### e) Prosódia semântica

Em termos de prosódia semântica, observou-se que *really* quando intensificador no *corpus* Br-ICLE, está associado com adjetivos de valor semântico negativo, enquanto que no *corpus* BAWE, ele está associado com adjetivos de valor semântico positivo.

26) *In fact, what is **really harmful** for kids is lack of attention from mother and from father due to their unsubstitutable role in growth and education. (Br-ICLE:a1059.txt).*

27) *It is **really interesting** to find out the reason behind this notice [...]*  
(BAWE:3017b.txt)

#### 4.3.3. Conclusões sobre o advérbio *really*

Levando em consideração que o *corpus* BAWE contenha os parâmetros de linguagem proficiente a ser usada como alvo pelos aprendizes e comparando a frequência de cada função no

gráfico, é possível concluir que, quanto maior o nível de proficiência na língua, menos o *really* irá desempenhar a função intensificadora e mais a função de comentário e /ou persuasão. Por exemplo, enquanto no Br-ICLE a função de comentário é exercida em 48%, no BAWE ela é exercida em cerca de 65%. Por outro lado, enquanto no Br-ICLE a função de intensidade é exercida em 50% das ocorrências, no BAWE ela é duas vezes menor, cerca de 25%. Ou seja, uma vez que os aprendizes usam mais a função intensificadora, o leitor pode ter a impressão de que o discurso deles é mais assertivo e hiperbólico.

Quanto ao LOCNESS, vê-se que seus sujeitos usam menos a função de intensidade e mais a de comentário e persuasão. No BAWE, as funções de comentário e intensidade apresentam quadros opostos àqueles apresentados pelo Br-ICLE: enquanto aprendizes intensificam mais, nativos intensificam menos; e enquanto nativos comentam mais, aprendizes comentam menos. Pela observação dos gráficos 2, 3 e 4 abaixo, a diferença no emprego das funções do *really* pelos aprendizes e pelos nativos torna-se evidente.

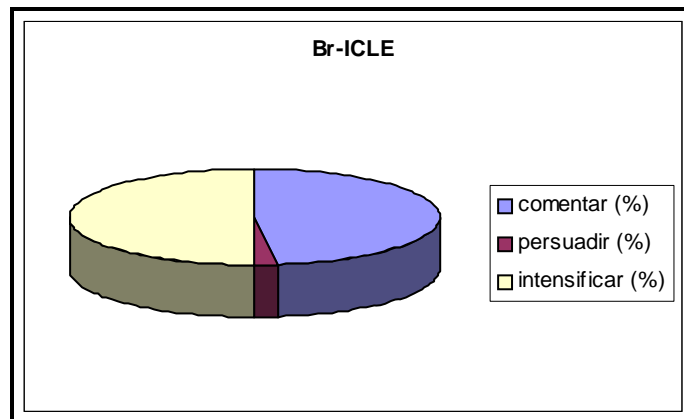


Gráfico 2: funções de *really* no Br-ICLE

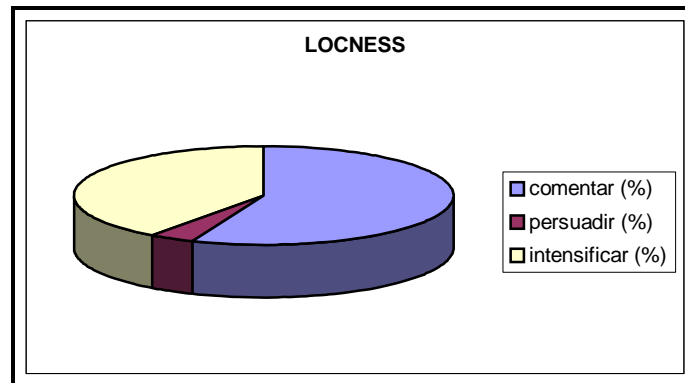


Gráfico 3: funções de *really* no LOCNESS

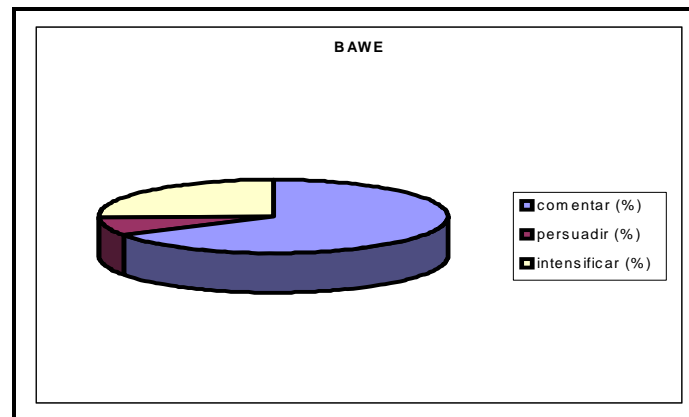


Gráfico 4: funções de *really* no BAWE

Independentemente de ser persuasivo, intensificador ou expressar um comentário, o advérbio *really* é sinalizador de função interpessoal, sendo modalizador

Além destas considerações acerca das funções do advérbio *really*, é possível notar duas diferenças principais entre a distribuição de colocados dos *corpora*, uma das quais pode ser atribuída ao fato de que o *corpus* Br-ICLE é um *corpus* pequeno. A primeira diferença é que os *corpora* de referência apresentam um maior leque de colocados tanto à direita quanto a esquerda de *really*. A segunda é que no *corpus* Br-ICLE, *really* é acompanhado de adjetivos de conotação negativa e no *corpus* BAWE, os adjetivos que acompanham o *really* têm prosódia semântica positiva.

Através da posição do advérbio nos *corpora*, nota-se que quando *really* é utilizado por aprendizes no início de orações, há a marcação de uma voz persuasiva. Nos *corpora* de nativos quando ocorre no início ou no final de orações é basicamente para fazer uma citação.

Com relação aos agrupamentos lexicais, o *corpus* Br-ICLE, apresentou apenas uma construção (*it is really*), enquanto que os *corpora* de nativos apresentaram inúmeras unidades multipalavras.

#### 4.4. *Simply*

##### 4.4.1. Função

O item *simply* é descrito por Halliday e Matthiessen (2004), além de Droga e Humphrey (2002), como adjunto de modo com função de **intensificador limitador**. Nos *corpora*, o advérbio parece limitar um evento ou uma ação, seja de uma pessoa, grupo de pessoas ou instituição, conforme ilustram os exemplos a seguir retirados dos *corpora*:

28) *Actually, some argue that it would be hypocrisy to try to censor programs, because they are **simply** a reflection of the society we live in. (Br-ICLE:a1076.txt)*

29) *However, it is very often to see this new generation of school, or even university students buying done works at internet, or **simply** copying-pasting it and then signing as if it was done by them. (Br-ICLE:a1097.txt).*

30) *Therefore, it became important to produce **not simply** the necessary for subsistence and, consequently, people have to work harder in order to produce more than necessary with the purpose to generate revenue. (Br-ICLE:a1051.txt)*

31) *Again, females are not inferior nor are males they are **simply** different. (LOCNESS:us132.txt)*

32) *It is **simply** impossible for the average person to get inside the mind of a criminal and see what he or she fears most. (LOCNESS:a\us3.txt)*

33) *Opponents to this issue might pose the question: if a parent did not want their children to view a particular show it would seem logical for the parent to **simply** turn the television off; [...](LOCNESS:us39.txt)*



34) *Nor a Socialist. Nor anything. **I am simply a writer.*** (BAWE:0419b.txt)

35) *Her characters are not out to reverse gender roles or stereotypes, they are **simply** looking for what creates them, attempting to unravel them.* (BAWE:0224a.txt)

36) *Thus it was **simply** part of life that they were literally seen as hot-headed and impetuous.* (BAWE:6066e.txt)

Apesar da frequência de *simply* nos *corpora* ser semelhante, isto é, 0,03% no Br-ICLE, 0,03% no LOCNESS e 0,02% no BAWE, e de praticamente não haver diferença de função do advérbio nos *corpora*, a diferença se apresenta nos padrões apresentados e na posição do advérbio nas orações, conforme será apresentado na seção 5.4.2 a seguir.

#### 4.4.2. Perfil léxico-gramatical

##### a) Colocados

Como se nota através da tabela 16 abaixo, o *corpus* Br-ICLE não apresentou nenhum padrão de colocado para *simply* seja imediatamente à direita ou à esquerda do advérbio. Isto quer dizer que os aprendizes usam *simply* com frequência semelhante aos seus pares, mas o fazem de modo individualizado.

Os itens em azul são aqueles compartilhados pelo *corpus* LOCNESS e BAWE. Nota-se ainda que apesar de serem dois *corpora* de nativos, o *corpus* LOCNESS apresentou colocados em menor número comparado ao BAWE.

N	LOCNESS	TOTAL	L1	R1	BAWE	TOTAL	L1	R1
1	to	17	6	11	is	194	185	9
2	is	13	13	0	the	179	9	170
3	the	13	0	13	to	134	49	85
4	they	8	5	3	not	130	89	41
5	a	5	1	4	a	69	17	52
6	are	5	5	0	it	58	28	30

7	for	5	0	5	or	55	50	5
8	it	5	4	1	was	51	51	0
9					as	48	27	21
10					that	45	8	37

Tabela 16: lista de colocados do advérbio *simply* nos *corpora* LOCNESS e BAWE

### b) Agrupamentos lexicais

Foram encontrados agrupamentos lexicais apenas nos *corpora* de referência. O único agrupamento compartilhado (dentro do limite de corte) pelos *corpora* é o *it is simply*, formado pelo pronome de referência *it* e pelo verbo *is* que pode estar associado a um processo material ou relacional. Os demais agrupamentos são formados por conjunção, advérbio, verbos lexical e auxiliar e modificador. Estes agrupamentos apontam para a formação de definições e particularizações.

N	LOCNESS	Total	BAWE	Total
1	<i>it is simply</i>	5	is simply a	40
2	simply because they	3	<i>it is simply</i>	29
3	simply wanted to	3	is not simply	25
4			rather than simply	20
5			is simply the	18
6			not simply a	16
7			is simply not	15
8			this is simply	15
9			simply do not	13
10			are not simply	12

Tabela 17: lista de agrupamentos lexicais nos *corpora* LOCNESS e BAWE.

### c) Padrões sintáticos

Com relação aos padrões sintáticos, foram notadas ocorrências apenas nos *corpora* de referência. No LOCNESS, os padrões 1L são com **verbo** *is* e *are* que podem representar processos materiais ou relacionais e com **pronome** *they*; os padrões 1R são com **preposição** *for* e *to* (que também pode ser partícula de infinitivo) apontando para um sintagma preposicional; e **artigo** *the* e *a*, que podem indicar um sintagma nominal.

37) *His existence is simply impossible. (LOCNESS:sh\br9.txt)*

38) *They have no ability to think, or be creative, or be spontaneous. They simply perform millions of calculations, very quickly, following instructions laid out by a human. (LOCNESS:Is\a53.txt)*

39) *Also, in Weisel's Night the Jews are exclusively evacuated from their land and dehumanized by the Germans simply for being Jews. (LOCNESS:us179.txt)*

40) *Monarchy is simply a tradition and keeping an institution merely for the sake of tradition is not right and against all reason. (LOCNESS:Is\a70.txt)*

No *corpus* BAWE, o advérbio é precedido (padrão 1L) por **verbo** *is* e *was*, indicando um processo material ou relacional; por um **modificador** *not*; e por **conjunção** *or* e *as* (que podem estabelecer uma relação de alternância ou de comparação respectivamente). Os elementos que sucedem o advérbio (padrão 1R) são: **preposição** ou **partícula de infinitivo** *to*, **artigo** *the* e *a*, que sugerem um sintagma nominal e **pronome** *it* e *that*.

41) *Whilst other architects were turning their buildings inside out in celebration of new technology Venturi was simply applying it in improvements to the way people live on both a personal and public scale. (BAWE:3096a.txt)*

42) *and failed to take justice and not simply legal considerations into account (BAWE:0360c.txt)*

43) *In his essay entitled, "Honor, Blood, and Poetry in Yerma", Gustavo Correa outlines the Spanish woman's physical boundaries as 'the walls of her house' and her occupation as simply 'matrimony'.(BAWE:6998e.txt)*

44) *Since it is not possible to observe the effects of WIMPs on surrounding objects, the only way to detect a WIMP is simply to witness one. (BAWE:6097c.txt)*

45) *After the politician has gained the approval of the masses to lead, Weber believed it is in turn the responsibility of the masses to submit themselves to the leader, "the following of such a leader must obey him blindly... this is **simply** the price paid for guidance by leaders" (FMW 113) (BAWE:0350a.txt).*

46) *This does not mean the disease was not real, **simply that** it was psychological and the label was placed on it by society*

Além das diferenças de padrões sintáticos, também nota-se que o advérbio *simply* se difere nos *corpora* quanto à sua posição nas orações, conforme ilustra tabela a seguir:

<i>SIMPLY</i>	início (%)	meio (%)	fim (%)
Br-ICLE	0	96	14
LOCNESS	5	95	0
BAWE	2	97,5	0,5

Tabela 18: posição do advérbio *simply* em orações dos *corpora*.

Apesar do advérbio *simply* ser mais frequentemente encontrado no meio das orações, os aprendizes do Br-ICLE não usam o advérbio para iniciar nenhuma oração, fato que não é notado no LOCNESS e no BAWE. Por outro lado, encontra-se o advérbio em posição final em 14% das ocorrências no Br-ICLE, frequência bem maior que a do BAWE que é de apenas 0,5%, caracterizando um sobreuso desta posição.

47) *We should have time to seize the opportunity of living, **simply**. (Br-ICLE: a1071.txt)*

48) *The fact is that society **simply** craves violence and capital punishment [...] (LOCNESS:a\us3.txt)*

49) *For the store, it is in terms of customers getting in the store, selecting products and buying them. **Simply** put the more customers Cost cutter can get through the store per hour the more revenue they will generate.(BAWE:0169g.txt)*

É provável que a diferença de posição do advérbio nas orações esteja relacionada com o fato de aprendizes utilizarem o advérbio de forma “errônea”, conforme mostra exemplo 47 acima, onde é possível que o aluno quisesse dizer “*living simply*”.

#### d) Padrões semânticos

No *corpus* Br-ICLE, o advérbio restringe processos representados por verbos. No LOCNESS, restringe tanto processos quanto uma relação de causa-efeito entre duas orações e no BAWE, além de restringir os mesmos elementos que os outros *corpora*, restringe também sintagmas, sejam nominais ou preposicionais.

50) *People simply can not bare leave television. (Br-ICLE:1105.txt)*

51) *We should not avoid literature with violence simply because it is violent. (LOCNESS:us181.txt)*

52) *The next stage is simply a repetition of what has just been done. (BAWE:6200b.txt)*

#### 4.4.3. Conclusões sobre o advérbio *simply*

Não há nenhuma ocorrência de padrão de colocados e agrupamentos lexicais no *corpus* Br-ICLE. Os padrões que ocorrem no *corpus* BAWE se mostram distintos, com maior frequência e variedade nele do que no *corpus* LOCNESS. Isto pode ser justificável pelo fato de que, mesmo se tratando de *corpora* de nativos, o LOCNESS é composto também de textos de pré-universitários, ou seja, aprendizes de uma variedade de inglês como língua materna, ou seja o inglês argumentativo de redações. A mesma variação é percebida nos *corpora* de nativos em termos de agrupamentos lexicais.

## 4.5. *Actually*

### 4.5.1. Função

*Actually*, de acordo com Hallyday e Matthiessen (2004), pode desempenhar três funções distintas, como por exemplo, a de **adjunto conjuntivo** com a função de verificador, a **de adjunto de comentário**, quando puder ser substituído por *really*, *in fact* e *as a matter of fact*; e de **adjunto de modo** com função intensificadora. Todas essas funções serão exemplificadas mais adiante nesta seção.

Quando empregado como adjunto de comentário ou intensidade desempenha a mesma função de *really* embora exista uma diferença na preferência das funções dos advérbios usados para comentar e usados para intensificar, como será explanada na seção 5.6.

Apos extraídas as ocorrências de *actually* em todos os três *corpora*, elas foram classificadas de acordo com as funções desempenhadas e um gráfico foi elaborado para que espelhasse as várias distribuições destas funções. No gráfico 5 abaixo, observa-se no *corpus* de aprendiz um considerável sobreuso de *actually* como conjuntivo (58% das ocorrências) e uma pequena frequência das demais funções (17% para adjunto de comentário e 25% para adjunto de intensidade). O mesmo gráfico mostra essas funções em comparação aos *corpora* de nativos: os sujeitos dos *corpora* BAWE e LOCNESS utilizam o advérbio como conjuntivo com pouca frequência (6% e 5% respectivamente) dando preferência, portanto, para o advérbio enquanto recurso intensificador (cerca de 35% no LOCNESS e 47% no BAWE) e de comentário (60% no LOCNESS e 47% no BAWE).

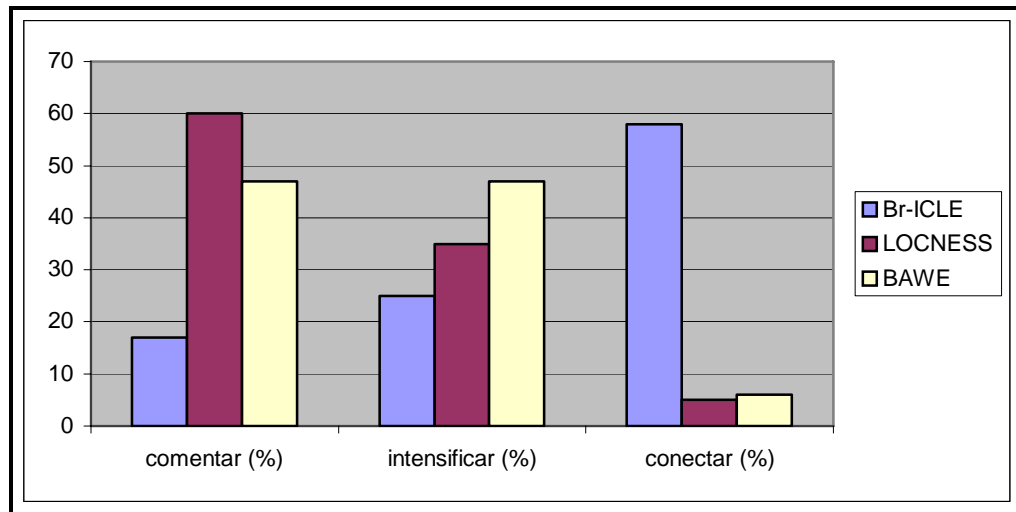


Gráfico 5:distribuição das funções de *actually* nos corpora

Assim como acontece com o advérbio *really*, *actually*, quando adjunto de comentário, marca um ponto de vista dentro de um contexto, conforme apontam exemplos 53, 54 e 55 a seguir.

53) *It is also certain that the capitalist system has found ways to exploit this social phenomenon too, men is allowed to have several women by our society, nowadays, they “may” have plenty of relationships without **actually** assuming all the responsibility and expenses of a family. (Br-ICLE:a1061.txt)*

54) *Many people believe that it **actually** places a lower value on human life. (LOCNESS:us11.txt)*

55) *Reading about the factory was easier than **actually** experiencing it. (BAWE:3012c.txt)*

*Actually*, enquanto recurso intensificador, pode ser substituído por *really*, marcando a veracidade da proposição, como se observa nos exemplos 56, 57 e 58 abaixo:

56) *Therefore, it is possible to see that the feminists **actually** brought more duties to women. (Br-ICLE:a1100.txt)*

57) *In my eyes and probably the eyes of others this is reason enough for the banning of boxing. The sport to say the least is **actually** a blood sport, the simple reason being that the idea of the sport is to 'knock your opposition out' [...]. (LOCNESS:Is\25.txt)*

58) *While claiming to be value-free on the grounds of its quasi-scientific approach to international relations, neo-realism is **actually** value-laden. (BAWE:0075d.txt)*

Como conjuntivo cumpre papel de tornar o discurso mais preciso, seja reafirmando o que foi dito na oração anterior ou abrindo espaço para um novo argumento.

59) *To sum up, as we could see from the data presented, crime in Brazil more than pays. **Actually** it is an invitation for those who want to get things easily without working hard. (Br-ICLE:a1093.txt)*

60) *Even though the pharmaceutical industry argues that medical pricing boards would raise prices and eliminate competition between companies, **actually** the opposite seems to be true. (LOCNESS:us10.txt)*

61) *If this resource is free then the server can use it. Otherwise, the server is blocked and waits for the release of this resource. **Actually**, the server waits for a synchronisation. (BAWE:3052e.txt).*

#### 4.5.2. Perfil léxico-gramatical.

##### a) Colocados

Apenas os *corpora* de referência apresentaram padrões de colocados imediatamente à direita e à esquerda do advérbio *really*. Contudo, apesar do *corpus* LOCNESS compartilhar das mesmas escolhas de padrões existentes no BAWE, este último ainda apresentou maior variedade de padrões (ver lista completa no apêndice).

N	LOCNESS	TOTAL	L1	R1	BAWE	TOTAL	L1	R1
1	the	14	0	14	is	160	151	9
2	is	13	13	0	the	123	3	120
3	a	7	0	7	was	70	68	2
4	not	7	6	1	are	68	66	2



5	he	6	6	0	in	68	3	65
6	in	5	1	4	not	62	55	7
7	it	5	3	2	and	60	25	35
8					it	48	26	22
9					a	40	2	38
10								

Tabela 19: lista de colocados do advérbio *actually* nos *corpora* LOCNESS e BAWE

### b) Agrupamentos lexicais

O *corpus* BAWE foi o único que apresentou padrões de agrupamentos lexicais. Esses agrupamentos são basicamente formados por verbos lexical, modal e auxiliar. Eles podem sugerir a marcação de um comentário ou podem ser usados como intensificador.

N	BAWE	Total
1	it is actually	17
2	may actually be	16
3	is actually a	15
4	is actually the	11
5	can actually be	10
6	is not actually	10
7	in terms of	8
8	what is actually	8
9	what they actually	8
10	do not actually	7

Tabela 20: lista de agrupamentos lexicais no *corpus* BAWE.

### c) Padrões sintáticos

Os padrões sintáticos observados a partir dos colocados do advérbio *actually* foram observados apenas nos *corpora* de referência.

O *corpus* LOCNESS apresentou padrões 1L em termos de **verbo** (*is*), **pronome** (*he/it*) e **modificador** (*not*). O padrão 1R foi com **artigo** (*the/a*), que pode introduzir um sintagma nominal.

62) *can desensitize one's sense of what is **actually** being portrayed* (LOCNESS:us106.txt)

63) *In a way he is driving for people to kill him, an in the end **he actually** wishes to be killed.*(LOCNESS:sh\br3.txt)

64) *Both are indeed seen as **not actually** representing the working class.* (LOCNESS:h\br10.txt)

65) *Even though the pharmaceutical industry argues that medical pricing boards would raise prices and eliminate competition between companies, **actually the** opposite seems to be true.* (LOCNESS:us10.txt)

No *corpus* BAWE, os padrões 1L observados foram com **verbo** (*is/are/was*), com **modificador** (*not*) e com **pronome** (*it*). E o padrão 1R foi com **artigo** (*the/a*) que pode introduzir um sintagma nominal.

66) *The reason of an emphasis on the identical surface structure of a sentence with scope ambiguity as mentioned above is that their deep structures **are actually** different although they have the same surface structure.* (BAWE:6058e.txt)

67) *The dynamic sublime involves an aesthetic estimation of a great natural force or power which has to be resistible and **not actually** threatening.* (BAWE:0407b.txt)

68) *We call an object red not because **it actually** contains the property of redness,* (BAWE:0294g.txt)

69) *This was **actually the** case initially as the English maintained good relations with Powhatan.* (BAWE:0029I.txt)

A posição do advérbio nas orações também aponta para algumas particularidades dos *corpora*, como mostrado na tabela 6:

<i>ACTUALLY</i>	início (%)	meio (%)
Br-ICLE	42	58
LOCNESS	3	97
BAWE	4	96

Tabela 21: posição do advérbio *actually* nos *corpora*

Os dados acima mostram que o advérbio é usado preferencialmente no meio de uma oração. Entretanto, quando o advérbio é posicionado no início de uma oração, pode assumir o papel de adjunto de comentário. No Br-ICLE, isto acontece em 100% das ocorrências. Nos *corpora* LOCNESS e BAWE isso acontece respectivamente em 67% e 65% das ocorrências do advérbio em início de oração.

70) *They are programs responsible for changing their behavior and stimulating their mind harmfully. **Actually**, the one who manipulates television is the principal answerable for the damage it causes in the viewer's mind*(Br-ICLE:a1076.txt)

71) *The battle flag was first flown in the year 1964. This was the year that the civil rights movement started. **Actually**, the flag was set up the day after the civil rights movement started* (LOCNESS:us127.txt)

72) *It encourages the learning process and allows the reader to experience the emotions involved without **actually** experiencing the action.* (LOCNESS:us185.txt)

73) *Specifically, one area of difficulty based on my expectations was the acquisition of the Persian pronunciation system. **Actually**, I made many mistakes by transferring L1 sounds to L2 system. For example, I kept on mispronouncing the[...]*(BAWE:6028a.txt)

74) *They have better planning and are easier to carry out the schemes which cope with the P-A problem. **Actually** solving the P-A problem is one of the means to reduce cost too.*(BAWE:6008h.txt)

#### **d) Padrões semânticos**

No *corpus* Br-ICLE, o advérbio *actually* é usado com maior frequência como adjunto conjuntivo;

Já no LOCNESS, o advérbio é usado como adjunto de comentário;

E no BAWE, o advérbio funciona tanto como adjunto de comentário quanto como adjunto de modo (intensidade).

#### 4.5.3. Conclusões sobre o advérbio *actually*

Como foi observado no início da seção 5.5.1, *actually* desempenha algumas funções nos *corpora*, tendo todavia, cada grupo (aprendizes e nativos) suas preferências.

No *corpus* Br-ICLE, os aprendizes parecem aplicar demasiadamente a função de conjuntivo do advérbio, dando menos ênfase para as funções de adjunto de modo (intensificador) e de adjunto de comentário.

No LOCNESS, os sujeitos aplicam o advérbio de maneira oposta aos aprendizes. *Actually* assume preferencialmente o papel de adjunto de comentário, seguido por sua função de adjunto de modo e por último como conjuntivo.

No BAWE, os sujeitos parecem utilizar *actually* com a mesma frequência em ambas funções seja de adjunto de comentário e ou de adjunto de modo, dando ênfase mínima para a sua função como conjuntivo.

A observação dos colocados de *actually* mostrou que os aprendizes fazem escolhas de colocados tão diferentes que não é possível notar nenhum padrão. Já os padrões de colocados apresentados pelo LOCNESS e pelo BAWE mostraram-se distintos., embora com algumas poucas semelhanças.

É possível que as diferenças estejam relacionadas com dois fatores: um, com o tamanho dos *corpora*, sendo o BAWE bem maior que o LOCNESS, pode ser que haja a maior variedade

de padrões; o outro pode estar relacionado com o nível educacional dos alunos, com o LOCNESS contendo textos de pré-universitários e o BAWE com pós-graduados.

A posição do advérbio nas orações mostrou que apesar do advérbio ser normalmente posicionado no meio de orações, as ocorrências em início de orações caracterizam um comentário.

A presença de agrupamentos lexicais apenas no BAWE aponta para questões provavelmente relacionadas à não internalização destes elementos linguísticos por parte dos aprendizes e à um processo progressivo de aprendizado por parte dos sujeitos do LOCNESS.

#### 4.6. Really X Actually

Embora os dois advérbios *really* e *actually* apresentem a mesma função quando se trata de comentar e intensificar (Halliday e Matthiessen ,2004:82,129-130 e Droga e Humphrey, 2002:73,98), parece haver algumas diferenças nos *corpora* quanto à preferência do advérbio que é usado para comentar e do que é usado para intensificar. Os gráficos 6, 7 e 8 foram obtidos a partir da comparação das funções exercidas pelos advérbios *really* e *actually* nos *corpora* conforme apresentados pelos gráficos 1 e 5 deste capítulo. As funções de intensificar e comentar foram isoladas e estão apresentadas a seguir:

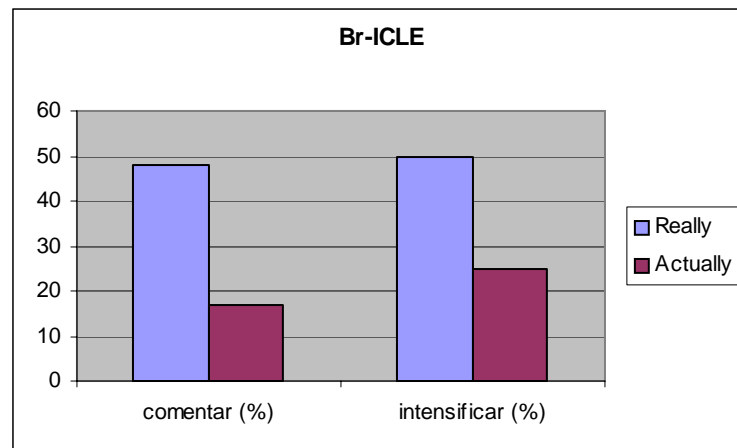


Gráfico 6: funções de *really* e *actually* no Br-ICLE

Como se observa no gráfico 6, ao se compararem as funções de cada advérbio, nota-se que *really* parece ser preferencialmente usado pelos sujeitos do Br-ICLE tanto para comentar quanto para intensificar.

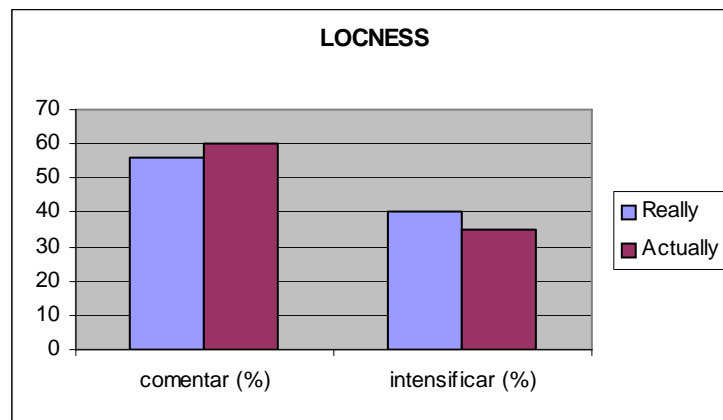


Gráfico 7: funções de *really* e *actually* no LOCNESS

No *corpus* LOCNESS, conforme aponta o gráfico 7, a diferença entre as funções dos advérbios é mínima. Entretanto, *actually* é usado para comentar em 60% das ocorrências e *really* é usado para intensificar em 40% das ocorrências.

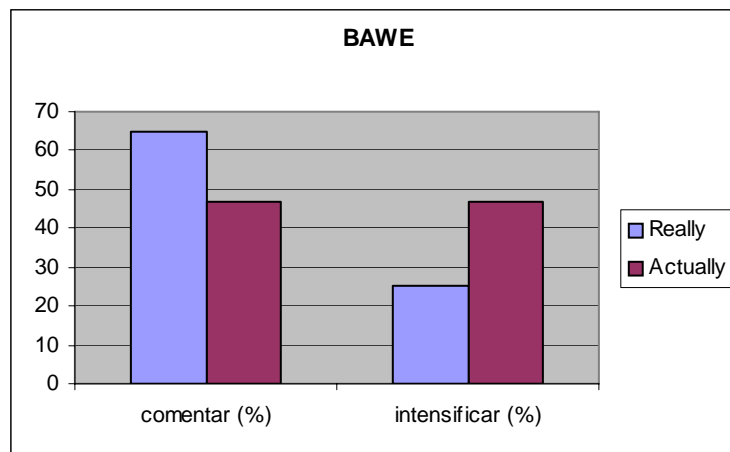


Gráfico 8: funções de *really* e *actually* no BAWE

Já no *corpus* BAWE, ocorre o oposto ao LOCNESS. Se o ponto de referência ou alvo é o *corpus* BAWE, pode-se dizer que preferencialmente usa-se *really* para comentar (65% das ocorrências) Por outro lado, *actually* é usado para intensificar em 47% das ocorrências.

#### 4.6.1. Conclusões sobre *really* X *actually*

*Really* é usado por aprendizes tanto para comentar quanto para intensificar. No *corpus* LOCNESS, *actually* é usado pelos nativos para comentar e *really* para intensificar. Por fim no *corpus* BAWE, *really* é usado para comentar e *actually* para intensificar.

Portanto, a análise de padrões de *really* e *actually* se mostrou importante para a comparação dos significados e funções desses itens aparentemente similares, porém usados com diferentes funções nos *corpora*.

### 4.7. *Consequently*.

#### 4.7.1. Função

O advérbio *consequently* é classificado como um **adjunto conjuntivo**, que pode ser substituído por *because of that*, *so*, *then*, *for*, e desta forma pode ser utilizado para unir idéias ou orações dentro de um contexto (Halliday e Matthiessen 2004), conforme ilustram os exemplos 75, 76, 77 a seguir.

Este advérbio, que aparece no Br-ICLE na posição 4 em cerca de 0,03% das ocorrências, só aparece no LOCNESS na posição 47 em 0,006% e no BAWE na posição 16 em 0,01 %. A interpretação desses dados nos diz que *consequently* é usado cinco vezes mais no Bricle em comparação com o LOCNESS e três vezes mais do que no *corpus* BAWE.

Como a função deste advérbio não se altera nos *corpora*, a diferença de uso, além é claro do fato de que *consequently* é usado em excesso, fica circunscrita aos padrões dos *corpora*, conforme segue seção 5.7.2.

75) *Some years ago humans started to worry about environmental problems as soon as it was headlined on television and they saw they were being affected they could suffer from cancer, drinking water could finish and **consequently** they would die. (Br-ICLE:a1031.txt)*

76) *The objects of one's consciousness are what the consciousness focuses on. They are therefore fundamentally different. There are **consequently** two modes of existence as Sartre sees it: existence without consciousness and therefore without meaning, and that with these. (LOCNESS:h\br12.txt)*

77) *Other stakeholders included the Public Affairs within the Premier Automotive Group and the wider Ford Motor Company, both of whom had some interest in the project due to it being an organisational first, but neither had the means to influence the project and conscious of increasing complexity I wanted to avoid their involvement. **Consequently** I provided a written update to the Directors of both respective teams upon project commencement and completion. (BAWE:0172c.txt)*

#### 4.7.2. Perfil léxico-gramatical.

##### a) Colocados

O único padrão observado no *corpus* Br-ICLE e no *corpus* LOCNESS, e que também foi encontrado no BAWE, foi com a conjunção *and*. Mais uma vez o *corpus* BAWE apresentou um número maior de padrões de colocados.

	Br- ICLE	TOTAL	L1	R1	LOCNESS	TOTAL	L1	R1	BAWE	TOTAL	L1	R1
1	and	10	9	1	and	7	6	1	and	210	206	4
2									is	62	17	45
3									the	51	0	51



4								to	26	1	25
5								a	20	0	20
6								was	17	10	7
7								it	15	5	10
8								would	15	7	8
9								are	14	2	12
10								can	14	2	12

Tabela 22: lista de colocados do advérbio *consequently* nos *corpora*

### b) Agrupamentos lexicais

Apenas no *corpus* BAWE foram encontrados agrupamentos lexicais. Estes são formados basicamente por conjunção, verbo lexical e artigo. Parecem estabelecer uma união com a oração que antecede dentro de um contexto de causa e consequência .

N	BAWE	Total
1	and consequently the	27
2	consequently it is	7
3	and consequently a	6
4	and consequently it	5
5	and is consequently	5
6	there is a	5
7	and consequently to	4
8	consequently lead to	4
9	it is not	4
10	the world and	4

Tabela 23: lista de agrupamentos lexicais no *corpus* BAWE.

### c) Padrões sintáticos

Br-ICLE e LOCNESS formaram apenas um padrão com **conjunção** (*and*) e na posição 1L, antecedendo o advérbio *consequently*. Isto significa que há preferência dos sujeitos do Br-ICLE e do LOCNESS em estabelecer uma ligação entre duas orações e também caracterizar essa ligação como causa e consequência.

78) *From an American perspective **and consequently** a Brazilian one, it is usual to believe that every person should detain a university degree. (Br-ICLE:a1106.txt)*

79) *A fully integrated market with Britain as a full member would be to the detriment of our political system **and consequently** leading to a loss of sovereignty. (LOCNESS:h\br59.txt)*

No *corpus* BAWE, o padrão 1L observado foi com **conjunção** (*and*) e com **verbo** (*was*). O padrão 1R foi formado por **verbo** (*is/are/can/would*), por **artigo** (*the/a*), por **pronome** (*it*) e por preposição (*to*).

80) *All in all, these critics show that Dickens provided his readers with a negative image of London **and consequently** of society in general [...]. (BAWE:3012b.txt)*

81) ***Consequently the** literature of identity was born which glorified countries, acclaimed resources and appraised people (BAWE:0261a.txt)*

82) ***Consequently it** is hard to have an ultrasound beam that can resolve fine detail deep into an object (BAWE:0311f.txt)*

83) *While on the one hand the interaction of the "transnational nucleus" (semi-periphery) with pre-existing societies (periphery) renders the latter aspiring for that lifestyle, the growing gap between the two leads to dissatisfaction between classes and **consequently to** social unrest (BAWE:0413c.txt)*

Quanto a posição do advérbio nas orações, apesar de todos os *corpora* apresentarem maior frequência do advérbio em início de orações (Br-ICLE 50%, LOCNESS 43% 3 BAWE 52%), a tabela mostra que a segunda maior frequência do advérbio em *corpora* de nativos é no meio de orações (LOCNESS 47% e BAWE 45%) e no *corpus* de aprendiz, entre vírgulas (33%) no meio das orações. No *corpus* LOCNESS há ainda uma pequena frequência do advérbio no final de orações (5%).

<i>CONSEQUENTLY</i>	início (%)	meio (%)	fim (%)	entre vírgulas (%)
Br-ICLE	50	17	0	33
LOCNESS	43	47	5	5
BAWE	52	45	0,5	21

Tabela 24: posição do *consequently* em orações dos *corpora*.

Os exemplos 84 a 94 a seguir referem-se à posição do advérbio nos *corpora*:

84) *Our society has changed over the last centuries and thanks to the science technology development and industrialisation we have a better standard of living. **Consequently** people have become more materialistic and have forgotten the basic principles to live happily. (Br-ICLE:a1017.txt)*

85) *From an American perspective **and consequently** a Brazilian one, it is usual to believe that every person should detain a university degree. (Br-ICLE:a1106.txt)*

86) *Therefore, it became important to produce not simply the necessary for subsistence and, **consequently**, people have to work harder in order to produce more than necessary with the purpose to generate revenue. (Br-ICLE:a1051.txt)*

87) *This money is mostly then invested into important institutions by the Heritage Fund. **Consequently** it has enabled the survival of many theatres and other places of historic and cultural interest. (LOCNESS:Is\48.txt)*

88) *Hugo came from a bourgeoisie background, **consequently**, having found a political cause, he was anxious to compensate for his background. (LOCNESS:sh\br8.txt)*

89) *If this did not happen then the UK would be in breach of the Treaty of Rome and would **consequently** be in breach of international law which would attract international responsibility. (LOCNESS:h\br72.txt)*

90) *She is very angry subconsciously and is very violent consequently; (LOCNESS:us180.txt)*

91) *The negative connotation of constant talking is still labelled to women, however, regardless of the countless amounts of data evidence which prove otherwise. **Consequently** and possibly rightly so, Spender holds the belief that*

*women are not respected in society in that their language is 'devalued to such an extent that they are required to be silent (BAWE:6120e.txt)*

92) *'He,' is denied an identity and **consequently** an existence, which suggests that masculine and religious hierarchy have become obsolete in the contemporary world. (BAWE:3008d.txt)*

93) *Ransley and Ingram (2001) believe that a valuable design can catch the attention of the targeted market segment and makes it possible for the hotel to set the prices **consequently**. (BAWE:3081b.txt)*

94) *In terms of class I will look at how working-class women received different treatments, as a result of the rise of medicine as a profession, and also how, **consequently**, the divisions between working-class and upper-class women increased. (BAWE:0001c.txt)*

#### **d) Padrões semânticos**

Os padrões semânticos observados nos *corpora* apontam para a adição de idéias numa relação de causa e consequência com a oração anterior.

95) *Parents work all day long questioning for money and, **consequently**, for a tranquil life, children go to school (Br-ICLE:a1071.txt)*

96) *Another reason in which society does not need a death penalty is because a death penalty really is not "full proof" and **consequently** is not the correct choice for a punishment. (LOCNESS:us16.txt)*

97) *Some chemical pesticides also enter food-chains and **consequently** threaten human health and a wide range of organisms. (BAWE:6113a.txt)*

#### 4.7.3. Conclusões sobre o advérbio *consequently*

A análise dos colocados do advérbio *consequently* permitiu identificar apenas um padrão no Br-ICLE (*and consequently*) que também foi encontrado no LOCNESS e no BAWE.

A observação da posição do advérbio nos *corpora* mostrou que *consequently* não aparece em final de oração no *corpus* de aprendiz, embora haja ocorrências em *corpora* de nativos. Entretanto, nota-se que o emprego ao final da oração é idiossincrático do *subcorpus* de inglês americano.

#### 4.8. *Epecially*.

##### 4.8.1. Função

*Epecially* é um **adjunto conjuntivo** com função de particularizador, ou seja, especifica um determinado elemento de uma proposição (Halliday e Matthiessen, 2004 e Droga 2002), assim como explicitado nos exemplos extraídos de nossos *corpora*, abaixo:

98) *In this situations, any idea or wrong concepts may be accepted as true. That is what is happening in most places nowadays, **especially** among young people and in poor countries.* (Br-ICLE:a1120.txt)

99) *The idiocy of his argument and the dry dogma he tries to enforce on others, **especially** Candide, serve to ridicule optimism and denounce it as having any philosophical worth.* (LOCNESS:h\br52.txt)

100) *Women may forgo certain activities, or have to modify/adapt their behaviour to reduce the risk of crime, **especially** after dark and when using public transport.* (BAWE:3134e.txt)

Nos exemplos acima, *especially* pode particularizar um grupo de pessoas (*especially among young people*, exemplo 98), uma pessoa (*especially Candide*, exemplo 99) e um circunstância (*especially after dark*, exemplo 100). Como a função deste advérbio não se altera nos *corpora*, isto é, os aprendizes e “nativos” tendem a usar *especially* com a função particularizadora, a diferença, mais uma vez, parece encontrar-se nos padrões de co-colocados dos *corpora*, conforme será discutido na seção 5.8.2 abaixo

Entretanto, através da observação dos *corpora*, notou-se que *especially* também pode funcionar como um **intensificador** quando anteceder um adjetivo (exemplos 101 e 102).

101) These changes are *especially important* to children who spend several hours a day in front of TV, watching from cartoons to inappropriate programs for their ages. (Br-ICLE:a1042.txt)

102) This can be *especially interesting* for hospitality and tourism organisations, which are known to have customers from all over the world, including many minorities. (BAWE:3022c.txt)

#### 4.8.2. Perfil léxico-gramatical.

##### a) Colocados

O *corpus* Br-ICLE não apresentou nenhum padrão de colocado seja à direita ou à esquerda do advérbio *especially*.

O *corpus* LOCNESS compartilhou todos os seus colocados com o BAWE, contudo, este último, apresentou maior variedade de colocados.

N	LOCNESS	TOTAL	L1	R1	BAWE	TOTAL	L1	R1
1	the	22	0	22	the	293	1	292
2	is	14	12	2	and	133	94	39
3	in	8	0	8	is	121	118	3
4	and	5	4	1	in	81	10	71
5					to	54	1	53
6					for	43	1	42
7					it	37	5	32
8					are	29	23	6
9					of	28	1	27
10					that	28	9	19

Tabela 25: lista de colocados do advérbio *especially* nos *corpora* LOCNESS e BAWE

##### b) Agrupamentos lexicais

Os agrupamentos lexicais encontrados nos *corpora* de referência LOCNESS e BAWE são formados por pronomes, verbos lexicais, adjetivos, artigos e preposição. Embora tenham ocorrido semelhanças entre os padrões de agrupamentos lexicais, o BAWE apresentou maior variedade de combinação de palavras. Não foi encontrado nenhum agrupamento no *corpus* Br-ICLE.

N	LOCNESS	Total (%)	BAWE	Total (%)
1	especially in the	8	especially in the	88
2	this is especially	7	this is especially	62
3	is especially true	4	is especially true	28
4	especially when there	3	is especially important	25
5	especially with the	3	especially when the	23
6			especially as the	16
7			and especially the	15
8			especially for the	15
9			especially if they	14
10			especially with the	14

Tabela 26: lista de agrupamentos lexicais nos *corpora* LOCNESS e no BAWE.

### c) Padrões sintáticos

Os padrões sintáticos (categorias gramaticais) encontrados em posição 1L no LOCNESS são com **verbo** (*is*) e com **conjunção** (*and*). Em posição 1R são com **artigo** (*the*) e **preposição** (*in*) que introduz um sintagma preposicional.

103) *This is especially true of the main theme: that life is futile and yet worth living to the full. (LOCNESS:sh\br6.txt)*

104) *British rail, now privatised, is under very fierce competition by cars and especially airplanes. (LOCNESS:els\4.txt)*

105) *THC affects the brain and circulatory system, especially the heart. It makes the heart beat faster-so. (LOCNESS:us75.txt)*

106) *For instance, surely the large rises in population and level of economic uncertainty - especially in the UK but also developed countries in general - must*

*have compelled many students to increase their educational level by going to university. (LOCNESS:Is\ab1.txt)*

No *corpus* BAWE, os padrões 1L são com **conjunção** (*and*) e com **verbo** (*is/are*). Os padrões 1R são com **artigo** (*the*), **pronome** (*it/that*) e com **preposição** (*in/for/of/to* – que pode ser também uma partícula de infinitivo)

*107) Many commentators argue that globalisation is just internationalisation of global trade and financial markets, having as its main actors the great powers **and especially** the U.S.A. (BAWE:0195a.txt)*

*108) A critical period is a time when synapses **are especially** plastic. (BAWE:0057c.txt)*

*109) By accepting the European material culture, the natives were thus impelled to accept European abstract culture, **especially the** European religion. (BAWE:0029n.txt)*

*110) The onset of illness, **especially that** which is not evidently self-limiting, fractures this social and cultural fabric, exposing the individual to threats of self-identity and a potentially damaging loss of control. (BAWE:0405a.txt)*

*111) The second section is very difficult to assess prior to any commercialization **especially for** a new intention. (BAWE:0296a.txt)*

Com relação à posição do advérbio nas orações, a diferença significativa encontra-se na posição do *especially* no final das orações e entre vírgulas, que são mais contempladas pelos aprendizes, com cerca de 4,5 % das ocorrências e 9% respectivamente, conforme ilustram tabela 27 abaixo e os exemplos 112, 113,114:

<i>ESPECIALLY</i>	início (%)	meio (%)	fim (%)	entre vírgulas (%)
Br-ICLE	4,5	82	4,5	9
LOCNESS	3	95	0	2
BAWE	3	96	0,5	0,5

Tabela 27: posição do advérbio *especially* em orações dos *corpora*.



112) *with the large masses. Finally, television, **especially**, for its strong advertising power (Br-ICLE:a1034.txt)*

113) *alised world, the logic of results and profit, **especially**. However, it is possible to ar (Br-ICLE:a1129.txt)*

114) *Toyota intends to improve the productivity for new models **especially**. (BAWE:0296d.txt)*

#### d) Padrões semânticos

No *corpus* Br-ICLE, predomina a função do advérbio como um particularizador de grupos nominais.

115) *Secondly, it is of great importance that society in general, but **especially the people who are in charge of making the laws** and of assuring that those laws be obeyed, start thinking about and try to find alternatives to the current prison system in order to substitute it by ways of rehabilitating the convicts. (Br-ICLE)*

No LOCNESS e no BAWE, o advérbio particulariza circunstâncias.

116) *This occurred for a variety of reasons which I will explore below but the point needs to be illustrated of how close Italy came to a real worker revolution and how **especially in the years 1919-20** they encountered little real resistance. (BAWE:0244f.txt)*

Um outro ponto a considerar se refere ao advérbio como intensificador de um adjetivo, em sua maioria, de conotação positiva, como *good, helpful, important, interesting, useful, valuable, true*, embora haja ocorrências com adjetivos de conotação negativa como *difficult, hard, problematic, vulnerable*.

117) *It is a technique which is less easy to utilize with other research methods, and is **especially useful** when discussing sensitive issues in political research such as certain government policies. (BAWE:0275b.txt)*

#### 4.8.3. Conclusões sobre o advérbio *especially*

Embora *especially* seja classificado como um adjunto conjuntivo, pode-se notar que o referido advérbio pode assumir a função de adjunto de modo (intensificador), quando anteceder um adjetivo.

A ocorrência de apenas um padrão de colocados no Br-ICLE aponta para a possibilidade de os aprendizes estarem fazendo escolhas lexicais muito diferentes das realizadas pelos nativos. A ausência de agrupamentos lexicais no *corpus* de aprendiz pode sugerir que, coletivamente, os aprendizes não tenham adquirido as variedades de combinações possíveis, ou em outras palavras as variedades ainda não tenham sido internalizadas.

#### 4.9. Conclusões sobre o capítulo de análise

Este capítulo buscou apresentar os dados extraídos dos *corpora* através da comparação de cinco advérbios em *-ly* mais frequentes do *corpus* de estudo com a ocorrência dos mesmos advérbios em dois *corpora* de referência. O capítulo começou apresentando o sobreuso de advérbios através do conceito de palavra-chave, após o que o trabalho mapeou a ocorrência e a distribuição de colocados desses mesmos advérbios, bem como seus agrupamentos lexicais, suas posições nas orações e em alguns casos, a prosódia semântica.

Vale ressaltar ainda que, embora a análise tenha se restringido a cinco advérbios, a Linguística de *Corpus* opera sob a perspectiva da linguagem enquanto probabilidade. Isto quer dizer que foi investigada neste trabalho a probabilidade de os usuários de inglês como língua materna usarem a forma A e não a forma B, quando escrevem academicamente. Da mesma forma, foi investigada a probabilidade de aprendizes brasileiros usarem coletivamente formas, combinações e determinada prosódia quando escrevem composições argumentativas. A pequena amostra de cinco advérbios sugere que as diferenças entre o uso em L1 e L2 pode se ater aos colocados ou ausência deles ou à posição nas orações, entre outros. Além disso, a observação dos 20 primeiros advérbios mais usados pelos sujeitos dos três *corpora* e dos advérbios-chave

mostra que os advérbios preferidos pelos aprendizes ocorrem com distribuição e frequência diferentes nos *corpora* de nativos e que há nesta lista ainda advérbios usados frequentemente apenas pelos aprendizes.

O próximo capítulo tratará, portanto, das conclusões a partir das análises feitas no presente capítulo, além de algumas possíveis aplicações pedagógicas do uso de *corpora* em sala de aula.

## 5. CONSIDERAÇÕES FINAIS

### 5.1. Percurso da dissertação e limitações

O presente capítulo faz um fechamento do trabalho, retomando os seus pontos principais, apontando algumas considerações e fazendo sugestões de pesquisa futura e possíveis aplicações pedagógicas dos resultados.

Conforme mencionado na Introdução, o presente trabalho pertence à área de pesquisa com *corpus* de aprendiz, que é na sua essência contrastiva: a produção do aprendiz é contrastada com a produção do ‘não aprendiz’, ou com as produções de vários grupos de aprendizes, ou com ambos. A importância de estudos nessa área se mostra de grande relevância, uma vez que permite identificar o léxico, a gramática e as características do discurso que diferenciam a produção de aprendizes de um possível alvo (seja ele de aprendizes cuja língua materna é a língua-alvo, ou de escritores publicados). Esses estudos são, portanto, uma rica fonte de dados para fins pedagógicos (Granger 2008).

Esta pesquisa teve como objetivo comparar o uso dos advérbios mais frequentes terminados em *-ly* em textos escritos em inglês por aprendizes brasileiros e os textos escritos por falantes nativos do inglês. Por ter usado um viés da Gramática Funcional, a pesquisa enfocou o uso desses advérbios como possíveis recursos para o enunciador criar espaço para fazer um comentário, passar julgamento sobre a veracidade ou probabilidade de uma proposição e estabelecer relações lógicas entre proposições. Ou seja, o advérbio visto como um recurso modalizador.

Para tanto, a pesquisa encontrou suporte teórico na Lingüística de *Corpus*, que é uma área de estudo que trata da coleta e exploração de *corpora*, com o objetivo de explorar a linguagem em sua manifestação real e autêntica. O trabalho lançou mão de alguns conceitos fundamentais desenvolvidos na Linguística de *Corpus* para o estudo do léxico, ou seja, os conceitos de colocação, de prosódia semântica, de agrupamentos lexicais ou clusters, de palavra-chave, isto tudo para dar conta de linguagem em uso por um grupo.

Além da Lingüística de *Corpus*, a pesquisa fundamentou-se na área de estudos chamada *Corpora* de Aprendizes, área pouco explorada no que tange a pesquisas sobre inglês de brasileiros, a qual conta com trabalhos de Granger (1998), Ringbom (1998), McEnery e Kifile (2000), Aijmer (2002) Römer (2004), Tankó (2004), Bolton, Nelson e Hung (2004), Nesselhauf (2004), Wei-yu Chen (2006), Neff et al (2006), Guo (2006), Gilquin, et al (2007), sobre o inglês de falantes de outras línguas.

A pesquisa também explorou advérbios sob a luz de duas gramáticas funcionais, uma teórica e a outra pedagógica de onde os parâmetros para análise dos dados foram tirados.

A metodologia empregada na pesquisa consistiu da comparação de um *corpus* de estudo, Br-ICLE, com dois *corpora* de referência, LOCNESS e BAWE - esses três *corpora* compostos por textos escritos em inglês, produzidos por aprendizes brasileiros e aprendizes nativos do inglês como língua materna e por falantes nativos proficientes respectivamente.

As perguntas básicas que orientaram esta dissertação foram:

a) que tipo de advérbios os aprendizes brasileiros de inglês como língua estrangeira investigados usam em ensaios argumentativos?

b) quais os mais freqüentes, como são posicionados na frase e com que função são usados?

c) como o seu uso se compara ao uso de advérbios em textos semelhantes em inglês como primeira língua, em termos de função e estrutura?

d) como os aprendizes brasileiros de língua inglesa constroem seu posicionamento no discurso através do uso de advérbios?

Os resultados indicaram, em resposta à primeira e à segunda perguntas de pesquisa, que os advérbios mais frequentes, de acordo com os critérios de corte apresentados no capítulo 4, são: *really*, que é normalmente usado pelos aprendizes no meio de orações, podendo aparecer em alguns casos no início. Sua função mais utilizada é como adjunto de modo (para intensificar); *simply*, que aparece no *corpus* de aprendiz em posições medial e final, é utilizado como um adjunto de modo (também para intensificar); *actually*, pode aparecer tanto em posição medial ou inicial e, sua função recorrente é como adjunto conjuntivo; *consequently*, ora é posicionado no início de orações, ora no meio ou entre vírgulas, mas nunca no final, e sua função básica é como adjunto conjuntivo; por fim, *especially* que apesar de ser amplamente utilizado no meio das orações pode ser também encontrado no início ou no fim das delas, funcionando em qualquer caso

como um adjunto conjuntivo. Isto posto, a utilização destes advérbios preferencialmente e nas porcentagens que foram usados, sugere um discurso focado na expressão de veracidade, realidade e intensidade, o que, por sua vez, empresta ao discurso características de assertividade.

Em relação à terceira pergunta, foi observado que os advérbios comparados e as diferenças notadas estão relacionados às porcentagens distintas dos advérbios encontrados nos *corpora* (seja por causa de sobreuso ou subuso destes elementos). Quando os itens comparados apresentam porcentagens semelhantes, contudo, o uso de colocados é “errôneo”, ou em termos de uso errôneo dos colocados propriamente ditos ou em termos de prosódia semântica. Entretanto, em termos de pesquisa linguística, considerar que as diferentes escolhas de advérbios, e as diferenças de frequência, função e colocados apresentadas no *corpus* de aprendiz seja um erro, é no mínimo um equívoco, se considerarmos que essas diferenças podem ser típicas do “inglês brasileiro”. Contudo, este é um fato que só pode ser atestado mediante uma pesquisa específica que considere essa variação.

Estas diferenças, marcadas coletivamente como sendo do inglês brasileiro podem estar ligadas a alguns fatores, tais como:

- a) O tamanho dos *corpora*: embora a linguística veja a linguagem como um sistema probabilístico, é fato que quanto maior o *corpus*, maior será a probabilidade de se observar padrões.
- b) A influência da língua materna na produção do texto em língua estrangeira. Uma pesquisa que compare textos do mesmo gênero escritos em português e inglês pode ajudar a identificar possíveis influências da língua materna.
- c) A internalização dos elementos linguísticos necessários para compor um texto em língua estrangeira. Considerando que a aquisição da linguagem é um processo, pode ser que os aprendizes ainda não tenham internalizado, por exemplo, as diferentes funções dos advérbios em língua inglesa.
- d) A falta de fluência dos aprendizes e o contexto de sala de aula das universidades formadoras de futuros professores de língua inglesa que provavelmente não estejam apresentando os advérbios como recursos importantes de modalização.

Dentro desta última possibilidade, pode-se dar como exemplo, a ementa a seguir, que refere-se ao estudo de modalização que é ensinado em Língua Inglesa III (terceiro período de Letras) de uma universidade pública do Rio de Janeiro..

- |       |   |
|-------|---|
| 1.    | SEMÂNTICA/PRAGMÁTICA                                  |
| 1.1   | Auxiliares modais                                     |
| 1.1.1 | Valores semânticos                                    |
| 1.1.2 | Referência temporal e aspectual                       |
| 1.1.3 | O alcance da negação em orações com auxiliares modais |
| 1.1.4 | O uso pragmático dos auxiliares modais                |

Quadro 1: ementa sobre estudo da modalização Língua Inglesa III

De acordo com a ementa apresentada, modalizar parece ser somente usar auxiliares modais, porém a presente pesquisa demonstrou que não é. Ainda dentro do contexto de sala de aula, a próxima ementa, que abrange o sintagma adverbial, é apresentada em Língua Inglesa VI (sexto período de Letras) na mesma universidade, após os alunos terem estudado um semestre sobre o ensaio literário.

- |       |   |
|-------|---|
| 1.1   | O sintagma adverbial  |
| 1.1.1 | A estrutura do sintagma adverbial   |
| 1.1.2 | Advérbios e adjuntos adverbiais   |
| 1.1.3 | Classificação dos adjuntos adverbiais de acordo com o grau de integração na estrutura da oração: adjuntos adverbiais de oração ("adjuncts" e "subjuncts") e adjuntos adverbiais de sentença ("disjuncts" e "conjuncts") |
| 1.1.4 | Classificação semântica dos adjuntos adverbiais   |
| 1.3.5 | Colocação dos adjuntos adverbiais   |

Quadro 2: ementa sobre sintagma adverbial

Nota-se também que não há sequer uma menção do advérbio como fator modalizador, ou seja, como marcador em potencial da voz de quem escreve. Usar um advérbio de forma sutil e informada significa ter a sua voz em qualquer texto.

E finalmente, em relação à quarta pergunta, a pesquisa revelou que a escolha dos aprendizes por aqueles advérbios que exprimem veracidade, realidade e intensidade cria a impressão de um discurso muito assertivo, distante do que se convencionou chamar de argumentação e deste modo temos que falar sobre uma das limitações importantes da presente dissertação, que é típica dos trabalhos com *corpora* – o fato de que em nenhum momento enfocou-se o tipo de texto produzido pelos aprendizes, ou seja a redação argumentativa. Apesar de esta dissertação não ter levado em conta o que é argumentar por escrito num texto minimamente acadêmico, a compilação dos *corpora* contempla isto. Pode ser que os alunos brasileiros não saibam o que é argumentar – argumentar não é asseverar, mas discutir e avaliar de forma concisa.

“O ensaio argumentativo é um gênero da escrita que requer que o aluno investigue um tópico, colete, gere e avalie evidências e estabeleça uma posição no tópico de maneira concisa (<http://owl.english.purdue.edu/owl/resource/685/05/>)”<sup>50</sup>.

## 5.2. Aplicações pedagógicas e futuras pesquisas

Uma possível aplicação pedagógica para os resultados desta pesquisa é a utilização de concordâncias a partir de *corpora* escritos em inglês, que podem ser apresentados na linguagem de sala de aula e/ou podem ser usados em materiais de ensino de língua inglesa. A concordância é empregada para exemplificar o uso de traços linguísticos e as situações nas quais ocorre e é um meio versátil de focalizar várias questões de sentido e estrutura. (Berber Sardinha 2004:273 e 276).

Alguns exemplos (ainda esporádicos) de como se trabalhar com *corpus*, podem ser encontrados, na *Natural Grammar* (Thornbury 2004:155 *apud* O’Keeffe 2007) onde é pedido que

---

<sup>50</sup> Tradução do original em inglês: The argumentative essay is a genre of writing that requires the student to investigate a topic, collect, generate, and evaluate evidence, and establish a position on the topic in a concise manner.



a palavra *there* seja identificada como pronome ou como advérbio, de acordo com as linhas de concordância apresentadas

## Exercises

① Look at these concordance lines, and identify the meaning of *there* in each case. Is it a pronoun (showing that something exists) or is it an adverb (saying where something is)?

- a There's a bar and a lecture room for guests' use.
- b There'd been another quake at 4am, a 6.5 shock.
- c It was only in my third year that I really felt happy there.
- d You say there's a certain amount of risk. How much?
- e I was there for her birth and it was the most exciting thing.
- f But there'll be no alcohol on sale.
- g He was standing there with Mrs Kasmin as she tried to give him tea.
- h He had been there since he left the Pit a year earlier.
- i He was confident there'd be no problem. So was I.

Figura 7: Extraído da *Natural Grammar* (Thornbury 2004:155 apud O'Keeffe 2007)

Sugerimos ainda, como outro exemplo de estudo com *corpora* em sala de aula, que os alunos comparem, por meio das linhas de concordância, advérbios como *probably* e *possibly*, considerados sinônimos, em termos de função e padrão de colocados em *corpora* de nativos. Através desta análise, é possível identificar semelhanças e diferenças entre esses advérbios quanto ao seu uso em contextos típicos destes *corpora*.

Este tipo de abordagem, orientada por um *corpus*, visa tornar o aluno um pesquisador, permitindo que busquem suas próprias respostas ao trabalhar com o computador ou com concordâncias impressas preparadas pelo professor (Berber Sardinha 2004:291).

Dito isto, a presente dissertação pode ser vista como uma singela contribuição para a área da Lingüística de *Corpus* aplicada à descrição lingüística baseada em *corpus* e abordou questões sobre a importância da descrição de *corpus* de aprendiz, para que pesquisadores e professores possam trabalhar em conjunto. Além disso, o presente trabalho visou fornecer uma contribuição original para a área de *Corpora* de Aprendiz, visto que há escassez de trabalhos que lidem com a

investigação das escolhas típicas da linguagem escrita por aprendizes brasileiros de língua inglesa. Como dissertação, entretanto, ela sinaliza, somente um início de investigação.

Para pesquisas futuras é importante primeiro que o *corpus* Br-ICLE seja aumentado. Compilar textos argumentativos dentro de universidades é uma tarefa sobremaneira difícil, o que pode ser constatado pelo número de redações até hoje coletada desde o início do projeto. Isso porque os universitários nem sempre querem contribuir para a pesquisa com seus textos. Se aumentado, o Br-ICLE poderia ser objeto de outros pontos de análise para um futuro foco, “onde os advérbios puderiam ser analisados sob o arcabouço da avaliatividade de Martin e White através do subsistema da graduação e do engajamento”<sup>51</sup>. Com o aumento do Br-ICLE poderia se considerar também a preparação de um material didático compatível com a população analisada (redação acadêmica em inglês por universitários).

---

<sup>51</sup> Vereza (2010) comunicação pessoal

## REFERÊNCIAS

- ADOLPHS, S. *Introducing Electronic Text Analysis: A practical guide for language and literary studies*. London: Routledge. 2006.
- AIJMER, K. Modality in advanced Swedish learners' written interlanguage. In GRANGER, S., HUNG, J., PETCH-TYSON (eds.). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam: John Benjamins. 2002.
- BARLOW, M. Computer-based analyses of learner language. In ELLIS, R.; BARKWIZEN, G. (eds.) *Analyzing learner language*. London: Oxford. 2005. p. 1-20.
- BERBER SARDINHA, T. A. *Linguística de Corpus*. São Paulo: Manole, 2004. passim
- BERNARDINI, S. "Corpora in the classroom – an overview and some reflections on the future developments". In SINCLAIR, J.M. (ed.) *How to use corpora in language teaching*. Amsterdam: John Benjamins. 2004. p. 15-36
- BIBER, D. et al. *Longman Grammar of Spoken and Written English*. 1999.
- BLOOR, T; BLOOR, M. *The Functional Analysis of English: A Hallidayan Approach*. London:Arnold. 1995.
- BOLTON, K, NELSON, G., HUNG, J. A corpus-based study of connectors in student writing. In HYLAND, K. *Teaching and researching writing*. London: Longman.2002. p. 49-84.
- BUTT, D. et al.. *Using Functional Grammar: An Explorer's Guide*. Sydney: Macquarie University. 1995
- CONRAD, S. Corpus linguistics, language variation and language teaching. In SINCLAIR, J.M. (ed.) *How to use corpora in language teaching*. Amsterdam: John Benjamins. 2004. p. 67-85.
- DROGA, L; HUMPHREY, S. *Getting Started with Functional Grammar*. Berry, Australia:2002 passim
- GILQUIN, G., GRANGER, S., PAQUOT, M. Learner corpora: the missing link in EAP pedagogy. *Journal of English or Academic Purposes*. V.6. October 2007.
- GRANGER, S. The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In AIJMER, K (ed.), *Corpora and Language Teaching*. Benjamins.2008

GREENBAUM, S. *The Oxford Grammar*. Oxford University Press:1996

GUO, X. Modal auxiliaries in phraseology – a contrastive study of learner English and NS english. In Danielsson P. and M. Wagenmakers (eds) *Proceedings from the Corpus Linguistics Conference Series*. V. 1, n. 1 *Corpus Linguistics* 2005.

KENNEDY, G. *An Introduction to Corpus Linguistics*. London: Longman . 1998

LEECH, G.; DEUCHAR, M.; HOOGEHRAAD, R. *English Grammar for Today*. London: Macmillan. 1982

McENERY, T. ; KIFLE, N.A. Epistemic modality in argumentative essays of second language writers. London: Pearson Education. 2002.

McENERY, T. ; WILSON, A. *Corpus Linguistics*. Edinburgh: Edinburgh University Press. 1996

MELINDA, T. Authentic language or language errors. *ELT Journal* V. 59/2. April 2005

NASSELHAUF, N. Learner *corpora* and their potential for language teaching. In SINCLAIR, J.M. (ed.) *How to use corpora in language teaching*. Amsterdam: John Benjamins. 2004. p. 125-152.

NESSSELHAUF, N. *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.2005

NEEF, J. et al. *Contrasting learner corpora: the use of modal and reporting verbs in the expression of writer stance*. In GRANGER, S. ; PETCH-TYSON, S. *Extending the scope of corpus-based research: new applications, new challenges*. Rodopi.2003

NEFF, J. et al. *Formulating writer stance: a contrastive study of EFL learner corpora*. In GERBIG A. ; MASON O. *Language, People, Numbers: Corpus Linguistics and Society*. Rodopi 2008.

O'KEEFFE, A; McCARTHY, M.; CARTER, R. *From Corpus to Classroom*. USA: Cambridge. 2007

RINGBOM, H. *Vocabulary frequencies in advanced learner English: a cross-linguistic approach*. In GRANGER, S. (ed). *Learner English on Computer*. London: Longman. 1998. p. 41-52.

RÖMER, U. *A corpus-driven approach to modal auxiliaries and their didactics*. In SINCLAIR, J.M. *How to use corpora in language teaching*. Amsterdam: John Benjamins. 2004. p. 185-199.

SCOTT, M. ; TRIBBLE, C. *Textual Patterns: key-words and corpus analysis in language education*. Amsterdam: John Benjamins Publishing Company. 2006

SIMON-VANDERBERG, A.M.; AIJMER, K. *The Semantic Field of Modal Certainty*. German: Mouton de Gruyter. 2007

TANKÓ, G. The use of adverbial connectors in Hungarian university students' argumentative essays. In SINCLAIR, J.M.(ed.) *How to use corpora in language teaching*. Amsterdam: John Benjamins. 2004. p. 157-181.

TOGNINI-BONELLI, E. *Corpus Linguistics at Work: Studies in Corpus Linguistics*. Amsterdam: John Benjamins Publishing Company. 2001

TSUI, A.B.M. What teachers have always wanted to know – and how *corpora* can help. In SINCLAIR, J.M. (ed.) *How to use corpora in language teaching*. Amsterdam: John Benjamins. 2004. p. 39-61

WEI-YU CHEN, C. The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, Amsterdam, v. 11, n 1, 2006

## ANEXO 1: descrição dos corpora

### ICLE DESCRIPTION (o corpus Br-ICLE segue os critérios do ICLE)

The ICLE learner profile has been created in order to provide researchers with information about contributors which will enable meaningful conclusions to be drawn from the results obtained when the corpus is analysed. Using the profile, it will be possible both to draw general conclusions about advanced learner writing, and also to examine subsections e.g. Spanish mother tongue learners, learners who speak some English at home, learners for whom German is the second language and English is the third language. It will also be possible to examine more sociolinguistic aspects such as for instance male/female comparisons. If the corpus is used as a basis for developing specifically adapted teaching tools, the potential advantages of this facility are clear.

#### Collect the right type of material

The corpus will consist entirely of essay writing. Two types of essay writing are useful:

Argumentative essay writing

Using titles such as the ones below :

- "Crime does not pay"
- "Feminism has done more harm to the cause of women than good"
- "Pollution : a silent conspiracy"

These essays may be done by students in their own time (untimed), using language reference tools (dictionaries, grammars, etc.) but should be entirely the students' own work, i.e. they should not draw on other articles, books for the essay and should not ask a native speaker of English for help. Alternatively, they may also be done under examination conditions.

Essays can be completed at home (untimed), should be at least 500 words long (up to 1,000), and handed in on disk - plain text). This reduces the time spent typing up student essays and minimises the risk of introducing errors into the text. Work should be entirely the students' own, no help should be sought from third parties, but they may use reference tools such as dictionaries and grammar books (use of reference tools should be indicated on the learner profile questionnaire).

**Important note :** the essays should be at least 500 words long (up to 1,000).

### 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

59,568 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





<b>AMERICAN ARGUMENTATIVE ESSAYS</b>
--------------------------------------

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

<b><u>Marquette University</u> (codes: ICLE-US-MRQ-0001.1-46.1)</b>
---

**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:

- |  |   |
|--|---|
| <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> </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> </ul> |
|--|---|

<b><u>Indiana University at Indianapolis</u> (codes: ICLE-US-IND-0001.1-28.1)</b>
---

**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

<b><u>Presbyterian College, South Carolina (codes: ICLE-US-PRB-0034.2-39.2)</u></b>
---

**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

<b><u>University of South Carolina</u></b>
--

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

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

### **13 essays**

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

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

<b><u>University of Michigan</u> (codes: ICLE-US-MICH-0001.1-45.1)</b>
--

**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.)



<b>AMERICAN LITERARY-MIXED ESSAYS</b>
---------------------------------------

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

<b><u>Presbyterian College, South Carolina</u></b>
--

**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').

## **BAWE DESCRIPTION**

### **1. Collection**

The collection process and decisions made during collection are described in detail in Alsop and Nesi (under review); a summary description of the most significant points is given here.

Student assignments were collected from three universities: Oxford Brookes, Reading and Warwick. Assignments were collected from 35 disciplines (see Table 2), in 4 broad disciplinary groupings (DGs: see Table 1), and from students in each of three undergraduate years and those on masters courses.

Table 1 shows the numbers of students, assignments, texts and words collected in each year in each DG. The difference in the numbers of assignments and texts is

due to the fact that some assignments (**compound assignments**, see 6.3) consist of more than one text.

**Table 1. Numbers of students, assignments, texts and words by disciplinary grouping and year**

disciplinary group		Yr 1	Yr 2	Yr 3	Masters	Total
<b>Arts and Humanities</b>	students	101	83	61	23	268
	assignments	239	228	160	78	705
	texts	259	231	161	83	734
	words	468,353	583,617	427,942	234,206	1,714,118
<b>Life Sciences</b>	students	74	71	42	46	233
	assignments	180	193	113	197	683
	texts	191	208	119	203	721
	words	299,370	408,070	263,668	441,283	1,412,391
<b>Physical Sciences</b>	students	73	60	56	36	225
	assignments	181	149	156	110	596
	texts	186	156	169	129	640
	words	300,989	314,331	426,431	339,605	1,381,356
<b>Social Sciences</b>	students	85	88	75	62	313 <sup>1</sup>
	assignments	207	197	162	202	777 <sup>2</sup>
	texts	218	202	169	204	802
	words	371,473	475,668	440,674	688,921	1,999,130 <sup>1</sup>
<b>Total students</b>		<b>333</b>	<b>302</b>	<b>234</b>	<b>167</b>	<b>1039<sup>1</sup></b>
<b>Total assignments</b>		<b>807</b>	<b>767</b>	<b>591</b>	<b>6587</b>	<b>2761<sup>2</sup></b>
<b>Total texts</b>		<b>854</b>	<b>797</b>	<b>618</b>	<b>619</b>	<b>2897<sup>3</sup></b>
<b>Total words</b>		<b>1,440,185</b>	<b>1,781,686</b>	<b>1,558,715</b>	<b>1,704,015</b>	<b>6,506,995<sup>4</sup></b>

<sup>1</sup> Includes 3 of unknown level.

<sup>2</sup> Includes 9 of unknown level.

<sup>3</sup> Includes 9 of unknown level.

<sup>4</sup> Includes 22,394 in texts of unknown level.

Table 2 shows the number of assignments by discipline and year.

**Table 2. Number of assignments by discipline and year**

disciplinary group	discipline	1	2	3	4	Total
<b>Arts and Humanities</b>	Archaeology	23	21	15	17	76
	Classics	33	27	15	7	82
	Comparative American Studies	29	26	13	6	74
	English	35	35	28	8	106
	History	30	32	31	3	96
	Linguistics	27	31	24	33	115
	Other	19	22	9	0	50
	Philosophy	43	34	25	4	106



	<b>Total</b>	239	228	160	78	705
<b>Life Sciences</b>	Agriculture	35	35	30	34	134
	Biological Sciences	52	50	26	41	169
	Food Sciences	26	36	32	30	124
	Health	35	33	12	1	81
	Medicine	0	0	0	80	80
	Psychology	32	39	13	11	95
	<b>Total</b>	180	193	113	197	683
<b>Physical Sciences</b>	Architecture	2	4	2	1	9
	Chemistry	23	24	29	13	89
	Computer Science	34	13	30	10	87
	Cybernetics & Electronics	4	4	13	7	28
	Engineering	59	71	54	54	238
	Mathematics	8	5	12	8	33
	Meteorology	6	9	0	14	29
	Other	0	1	0	0	1
	Physics	37	14	14	3	68
	Planning	8	4	2	0	14
	<b>Total</b>	181	149	156	110	596
<b>Social Sciences</b>	Anthropology	14	12	6	17	49
	Business	32	33	31	50	146
	Economics	30	30	23	13	96
	HLT/M	14	21	29	29	93
	Law	37	37	31	28	134*
	Other	0	2	3	4	9
	Politics	37	33	15	25	110
	Publishing	11	4	0	15	30
	Sociology	32	25	24	21	110
	<b>Total</b>	207	197	162	202	777
<b>Total</b>		807	767	591	587	2761

\* Includes 1 of unknown year.

† Includes 8 of unknown year.

‡ Includes 9 of unknown year.

Students were paid £3 for each assignment submitted (towards the end of the collection period students in some underrepresented disciplines were rewarded at £5 and even £10 per assignment) and were asked to sign disclaimer forms assigning copyright to the respective university (disclaimer forms are given in Appendix 1). Contextual information about the student and the assignment were also collected at this stage. Students were asked to supply all the contextual information to be contained in the XML file headers (see below).

After all assignments were collected, a number of analyses were carried out, including a Multi-Dimensional Analysis (Conrad and Biber 2001) of the registers of the texts, to be described in future publications, and a genre analysis described in the following section.

## 2. Genre and genre family

All assignments in the corpus were scrutinised for generic properties, and a large number of genres were identified. These were collected into 13 **genre families** (GFs), classes of genres sharing functional and structural properties. The full set of genres and GFs is given in Appendix 2. A list of the GFs, and their distribution by DG, is shown in Table 3.<sup>1</sup> The distribution of GFs by discipline is shown in the following tables.

**Table 3. Distribution of GFs by DG**

	Arts and Humanities	Life Sciences	Physical Sciences	Social Sciences	Total
case study	0	91	37	66	194
critique	48	84	76	114	322
design specification	1	2	87	3	93
empathy writing	4	19	9	3	35
essay	602	127	65	444	1238
exercise	14	33	49	18	114
explanation	9	117	65	23	214
literature review	7	14	4	10	35
methodology recount	18	158	170	16	362
narrative recount	10	25	21	19	75
problem question	0	2	6	32	40
proposal	2	26	19	29	76
research report	9	22	16	14	61
Total	724	720	624	791	2859

**Table 4. Distribution of GFs by discipline, Arts and Humanities disciplines**

AH	Arch	Classics	CAS	English	History	Ling	Other	Phil	Total
case study	0	0	0	0	0	0	0	0	0
critique	15	2	2	1	1	21	0	6	48
design specification	1	0	0	0	0	0	0	0	1
empathy writing	0	0	0	4	0	0	0	0	4
essay	49	78	71	89	94	75	48	98	602
exercise	1	0	0	7	1	5	0	0	14
explanation	2	0	0	0	0	7	0	0	9
literature review	0	1	0	5	0	0	0	1	7
methodology recount	7	0	0	0	0	9	2	0	18
narrative recount	1	0	0	4	0	5	0	0	10
problem question	0	0	0	0	0	0	0	0	0
proposal	0	1	0	1	0	0	0	0	2
research report	1	1	1	0	0	5	0	1	9
Total	77	83	74	111	96	127	50	106	724

**Table 5. Distribution of GFs by discipline, Life Sciences disciplines**

LS	Agriculture	BioSci	FoodSci	Health	Med	Psych	Total
case study	12	0	2	8	69	0	91
critique	37	20	9	9	1	8	84
design specification	1	0	0	0	0	1	2
empathy writing	2	0	13	1	1	2	19
essay	27	11	7	15	10	57	127
exercise	7	7	18	0	0	1	33
explanation	30	63	7	13	1	3	117
literature review	4	3	4	3	0	0	14
methodology recount	7	58	82	1	0	10	158

narrative recount	1	2	0	20	2	0	25
problem question	0	0	1	1	0	0	2
proposal	6	3	3	13	0	1	26
research report	1	8	0	1	0	12	22
Total	135	175	146	85	84	95	720

**Table 6. Distribution of GFs by discipline, Physical Sciences disciplines**

PS	Arch	Chem	CS	C&E	Eng	Math	Met	Phys	Plan	Total
case study	0	2	2	0	33	0	0	0	0	37
critique	1	11	8	3	31	6	4	10	2	76
design specification	0	0	41	13	28	0	3	2	0	87
empathy writing	1	0	2	0	1	4	0	1	0	9
essay	4	6	9	2	16	4	0	12	12	65
exercise	1	6	8	4	10	15	4	1	0	49
explanation	0	10	16	0	16	3	5	15	0	65
literature review	0	0	1	2	1	0	0	0	0	4
methodolog y recount	0	51	3	1	85	0	14	18	0	170
narrative recount	1	2	2	1	12	1	0	1	0	21 <sup>1</sup>
problem question	0	0	0	0	6	0	0	0	0	6
proposal	1	0	6	0	10	1	0	1	0	19
research report	0	1	2	2	4	0	0	7	0	16
Total	9	89	100	28	251	34	30	68	14	624 <sup>2</sup>

<sup>1</sup> Includes 1 from another discipline

<sup>2</sup> Includes 1 from another discipline

**Table 7. Distribution of GFs by discipline, Social Sciences disciplines**

SS	Anth	Bus	Econ	HLTM	Law	Pol	Publ	Soc	Total
case study	0	31	1	27	1	0	5	1	66

critique	8	29	17	11	25	11	2	11	114
design specification	0	2	0	0	0	0	0	0	3
empathy writing	0	0	0	1	0	0	2	0	3 <sup>1</sup>
essay	27	49	55	29	85	97	4	91	444
exercise	0	12	5	0	0	0	1	0	18
explanation	3	5	0	5	0	0	10	0	23
literature review	4	0	0	2	1	2	1	0	10
methodology recount	2	2	10	1	0	0	0	1	16 <sup>2</sup>
narrative recount	2	4	0	5	0	0	5	2	19
problem question	0	9	2	2	19	0	0	0	32
proposal	3	2	0	12	3	0	7	2	29 <sup>3</sup>
research report	0	1	7	3	0	0	0	2	14
Total	49	146	97	98	134	110	38	110	791 <sup>4</sup>

<sup>1</sup> Includes 7 from other disciplines.

<sup>2</sup> Includes 1 from another discipline.

<sup>3</sup> Includes 1 from another discipline.

<sup>4</sup> Includes 9 from other disciplines.

**ANEXO 2:** lista completa dos colocados e agrupamentos lexicais dos advérbios da pesquisa no corpus BAWE

As tabelas devem ser lidas da seguinte forma: a coluna *Word* representa as palavras que ocorrem ao redor da palavra de busca (posição 1). A coluna *Total* representa quantas vezes a palavra apareceu no corpus junto da palavra de busca. A coluna *Left*, representa quantas vezes a palavra apareceu à esquerda da palavra de busca e a coluna *Right*, quantas vezes ela apareceu à direita da palavra de busca. A coluna com asterisco (\*) representa a palavra de busca propriamente dita.

As tabelas dos agrupamentos lexicais são compostas por grupos de palavras formados por três elementos incluindo a palavra de busca. Estão listados em ordem decrescente de ocorrência.

Colocados do advérbio *really*

N	WORD	TOTAL	LEFT	RIGHT	L1	*	R1
1	REALLY	764	2	0	2	762	0
2	NOT	90	85	5	85	0	5
3	IS	83	75	8	75	0	8
4	TO	80	14	66	14	0	66
5	THE	69	4	65	4	0	65
6	IT	45	32	13	32	0	13
7	IN	35	0	35	0	0	35
8	ONLY	33	28	5	28	0	5
9	WAS	33	31	2	31	0	2
10	A	32	10	22	10	0	22
11	WHAT	27	16	11	16	0	11
12	ARE	26	24	2	24	0	2
13	THIS	26	15	11	15	0	11
14	AND	24	4	20	4	0	20
15	I	24	16	8	16	0	8
16	NEVER	23	23	0	23	0	0
17	FOR	21	0	21	0	0	21
18	THEY	20	17	3	17	0	3
19	THAT	18	6	12	6	0	12
20	OF	15	4	11	4	0	11
21	CAN	14	14	0	14	0	0
22	OR	14	4	10	4	0	10
23	WE	14	12	2	12	0	2
24	AS	11	0	11	0	0	11
25	BE	11	9	2	9	0	2
26	CANNOT	11	11	0	11	0	0
27	ABOUT	10	1	9	1	0	9
28	ON	10	0	10	0	0	10
29	DIDN'T	9	9	0	9	0	0
30	HE	9	7	2	7	0	2
31	WERE	9	8	1	8	0	1
32	WHEN	9	3	6	3	0	6

33	THERE	8	8	0	8	0	0
34	WHO	8	5	3	5	0	3
35	WITH	8	0	8	0	0	8
36	AT	7	0	7	0	0	7
37	BY	7	0	7	0	0	7
38	DON'T	7	7	0	7	0	0
39	HAVE	7	5	2	5	0	2
40	HAS	6	6	0	6	0	0
41	IF	6	1	5	1	0	5
42	ISN'T	6	6	0	6	0	0
43	SHOULD	6	5	1	5	0	1
44	AN	5	0	5	0	0	5
45	ANY	5	0	5	0	0	5
46	ME	5	0	5	0	0	5
47	PEOPLE	5	2	3	2	0	3

Agrupamentos lexicais com advérbio *really*

N	cluster	Freq.
1	is not really	23
2	it is really	16
3	does not really	13
4	are not really	11
5	only really be	11
6	not really a	10
7	can only really	9
8	it really is	8
9	it is not	6
10	was not really	6
11	can be really	5
12	did not really	5
13	do not really	5
14	is it really	5
15	is really the	5
16	really have to	5
17	really interested in	5
18	really want to	5
19	they really are	5
20	this is really	5
21	what is really	5
22	and it is	4
23	can never really	4
24	can really be	4
25	has never really	4
26	in order to	4
27	is really a	4
28	is really being	4

29	is really interesting	4
30	is this really	4
31	it does not	4
32	it really was	4
33	not really need	4
34	really need a	4
35	really need to	4
36	really possible to	4
37	that it is	4
38	they never really	4
39	to really understand	4
40	we are really	4
41	we cannot really	4
42	what really is	4
43	and colonies really	3
44	and not really	3
45	are coal and	3
46	are really the	3
47	as it really	3
48	can we really	3
49	coal and colonies	3
50	colonies really crucial	3
51	depends on the	3
52	didn't really have	3
53	how it really	3
54	i really enjoyed	3
55	is politics really	3
56	is there really	3
57	it's not really	3
58	know what to	3
59	not really possible	3
60	of what really	3
61	one can really	3
62	order to really	3
63	politics really about	3
64	really about people	3
65	really is a	3
66	really is to	3
67	really know what	3
68	really the case	3
69	really the key	3
70	really the only	3
71	really thought of	3
72	really understand what	3
73	really wanted to	3
74	should really be	3
75	than it really	3
76	that they really	3



77	the idea of	3
78	this only really	3
79	was really like	3
80	was really needed	3
81	we can really	3
82	what really happened	3
83	what was really	3
84	whether or not	3

Colocados do advérbio *simply*

N	WORD	TOTAL	LEFT	RIGHT	L1	*	R1
1	SIMPLY	1485	1	1	1	1483	1
2	IS	194	185	9	185	0	9
3	THE	179	9	170	9	0	170
4	TO	134	49	85	49	0	85
5	NOT	130	89	41	89	0	41
6	A	69	17	52	17	0	52
7	IT	58	28	30	28	0	30
8	OR	55	50	5	50	0	5
9	WAS	51	51	0	51	0	0
10	AS	48	27	21	27	0	21
11	THAT	45	8	37	8	0	37
12	OF	44	10	34	10	0	34
13	THAN	44	44	0	44	0	0
14	ARE	42	40	2	40	0	2
15	BY	36	28	8	28	0	8
16	AND	34	19	15	19	0	15
17	MAY	32	32	0	32	0	0
18	WERE	30	30	0	30	0	0
19	BE	26	17	9	17	0	9
20	WE	26	21	5	21	0	5
21	THEY	24	16	8	16	0	8
22	BUT	22	20	2	20	0	2
23	THIS	21	12	9	12	0	9
24	COULD	20	17	3	17	0	3
25	WILL	19	19	0	19	0	0
26	ON	18	1	17	1	0	17
27	HE	17	11	6	11	0	6
28	WOULD	17	17	0	17	0	0
29	CAN	16	16	0	16	0	0
30	IN	16	3	13	3	0	13
31	QUITE	16	16	0	16	0	0
32	FROM	14	7	7	7	0	7

33	THEM	14	3	11	3	0	11
34	FOR	13	2	11	2	0	11
35	THEIR	13	0	13	0	0	13
36	WITH	13	3	10	3	0	10
37	PUT	12	12	0	12	0	0
38	THERE	12	6	6	6	0	6
39	BEAM	11	0	11	0	0	11
40	CANNOT	11	10	1	10	0	1
41	I	11	6	5	6	0	5
42	SHOULD	11	11	0	11	0	0
43	AN	10	0	10	0	0	10
44	HAD	10	8	2	8	0	2
45	MORE	10	4	6	4	0	6
46	OUT	10	1	9	1	0	9
47	WHICH	10	9	1	9	0	1
48	ALL	8	0	8	0	0	8
49	DONE	8	6	2	6	0	2
50	WOMEN	8	5	3	5	0	3
51	AT	7	0	7	0	0	7
52	IF	7	0	7	0	0	7
53	NO	7	1	6	1	0	6
54	THESE	7	1	6	1	0	6
55	WHAT	7	0	7	0	0	7
56	BEING	6	3	3	3	0	3
57	CASE	6	3	3	3	0	3
58	HAS	6	6	0	6	0	0
59	HAVE	6	6	0	6	0	0
60	MATTER	6	0	6	0	0	6
61	PEOPLE	6	3	3	3	0	3
62	THEN	6	5	1	5	0	1
63	WHO	6	6	0	6	0	0
64	ANY	5	2	3	2	0	3
65	ENOUGH	5	0	5	0	0	5
66	HER	5	0	5	0	0	5
67	INTO	5	1	4	1	0	4
68	SHE	5	5	0	5	0	0
69	SOME	5	1	4	1	0	4

Agrupamentos lexicais com o advérbio *simply*

N	cluster	Freq.
1	is simply a	40
2	it is simply	29
3	is not simply	25
4	rather than simply	20
5	is simply the	18
6	not simply a	16

7	is simply not	15
8	this is simply	15
9	simply do not	13
10	are not simply	12
11	more than simply	12
12	may simply be	11
13	simply because it	11
14	was not simply	11
15	simply because of	10
16	that it is	10
17	is simply an	9
18	it was simply	9
19	simply supported beam	9
20	as simply a	8
21	not simply the	8
22	than simply a	8
23	was simply not	8
24	a matter of	7
25	a simply supported	7
26	because of the	7
27	can simply be	7
28	could simply be	7
29	program should simply	7
30	simply be a	7
31	simply because the	7
32	simply because they	7
33	simply did not	7
34	simply-supported beam	7
35	the program should	7
36	was simply a	7
37	and not simply	6
38	are simply a	6
39	cannot simply be	6
40	does not simply	6
41	is simply because	6
42	is simply too	6
43	or simply a	6
44	simply a matter	6
45	simply part of	6
46	than simply the	6
47	the fact that	6
48	the use of	6
49	there is simply	6
50	they are simply	6
51	to simply say	6
52	are simply not	5
53	be done simply	5
54	because it is	5

55	for the sake	5
56	is simply to	5
57	it could simply	5
58	it is not	5
59	it may simply	5
60	it simply means	5
61	it was not	5
62	simply by the	5
63	simply does not	5
64	simply due to	5
65	simply for the	5
66	simply say that	5
67	simply takes the	5
68	that we simply	5
69	they were simply	5
70	to simply be	5
71	are simply the	4
72	as opposed to	4
73	because they are	4
74	by a user	4
75	called by a	4
76	do not simply	4
77	have been simply	4
78	i will simply	4
79	is simply connected	4
80	it simply takes	4
81	not simply to	4
82	of a simply	4
83	opposed to simply	4
84	or simply because	4
85	possible to simply	4
86	simply and solely	4
87	simply as a	4
88	simply be the	4
89	simply because he	4
90	simply because there	4
91	simply concerned with	4
92	simply means that	4
93	simply on the	4
94	simply one of	4
95	simply refer to	4
96	simply unable to	4
97	takes the two	4
98	that they are	4
99	the simply-supported	4
100	the two lists	4
101	this is not	4
102	to put it	4

103	two lists to	4
104	was simply the	4
105	which is simply	4
106	a case of	3
107	a result of	3
108	a simply-supported	3
109	an extension of	3
110	and are not	3
111	as simply the	3
112	be done by	3
113	be seen as	3
114	because there is	3
115	between two objects	3
116	but it is	3
117	can be simply	3
118	cannot be simply	3
119	do not know	3
120	due to the	3
121	in the first	3
122	in this situation	3
123	instead of simply	3
124	is simply and	3
125	is simply defined	3
126	is simply no	3
127	it as a	3
128	it could be	3
129	may be simply	3
130	measurements on the	3
131	member of the	3
132	not possible to	3
133	not simply because	3
134	not to be	3
135	of all the	3
136	of the simply	3
137	of the system	3
138	on the other	3
139	on the simply	3
140	or simply the	3
141	or simply to	3
142	part of the	3
143	put it simply	3
144	referred to as	3
145	simply a result	3
146	simply a tool	3
147	simply an extension	3
148	simply be seen	3
149	simply being a	3
150	simply changing the	3

151	simply connected if	3
152	simply consists of	3
153	simply could not	3
154	simply defined as	3
155	simply had to	3
156	simply has to	3
157	simply in the	3
158	simply is not	3
159	simply not enough	3
160	simply referred to	3
161	simply that it	3
162	simply that the	3
163	simply the ability	3
164	simply the need	3
165	simply to be	3
166	that it was	3
167	that we do	3
168	the ability to	3
169	the death penalty	3
170	the idea that	3
171	the number of	3
172	the other hand	3
173	the sake of	3
174	the simply support	3
175	then we simply	3
176	they will simply	3
177	this was simply	3
178	to be a	3
179	to be made	3
180	was simply an	3
181	we simply do	3
182	were simply the	3
183	would simply be	3

Colocados do advérbio *actually*

N	WORD	TOTAL	LEFT	RIGHT	L1	*	R1
1	ACTUALLY	1180	0	0	0	1180	0
2	IS	160	151	9	151	0	9
3	THE	123	3	120	3	0	120
4	TO	108	26	82	26	0	82
5	WAS	70	68	2	68	0	2
6	ARE	68	66	2	66	0	2
7	IN	68	3	65	3	0	65
8	NOT	62	55	7	55	0	7
9	AND	60	25	35	25	0	35
10	IT	48	26	22	26	0	22

11	A	40	2	38	2	0	38
12	MAY	34	34	0	34	0	0
13	CAN	33	32	1	32	0	1
14	BY	31	5	26	5	0	26
15	THAT	31	14	17	14	0	17
16	THEY	30	28	2	28	0	2
17	WHAT	30	21	9	21	0	9
18	HAS	29	27	2	27	0	2
19	WERE	29	28	1	28	0	1
20	OF	25	9	16	9	0	16
21	THIS	22	6	16	6	0	16
22	AS	21	3	18	3	0	18
23	WHO	21	20	1	20	0	1
24	WILL	19	18	1	18	0	1
25	BUT	17	15	2	15	0	2
26	HE	17	15	2	15	0	2
27	THEM	16	2	14	2	0	14
28	WE	15	12	3	12	0	3
29	WITHOUT	14	13	1	13	0	1
30	COULD	13	12	1	12	0	1
31	MORE	13	0	13	0	0	13
32	NEVER	13	13	0	13	0	0
33	WITH	13	0	13	0	0	13
34	HAVE	12	12	0	12	0	0
35	THEIR	12	0	12	0	0	12
36	THERE	12	4	8	4	0	8
37	WHICH	11	8	3	8	0	3
38	FOR	10	2	8	2	0	8
39	FROM	10	1	9	1	0	9
40	ON	8	0	8	0	0	8
41	WHEN	8	4	4	4	0	4
42	HAD	7	7	0	7	0	0
43	SHE	7	6	1	6	0	1
44	SOMETHING	7	2	5	2	0	5
45	WOULD	7	7	0	7	0	0
46	ABOUT	6	0	6	0	0	6
47	AT	6	0	6	0	0	6
48	OUT	6	1	5	1	0	5
49	PEOPLE	6	4	2	4	0	2
50	THAN	6	5	1	5	0	1
51	AN	5	0	5	0	0	5
52	ANY	5	0	5	0	0	5
53	BECAUSE	5	0	5	0	0	5
54	EITHER	5	4	1	4	0	1
55	GOOD	5	1	4	1	0	4
56	HER	5	1	4	1	0	4
57	I	5	3	2	3	0	2
58	LANGUAGE	5	2	3	2	0	3

59	NO	5	1	4	1	0	4
60	NOBODY	5	5	0	5	0	0
61	ONLY	5	2	3	2	0	3

Agrupamentos lexicais com o advérbio *actually*

N	cluster	Freq.
1	it is actually	17
2	may actually be	16
3	is actually a	15
4	is actually the	11
5	can actually be	10
6	is not actually	10
7	in terms of	8
8	what is actually	8
9	what they actually	8
10	do not actually	7
11	did not actually	6
12	it actually was	6
13	it was actually	6
14	than they actually	6
15	there is no	6
16	they are actually	6
17	this is actually	6
18	was not actually	6
19	whether it is	6
20	which is actually	6
21	has actually been	5
22	in order to	5
23	that they actually	5
24	this was actually	5
25	was actually a	5
26	was actually the	5
27	what actually happened	5
28	what actually happens	5
29	will actually be	5
30	actually able to	4
31	actually has a	4
32	actually part of	4
33	actually use them	4
34	and what they	4
35	does not actually	4
36	due to the	4
37	either actually or	4
38	is actually an	4
39	it actually is	4
40	may not actually	4



41	nobody actually knows	4
42	not actually exist	4
43	not actually the	4
44	rather than what	4
45	that it actually	4
46	that was actually	4
47	the problem of	4
48	they actually did	4
49	they actually do	4
50	they have actually	4
51	we actually have	4
52	what he actually	4
53	whether or not	4
54	who is actually	4
55	actually changed the	3
56	actually due to	3
57	actually eat more	3
58	actually happened in	3
59	actually happens in	3
60	actually have a	3
61	actually is and	3
62	actually know about	3
63	actually knows what	3
64	actually led to	3
65	actually or inferentially	3
66	actually result in	3
67	actually the case	3
68	and do not	3
69	are actually something	3
70	are actually the	3
71	are actually very	3
72	are not actually	3
73	as it actually	3
74	as well as	3
75	but do not	3
76	can actually see	3
77	could actually be	3
78	eat more than	3
79	have not actually	3
80	in this case	3
81	is actually being	3
82	is actually not	3
83	is actually there	3
84	knows what dark	3
85	metabolic rate and	3
86	more than they	3
87	more than those	3
88	not actually a	3

89	not actually eat	3
90	only # of	3
91	out of the	3
92	part of the	3
93	question of whether	3
94	rate and do	3
95	than those of	3
96	than what is	3
97	that he actually	3
98	that we actually	3
99	the fact that	3
100	the majority of	3
101	there actually is	3
102	there are actually	3
103	they can actually	3
104	they were actually	3
105	were not actually	3
106	what dark matter	3
107	what has actually	3
108	would actually be	3

Colocados do advérbio *consequently*

N	WORD	TOTAL	LEFT	RIGHT	L1	*	R1
1	CONSEQUENTLY	807	0	0	0	807	0
2	AND	210	206	4	206	0	4
3	IS	62	17	45	17	0	45
4	THE	51	0	51	0	0	51
5	TO	26	1	25	1	0	25
6	A	20	0	20	0	0	20
7	WAS	17	10	7	10	0	7
8	IT	15	5	10	5	0	10
9	WOULD	15	7	8	7	0	8
10	ARE	14	2	12	2	0	12
11	CAN	14	2	12	2	0	12
12	MAY	11	5	6	5	0	6
13	OF	11	0	11	0	0	11
14	WHICH	11	10	1	10	0	1
15	THAT	10	3	7	3	0	7
16	AS	9	2	7	2	0	7
17	IN	9	0	9	0	0	9
18	THIS	9	6	3	6	0	3
19	HAS	8	2	6	2	0	6
20	WERE	8	3	5	3	0	5
21	HAVE	7	0	7	0	0	7
22	WILL	7	0	7	0	0	7
23	ALL	6	3	3	3	0	3

24	RESULTS	6	1	5	1	0	5
25	THEM	6	3	3	3	0	3
26	STATE	5	3	2	3	0	2
27	TIME	5	5	0	5	0	0

Agrupamentos lexicais com o advérbio *consequently*

N	cluster	Freq.
1	and consequently the	27
2	consequently it is	7
3	and consequently a	6
4	and consequently it	5
5	and is consequently	5
6	there is a	5
7	and consequently to	4
8	consequently lead to	4
9	it is not	4
10	the world and	4
11	and consequently are	3
12	and consequently he	3
13	and consequently more	3
14	and consequently only	3
15	are able to	3
16	consequently leads to	3
17	in order to	3
18	it can be	3
19	it is clear	3
20	it would be	3
21	may consequently be	3
22	the principles of	3

Colocados do advérbio *especially*

N	WORD	TOTAL	LEFT	RIGHT	L1	*	R1
1	ESPECIALLY	1698	1	0	1	1697	0
2	THE	293	1	292	1	0	292
3	AND	133	94	39	94	0	39
4	IS	121	118	3	118	0	3
5	IN	81	10	71	10	0	71
6	TO	54	1	53	1	0	53
7	FOR	43	1	42	1	0	42
8	IT	37	5	32	5	0	32
9	ARE	29	23	6	23	0	6
10	OF	28	1	27	1	0	27
11	THAT	28	9	19	9	0	19
12	A	26	0	26	0	0	26

13	THEY	26	2	24	2	0	24
14	AS	21	4	17	4	0	17
15	WITH	20	1	19	1	0	19
16	AN	19	9	10	9	0	10
17	BE	18	17	1	17	0	1
18	WAS	18	16	2	16	0	2
19	THIS	16	1	15	1	0	15
20	THOSE	16	0	16	0	0	16
21	WERE	15	13	2	13	0	2
22	WOMEN	14	6	8	6	0	8
23	COUNTRIES	12	8	4	8	0	4
24	FROM	12	1	11	1	0	11
25	PEOPLE	12	9	3	9	0	3
26	TERMS	12	2	10	2	0	10
27	WHICH	12	2	10	2	0	10
28	BUT	11	10	1	10	0	1
29	THEIR	11	0	11	0	0	11
30	BECAUSE	10	0	10	0	0	10
31	AT	9	0	9	0	0	9
32	HE	9	0	9	0	0	9
33	ON	9	1	8	1	0	8
34	WHEN	9	0	9	0	0	9
35	WHO	9	1	8	1	0	8
36	EUROPE	8	5	3	5	0	3
37	MEN	8	5	3	5	0	3
38	RELATION	8	0	8	0	0	8
39	THESE	8	1	7	1	0	7
40	TIME	8	7	1	7	0	1
41	CASE	7	2	5	2	0	5
42	HIGH	7	3	4	3	0	4
43	HISTORY	7	6	1	6	0	1
44	INFORMATION	7	6	1	6	0	1
45	ITS	7	0	7	0	0	7
46	MORE	7	4	3	4	0	3
47	NOT	7	6	1	6	0	1
48	WE	7	0	7	0	0	7
49	WELL	7	6	1	6	0	1
50	WORLD	7	7	0	7	0	0
51	AREAS	6	1	5	1	0	5
52	COMPANIES	6	3	3	3	0	3
53	DIFFICULT	6	6	0	6	0	0
54	GROUPS	6	6	0	6	0	0
55	HOSPITALITY	6	1	5	1	0	5
56	IMPORTANT	6	6	0	6	0	0
57	MARKET	6	6	0	6	0	0
58	ONES	6	2	4	2	0	4
59	SOME	6	1	5	1	0	5
60	SUCH	6	0	6	0	0	6

61	THEM	6	6	0	6	0	0
62	ABILITY	5	1	4	1	0	4
63	ANALYSIS	5	5	0	5	0	0
64	BEEN	5	4	1	4	0	1
65	BEING	5	5	0	5	0	0
66	DEVELOPING	5	0	5	0	0	5
67	ECONOMY	5	5	0	5	0	0
68	FUTURE	5	5	0	5	0	0
69	HIS	5	0	5	0	0	5
70	I	5	2	3	2	0	3
71	IF	5	0	5	0	0	5
72	IMPORTANCE	5	5	0	5	0	0
73	INSTITUTIONS	5	4	1	4	0	1
74	ISSUE	5	2	3	2	0	3
75	LIFE	5	5	0	5	0	0
76	LOW	5	3	2	3	0	2
77	POOR	5	2	3	2	0	3
78	PROBLEMS	5	5	0	5	0	0
79	SEEMS	5	4	1	4	0	1
80	SYSTEM	5	5	0	5	0	0
81	WORDS	5	4	1	4	0	1

#### Agrupamentos lexicais do advérbio *especially*

N	cluster	Freq.
1	especially in the	88
2	this is especially	62
3	is especially true	28
4	is especially important	25
5	especially when the	23
6	especially as the	16
7	and especially the	15
8	especially for the	15
9	especially if they	14
10	especially with the	14
11	it is especially	12
12	especially after the	11
13	especially due to	11
14	especially if the	11
15	especially important for	11
16	especially as it	10
17	especially in terms	10
18	in terms of	10
19	due to the	9
20	especially since the	9
21	especially true for	9
22	especially during the	8

23	especially in relation	8
24	especially true in	8
25	especially when it	8
26	in relation to	8
27	especially at the	7
28	especially for those	7
29	especially from the	7
30	especially important in	7
31	especially those from	7
32	especially to the	7
33	especially within the	7
34	is especially useful	7
35	which is especially	7
36	especially considering that	6
37	especially evident in	6
38	especially for a	6
39	especially given the	6
40	especially if it	6
41	especially in a	6
42	especially in this	6
43	especially on the	6
44	especially that of	6
45	especially vulnerable to	6
46	in light of	6
47	is an especially	6
48	is especially so	6
49	the use of	6
50	this was especially	6
51	a lot of	5
52	and is especially	5
53	as well as	5
54	especially important to	5
55	especially so for	5
56	especially the case	5
57	especially useful in	5
58	if they are	5
59	important for the	5
60	in developing countries	5
61	in the case	5
62	light of the	5
63	to be especially	5
64	with regard to	5
65	a negative impact	4
66	and especially in	4
67	because of the	4
68	developing countries since	4
69	especially among the	4
70	especially as a	4

71	especially by the	4
72	especially in hospitality	4
73	especially in such	4
74	especially the last	4
75	especially those that	4
76	especially those who	4
77	especially those with	4
78	especially through the	4
79	especially when we	4
80	especially with regard	4
81	evident in the	4
82	has a negative	4
83	impact especially in	4
84	in the north	4
85	in the uk	4
86	is especially evident	4
87	is especially shown	4
88	it is a	4
89	members of the	4
90	negative impact especially	4
91	the case for	4
92	the case of	4
93	the end of	4
94	the fact that	4
95	those from the	4
96	those who are	4
97	and especially with	3
98	are especially important	3
99	as a whole	3
100	as being especially	3
101	as it is	3
102	as it was	3
103	between the two	3
104	can be especially	3
105	countries since they	3
106	end of the	3
107	especially apparent in	3
108	especially as he	3
109	especially as this	3
110	especially because of	3
111	especially concerning the	3
112	especially hard for	3
113	especially important when	3
114	especially in an	3
115	especially in cases	3
116	especially in developing	3
117	especially in light	3
118	especially in rural	3

119	especially in those	3
120	especially not to	3
121	especially so in	3
122	especially taking into	3
123	especially the unskilled	3
124	especially those in	3
125	especially those of	3
126	especially true of	3
127	especially when considering	3
128	globalisation has a	3
129	has been especially	3
130	in especially the	3
131	in rural areas	3
132	in the context	3
133	in the future	3
134	is especially apparent	3
135	is especially relevant	3
136	is especially the	3
137	it can be	3
138	it is an	3
139	it was not	3
140	need to be	3
141	part of the	3
142	relation to the	3
143	such as the	3
144	taking into account	3
145	that it is	3
146	the area of	3
147	the context of	3
148	the european union	3
149	the hospitality industry	3
150	the nature of	3
151	this is an	3
152	true for the	3
153	with regards to	3
154	would be especially	3