

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS**

REJANE WAGNER PROTZNER SILERO

**OS QUANTIFICADORES *A FEW* E *FEW*: QUESTÕES DE
INTERLÍNGUA E PROSÓDIA SEMÂNTICA EM *CORPUS* DE
APRENDIZES**

**BELO HORIZONTE
FACULDADE DE LETRAS DA UFMG
2014**

Rejane Wagner Protzner Silero

OS QUANTIFICADORES *A FEW* E *FEW*: questões de interlíngua e prosódia semântica em *corpus* de aprendizes

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

Área de Concentração: Linguística Aplicada
Linha de Pesquisa: Ensino/Aprendizagem de Línguas Estrangeiras

Orientadora: Profa. Dra. Deise Prina Dutra

Belo Horizonte
Faculdade de Letras da UFMG
2014

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

S582q Silero, Rejane Wagner Protzner.
Os quantificadores *A Few* e *Few* [manuscrito] : questões de Interlíngua e prosódia semântica em corpus de aprendizes / Rejane Wagner Protzner Silero. – 2014.
107 f., enc. : il., grafis (p&b)
Orientadora: Daise Prina Dutra.
Área de concentração: Linguística Aplicada.
Linha de Pesquisa: Ensino/Aprendizagem de Línguas Estrangeiras.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.
Bibliografia: f. 72-79.
Anexos: 80-87.

1. Interlíngua (Aprendizagem de línguas) – Teses. 2. Língua inglesa – Estudo e ensino – Falantes estrangeiros – Teses. 3. Língua inglesa – Semântica – Teses. 4. Aquisição da segunda língua – Teses. I. Dutra, Daise Prina. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD: 420.7




FOLHA DE APROVAÇÃO

OS QUANTIFICADORES A FEW E FEW: questões de interlíngua e prosódia semântica em corpus de aprendizes

REJANE WAGNER PROTZNER SILERO

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA APLICADA, linha de pesquisa Ensino/Aprendizagem de Línguas Estrangeiras.

Aprovada em 06 de junho de 2014, pela banca constituída pelos membros:


Prof(a). Deise Prina Dutra - Orientador
UFMG


Prof(a). Tania Maria Granja Shepherd
UERJ


Prof(a). Heliana Ribeiro de Mello
UFMG

Belo Horizonte, 6 de junho de 2014.

Ao meu avô (*in memoriam*), cuja presença em minha vida foi pequena em número de anos, mas suficientemente grande para impactar a minha visão de mundo e falar alto em meu coração, pelo resto da minha existência, sobre a serenidade da sabedoria inteligente e da inteligência sábia. À minha mãe, pelo apoio incondicional, por ser esteio manifesto em firme cuidado e amor, e meu incomparável exemplo de força e fé.

AGRADECIMENTO

Em primeiro lugar a Deus, porque pude experimentar de uma maneira amorosa e singularmente impressionante a força que Ele imprimiu em meu ser, a qual só poderia ter vindo dEle, e que foi a única responsável pelo meu mover e pelos passos que eu consegui dar ao longo dessa jornada. Expresso, portanto, reconhecendo em atitude de gratidão e com humildade, a ação propulsora de Cristo Jesus em minha vida, que me levantou em momentos de fraqueza e me fez chegar até o final. Se o Verbo de Deus tivesse se calado, eu nada teria feito. Louvado seja o Senhor!

À minha mãe, cujo apoio e firmeza me fizeram permanecer de pé e me auxiliaram incessantemente nesse tempo. Não fosse a sua paciência, o seu auxílio e os seus ouvidos sempre disponíveis, eu não teria me estruturado emocionalmente durante esse processo.

À professora Deise Prina Dutra, que foi mentora inestimável para mim, desde os tempos de graduação. Suas aulas me fizeram conhecer mais profundamente as teorias da linguística aplicada ao ensino e os métodos provenientes da linguística de *corpus*. Tais conhecimentos influenciaram de forma substancial a minha visão sobre aquisição e descrição de línguas, e continuamente me fazem repensar o lugar ‘sala de aula’, e as posições ‘professor’ e ‘aluno’. Agradeço também pelas suas valiosas instruções e comentários pertinentes durante a orientação que eficientemente me prestou.

À inspiradora professora Heliana Mello, que como uma fonte inesgotável de saberes, emana-os naturalmente em meio à humilde sobriedade. Ministradas com requintado grau de competência, suas aulas sempre refletiram todo o ressoar da palavra ‘educação’, pois integralmente visavam ao crescimento dos alunos nas mais diversas esferas. Posso afirmar com segurança que, além de adquirir conhecimentos acadêmicos, também me deparei com momentos em que me vi crescendo como pessoa e ativando positivamente o meu senso crítico.

Ao CCAA Santo Agostinho, pois embora não tenha tido participação direta em minha formação acadêmica, foi o meu ‘trabalho-casa’, onde exerci minha profissão ao longo desses anos, em contato com diversos alunos preciosos, e onde eu construí verdadeiras amizades. No CCAA eu percebi que é possível trabalhar harmoniosamente em grupo, desfrutar de discussões profissionais e de momentos de descontração.

À Escola Lúcia Casasanta, que desenhou o mais belo bordado em minha infância, ensinou-me a criar gosto pela leitura e pelos estudos, e cujas mais belas histórias foram plantadas em minha memória para nunca mais deixarem de dar fruto.

Ao Edson, pela tolerância, palavras de incentivo, e por ter sempre me lembrado de que no final daria tudo certo.

Ao meu pai e à tia Sossô, pelo carinho com minha vida e perseverante oração.

Aos colegas de faculdade e companheiros de profissão, pelas conversas e trocas de experiências, pelos momentos de estudo e também pelos lanches na cantina.

RESUMO

A pesquisa aqui descrita foi conduzida à luz do que já foi realizado por pesquisadores (GRANGER; TRIBBLE, 1998) que se interessam pela produção escrita de aprendizes de inglês. O objetivo geral foi o de verificar e descrever a interlíngua dos aprendizes brasileiros de inglês em nível intermediário-avançado, especificamente quanto ao uso que eles fazem das expressões quantificadoras *a few* e *few* em suas produções escritas. Utilizando os dados do Br-ICLE, subcorpus brasileiro do *International Corpus of Learner English* (ICLE), foram feitas análises que permitiram vislumbrar esse uso não somente em nível gramatical como também em nível de prosódia semântica. Similarmente, foram analisados dados provenientes do corpus de nativos equivalente ao ICLE, o *Louvain Corpus of Native English Essays* (LOCNESS) para fins comparativos. Recorreu-se, também, ao COCA acadêmico como outro *corpus* de referência para a descrição do uso que os nativos fazem dos quantificadores em estudo. Os dados do Br-ICLE e do LOCNESS foram tratados qualitativamente por meio de análises manuais do contexto expandido das linhas de concordância com o nódulo *few*. Foram identificadas e comparadas colocações que se mostraram adequadas, em consonância com o uso dos nativos, e que se mostraram inadequadas, apresentando desvios que foram sistematizados e comentados. O tipo de inadequação que recebeu tratamento especial foi a concernente à manifestação da prosódia semântica. Foi detectado que no *corpus* dos nativos houve um considerável número de ocorrências da colocação *very few* para expressar prosódia semântica negativa, fato não verificado no Br-ICLE, atestando subuso dessa colocação pelos aprendizes. Quantitativamente, os resultados obtidos nesse trabalho foram descritos em porcentagens e o subuso identificado recebeu validação por meio de teste estatístico. A fim de averiguar a extensão da influência da língua materna no uso que os aprendizes fizeram dos quantificadores, especialmente mediante a interface com a prosódia semântica, foi compilado o Corpus de Redações Acadêmicas em Português (Co-RAP), de tamanho semelhante ao Br-ICLE e ao LOCNESS, e também foi utilizado o Corpus do Português (*online*) para demais consultas. A conclusão foi a de que os falantes brasileiros, em suas produções escritas, não fazem um vasto uso de ‘muito poucos(as)’, que seria a colocação correspondente a *very few*, e isso pode interferir em suas escolhas lexicais na língua-alvo. Por fim, foi sugerida, aos moldes dos princípios da Aprendizagem Movidada por Dados (JOHNS, 1994), uma atividade que estimula a identificação de padrões léxico-gramaticais para os quantificadores investigados, e também foram sugeridos caminhos para próximas pesquisas.

Palavras-chave: Corpus de aprendizes. Interlíngua. Prosódia semântica. Léxico-gramática.

ABSTRACT

The investigation described here was carried out under the light of what has been done by researchers (GRANGER; TRIBBLE, 1998) who are interested in learner English writing. The main goal was to verify and describe the upper-intermediate Brazilian learners' interlanguage, specifically concerning the use of the quantifying expressions **a few** and **few**. Data from Br-ICLE, the Brazilian subcorpus of the International Corpus of Learner English (ICLE) was analyzed so as to enable description both on the grammatical level and on the level of semantic prosody. Similarly, data from the Louvain Corpus of Native English Essays (LOCNESS), which was built to be the native equivalent for ICLE, was analyzed and used for comparisons. It was also necessary to resort to the academic section of COCA as another reference corpus in order to describe the way native speakers use the quantifiers being studied. The data from both Br-ICLE and LOCNESS were treated qualitatively, through manual analyses of the expanded context of concordance lines whose node word was **few**. The following sorts of collocations were identified and compared: the collocations that appeared adequate, that is, compatible with native speakers' use; and the ones that appeared inadequate, that is, showing deviations that were further systematized and discussed. The type of inadequacy that received special treatment was the one related to the surfacing of semantic prosody. In the native speakers' corpus, the collocation **very few** was often used to express negative semantic prosody; however, this was not attested in the Br-ICLE, which leads to the understanding that such collocation was underused by the learners in similar contexts. Quantitatively, the results of this research were described in percentages, and the diagnosed underuse was validated through statistical testing. In order to gauge the extent to which the mother tongue influences the expression of semantic prosody in the target language, another corpus, whose size is on a par with Br-ICLE and LOCNESS, was compiled: the Co-RAP, which stands for Corpus of Academic Essays in Portuguese (Brazilian Portuguese). Also, the online platform of Corpus do Português was used for further necessary searches. The conclusion drawn was that, in their writing, Brazilians do not seem to widely use *muito poucos(as)*, which is the correspondent collocation in Portuguese to **very few**, and this may influence their lexical choices when writing in English. Lastly, following the tenets of Data-Driven Learning (JOHNS, 1994), a classroom activity that encourages the recognition of lexico-grammatical patterns was suggested, and promising pathways to future research were mentioned as well.

KeyWords: Learner corpus. Interlanguage. Semantic prosody. Lexico-grammar.

LISTA DE FIGURAS

Figura 1 – O uso de <i>corpora</i> no ensino e aprendizagem de segunda língua (L2) – tradução nossa	04
Figura 2 – Quadrado de oposições de quantificadores proposto por Aristóteles	12
Figura 3 – Composição dos dados da parte do LOCNESS e do Br-ICLE para a aplicação do Teste Exato de Fisher	48
Figura 4 – Resultado do Teste Exato de Fisher para os dados fornecidos	49
Figura 5 – Obtida pela busca do item lexical <i>few</i> no dicionário online Longman	63
Figura 6 – About the Cambridge Learner Corpus	80
Figura 7 – About the Longman Learner Corpus.....	81
Figura 8 – Longman Student Grammar (Moderate or small quantity).....	82
Figura 9 – Practical English Usage (Use of <i>a</i>)	83

LISTA DE GRÁFICOS

Gráfico 1 – Proporção das ocorrências de <i>past few</i> e <i>last few</i> na parte do LOCNESS e no Br-ICLE.....	42
Gráfico 2 – Ocorrências de <i>very few</i> na parte do LOCNESS e no Br-ICLE	43
Gráfico 3 – Parte inicial da curva de Zipf para a distribuição de palavras no romance <i>Moby Dick</i>	85

LISTA DE QUADROS

Quadro 1 – Porcentagem de ocorrências com <i>a few</i> e <i>few</i> no Br-ICLE	37
Quadro 2 – Porcentagens no Br-ICLE dos colocados à direita de <i>a few</i> e de <i>few</i> , separados por domínio semântico.....	38
Quadro 3 – Porcentagens na parte do LOCNESS dos colocados à direita de <i>a few</i> e de <i>few</i> , separados por domínio semântico.....	39
Quadro 4 – 12 primeiros colocados à direita de <i>few</i> na seção acadêmica do COCA.....	39
Quadro 5 – 12 primeiros colocados à direita de <i>a few</i> na seção acadêmica do COCA.....	40
Quadro 6 – 12 primeiros colocados à esquerda de <i>few</i> na seção acadêmica do COCA, destacando-se <i>very</i> , <i>past</i> e <i>last</i>	41
Quadro 7 – Ocorrências de <i>past few</i> no Br-ICLE.....	41
Quadro 8 – Ocorrências de <i>very few</i> no Br-ICLE	41
Quadro 9 – Ocorrências de <i>very few</i> no Br-ICLE, com os colocados à direita em destaque ...	42
Quadro 10 – Ocorrências de <i>very few</i> na parte do LOCNESS, com os colocados à direita em destaque	42
Quadro 11 – 12 primeiros colocados à direita de <i>very few</i> na seção acadêmica do COCA	43
Quadro 12 – Grupo de linhas de concordância provenientes da parte do LOCNESS que apresentam prosódia semântica negativa.....	45
Quadro 13 – Grupo de linhas de concordância provenientes do Br-ICLE que apresentam prosódia semântica negativa.....	47
Quadro 14 – 12 primeiros colocados à esquerda de <i>few</i> no COCA	50
Quadro 15 – 12 primeiros colocados à esquerda de poucos no Corpus do Português.....	51
Quadro 16 – 12 primeiros colocados à esquerda de poucas no Corpus do Português.....	51
Quadro 17 – Ocorrências aleatórias de pouquíssimos no Corpus de Português	52
Quadro 18 – Ocorrências aleatórias de pouquíssimas no Corpus de Português	52
Quadro 19 – Imagem gerada pelo programa AntConc da única ocorrência de pouquíssimas no Co-RAP	53
Quadro 20 – Porcentagens das linhas de concordância adequadas e inadequadas do Br-ICLE	54
Quadro 21 – 1º quadro com 16 linhas de concordância com <i>a few</i> e <i>few</i> no Br-ICLE cuja avaliação léxico-gramatical é adequada	55

Quadro 22 – Itens lexicais do contexto expandido (referentes às linhas de concordância do quadro 21) que contribuíram para a emergência da prosódia semântica negativa no Br-ICLE	56
Quadro 23 – 2º quadro com 16 linhas de concordância com <i>a few</i> e <i>few</i> no Br-ICLE cuja avaliação léxico-gramatical é adequada.....	56
Quadro 24 – Itens lexicais do contexto expandido (referentes às linhas de concordância do quadro 23) que contribuíram para a emergência da prosódia semântica negativa no Br-ICLE	57
Quadro 25 – 3º quadro com 16 linhas de concordância com <i>a few</i> e <i>few</i> no Br-ICLE cuja avaliação léxico-gramatical é adequada.....	57
Quadro 26 –Itens lexicais do contexto expandido (referentes às linhas de concordância do quadro 25) que contribuíram para a emergência da prosódia semântica negativa no Br-ICLE	558
Quadro 27 – 4º quadro com 23 linhas de concordância com <i>a few</i> e <i>few</i> no Br-ICLE cuja avaliação léxico-gramatical é inadequada.....	59
Quadro 28 – Divisão em categorias das inadequações cometidas pelos participantes do Br-ICLE.....	60
Quadro 29 – Linhas de concordância com o padrão <i>to just a few</i> , aleatoriamente retiradas do COCA acadêmico	61
Quadro 30 – As linhas de concordância do Br-ICLE cujo nóculo é <i>a few</i> e os colocados à direita possuem sentido temporal.....	86
Quadro 31 – As linhas de concordância do Br-ICLE cujo nóculo é <i>few</i> e os colocados à direita possuem sentido temporal.....	87

SUMÁRIO

1	INTRODUÇÃO	01
1.1	Historicizando as motivações para este trabalho	01
1.2	Os objetivos da pesquisa	02
1.3	As perguntas de pesquisa.....	03
1.4	A justificativa.....	03
1.5	A organização da dissertação	06
2	REVISÃO DE LITERATURA	07
2.1	Fundamentos-base.....	07
2.2	Panorama histórico e gramatical sobre os quantificadores.....	11
2.2.1	Um breve histórico	11
2.2.2	Quantificadores Generalizados.....	12
2.2.3	Monotonicidade dos quantificadores.....	14
2.2.4	Os quantificadores em gramáticas.....	15
<u>2.2.4.1</u>	<u>As gramáticas do inglês</u>	<u>16</u>
<u>2.2.4.2</u>	<u>As gramáticas do português.....</u>	<u>17</u>
2.3	Panorama sobre prosódia semântica	18
2.3.1	Apanhado de conceitos de prosódia semântica	18
2.3.2	Problematização dos conceitos.....	20
2.3.3	Estudos relevantes de prosódia semântica e o entrelaçar com o ensino.....	21
2.4	Noções sobre erro e conscientização linguística.....	22
2.4.1	Conceitos de erro.....	23
2.4.2	O papel da conscientização linguística.....	25
2.5	Panorama de trabalhos com <i>corpora</i> de aprendizes e sobre quantificadores em <i>corpora</i> e <i>corpora</i> de aprendizes	27
2.5.1	Um micro retrato dos estudos com <i>corpora</i> de aprendizes	27
2.5.2	Trabalhos sobre quantificadores em <i>corpora</i>	28
3	METODOLOGIA	30
3.1	Dos <i>corpora</i> utilizados.....	30
3.2	Das análises realizadas.....	34
3.2.1	O apanhado geral.....	34
3.2.2	O apanhado com nuances de prosódia semântica	35
3.2.3	O detalhamento dos dados dos aprendizes	36

4	RESULTADOS	37
4.1	Análises: o apanhado geral	37
4.1.1	Os colocados à direita	37
4.1.2	Os colocados à esquerda	40
4.1.3	Os padrões léxico-gramaticais relacionados a tempo	44
4.2	Análises: a interface com a prosódia semântica.....	45
4.2.1	O primeiro argumento.....	48
4.2.2	O segundo argumento	48
4.2.3	O terceiro argumento	50
4.3	Análises: o detalhamento dos dados dos aprendizes.....	54
4.3.1	Os três quadros com as adequações.....	55
4.3.2	O quadro com as inadequações.....	59
4.4	Implicações dos resultados.....	63
4.5	Activity - Patterns & Quantifiers	66
5	CONCLUSÃO.....	69
	REFERÊNCIAS.....	72
	ANEXOS A – ABOUT THE CAMBRIDGE LEARNER CORPUS.....	80
	ANEXO B – ABOUT THE LONGMAN LEARNER CORPUS	81
	ANEXO C – LONGMAN STUDENT GRAMMAR (MODERATE OR SMALL	
	QUANTITY).....	82
	ANEXO D – PRACTICAL ENGLISH USAGE (USE OF A).....	83
	ANEXO E – LP (FORMULÁRIO PREENCHIDO PELOS PARTICIPANTES DO BR-	
	ICLE).....	84
	ANEXO F – PARTE INICIAL DA CURVA DE ZIPF PARA A DISTRIBUIÇÃO DE	
	PALAVRAS NO ROMANCE <i>MOBY DICK</i>	85
	ANEXO G – LINHAS DE CONCORDÂNCIA DO BR-ICLE CUJO NÓDULO É A	
	<i>FEW</i> E OS COLOCADOS À DIREITA POSSUEM SENTIDO TEMPORAL	86
	ANEXO H – LINHAS DE CONCORDÂNCIA DO BR-ICLE CUJO NÓDULO É <i>FEW</i>	
	E OS COLOCADOS À DIREITA POSSUEM SENTIDO TEMPORAL	87

1 INTRODUÇÃO

1.1 Historicizando as motivações para este trabalho

Quando ainda aluna na graduação, e participando do grupo de pesquisas durante o período de Iniciação Científica, eu me deparei com o mundo novo da Linguística de Corpus, doravante LC, cujos primeiros conceitos foram apresentados a mim pela Professora Doutora Deise Prina Dutra. Desde então, a minha fascinação por este universo foi despertada e, como professora de inglês, senti bater mais forte a necessidade de entender melhor como o ensino poderia ser aprimorado com o auxílio das ferramentas da LC.

Em breves palavras, fui percebendo como as investigações com *corpora* abriam caminhos para novas descobertas envolvendo a descrição das línguas e o estudo de mudanças linguísticas em diferentes níveis. O número de pesquisas nessa área foi crescendo¹, bem como o número de aparatos tecnológicos que dão suporte às metodologias propostas pela LC. Atualmente, há diversos *corpora* que podem ser acessados gratuitamente (como o *Brown Corpus* e o *Corpus of Contemporary American English (COCA)*², e também programas de tratamento de dados – concordanceadores, por exemplo – que podem ser baixados gratuitamente, tais como o TextStat 3.0 e o AntConc 3.0. Com isso, as investigações com *corpora* se tornaram mais acessíveis não somente a pesquisadores, mas também a professores e aos próprios alunos.

Em recordações da minha própria experiência, nos anos de 2008 e 2009, quando ainda cursava disciplinas na graduação, aproximadamente na metade do Curso de Letras, fui sendo apresentada a alguns *corpora online*, como o *Compleat Lexical Tutor* (<http://www.lextutor.ca/>) e o COCA (<http://www.americancorpus.org>). Assim, fui encorajada a utilizá-los como ferramentas de consulta que auxiliariam a compor as discussões e as atividades propostas nas disciplinas, como a de *Corpora* e Aprendizagem, ministrada pela Professora Doutora Deise Dutra, a de História do Inglês (para qual também utilizamos o *Corpus of Historical American English (COHA)* – (<http://corpus.byu.edu/coha/>), ministrada

¹ Alves e Tagnin (2012, p. 17) elencam cinco áreas de pesquisa que utilizam suporte teórico-metodológico da LC, cuja pluralidade de aplicações como ferramenta para pesquisa linguística é evidenciada. As cinco áreas identificadas pelos autores são a tradução, a descrição da linguagem, o ensino, a terminologia e o Processamento de Linguagem Natural (PNL). Pode-se atestar o profícuo crescimento da LC também pela reunião de vários pesquisadores da área em seminários e congressos internacionais e nacionais, respectivamente exemplificados pelo UCREL Corpus Research Seminar (<http://ucrel.lancs.ac.uk/crs/>) e pelo ELC (Encontro de Linguística de Corpus), este último já com mais de dez edições no país.

² Eis os links para o Brown Corpus e para o COCA, respectivamente: Disponível em: <http://www.helsinki.fi/varieng/CoRD/corpora/BROWN/>. e Disponível em: <http://www.americancorpus.org>.

pela Professora Doutora Heliana Ribeiro de Mello, e a de Sintaxe, ministrada pelo Professor Doutor Ricardo Augusto de Souza.

Direcionando agora a atenção para a habilidade da escrita, ao longo de quase dez anos dando aulas de inglês em curso livre e corrigindo as redações dos alunos, sempre me deparei com inadequações complicadas de serem resolvidas, pois são de difícil explicação para o professor e difícil captação pelos alunos. Desse modo, dentre as grandes motivações deste trabalho está a possibilidade de utilizar as ferramentas da LC para elucidar questões problemáticas de escrita. Tais focos de problema podem ser reduzidos caso haja, em materiais didáticos e na instrução fornecida pelo professor, um melhor detalhamento dessas questões, ilustradas com exemplos autênticos provenientes de *corpora* e discutidas à luz de uma perspectiva léxico-gramatical da linguagem. Biber *et al.* (1999, p. 13) explicam que a maioria das estruturas sintáticas tendem a ter um conjunto de palavras ou frases associadas a elas; logo, a visão da léxico-gramaticalidade, encorajada nesse trabalho, promove a identificação de padrões linguísticos recorrentes para expressar um determinado sentido.

1.2 Os objetivos da pesquisa

Segundo Harmer (1998), além de funcionar como um reforço das demais habilidades, a escrita promove o desenvolvimento linguístico do aluno, ajudando-o a solidificar o uso de novas estruturas e vocabulário. Além disso, a escrita, assim como a fala, é reflexo direto da expressão e do conhecimento do aprendiz, sendo uma amostra de sua própria produção. Entendendo que a interlíngua³ se manifesta no decorrer do aprendizado dos alunos, e observando algumas inadequações cometidas quanto ao uso das expressões quantificadoras *a few* e *few*, o objetivo geral desse trabalho é contribuir com uma descrição dessa interlíngua, especificamente de aprendizes brasileiros da língua inglesa cujo nível é entre intermediário e avançado, levando em conta a maneira como eles usam essas expressões. As noções de **sobreuso** e **subuso**, já consolidadas na literatura que lida com pesquisas de *corpora* de aprendizes, também são invocadas nessa descrição, e serão devidamente explicadas no capítulo 2.

³ Selinker (1969, apud SONG, 2012) define o termo interlíngua como as gramáticas intermediárias construídas pelos aprendizes de segunda língua na sua caminhada rumo à língua-alvo. Assim, é criado um sistema linguístico separado, que é resultante da tentativa de produção, por parte dos aprendizes, da norma que preside a língua-alvo.

Os objetivos específicos são:

- (i) descrever como os aprendizes brasileiros utilizam os quantificadores *a few* e *few* em termos gramaticais e em termos de **prosódia semântica**⁴;
- (ii) descrever, similarmente, como os falantes nativos utilizam esses mesmos quantificadores;
- (iii) verificar os tipos de inadequações cometidas pelos aprendizes;
- (iv) propor um modelo de atividade que pode ser aplicada como uma intervenção pedagógica direta que incentive a percepção da noção de léxico-gramaticalidade, isto é, a identificação de padrões estruturais que estão atrelados a um determinado sentido/uso pragmático dos quantificadores.

1.3 As perguntas de pesquisa

Após feitas análises manuais das linhas de concordância dos *corpora* utilizados como fontes de dados para esse trabalho, foram formuladas as seguintes perguntas:

- a) Há semelhanças entre o uso que os aprendizes brasileiros fazem das expressões quantificadoras *a few* e *few* e o uso que os nativos do inglês fazem das mesmas? Existe **sobreuso** ou **subuso** de alguma colocação⁵?
- b) Há emergência clara de **prosódia semântica** quando do uso de colocações com essas expressões quantificadoras por parte dos nativos? Como isso se dá nos dados provenientes dos aprendizes, considerando a sua interlíngua? Qual a atuação da língua materna nesse tocante?

1.4 A justificativa

Segundo Römer (2011), em se tratando de aplicações de *corpora* no ensino de uma língua-alvo (L2), deve-se incluir ambos o uso de ferramentas de *corpus* (os conjuntos de dados em si e os *softwares* de tratamento desses dados) e o uso de métodos de *corpus*

⁴ Esse conceito será amplamente desenvolvido no capítulo 2.

⁵ Firth (1957, p. 11) cunhou o termo colocação por meio de sua célebre frase “*You shall know a word by the company it keeps*” (Conhece-se uma palavra pelas suas companhias – tradução minha). Outros linguistas, como Sinclair (1991) e Cowie (1994) também definiram o termo. O primeiro reconhece colocação como a ocorrência de uma ou mais palavras dentro de um texto com um espaço curto entre elas; o segundo a concebe como um tipo de combinação de palavras a qual pode ser distinguida de outros tipos de agrupamentos lexicais, como as combinações livres e os idiomas.

(técnicas de análise que são empregadas ao lidar com os dados). Recuperando Leech (1997), para classificar aplicações pedagógicas de *corpus*, isto é, o uso de ferramentas e métodos de *corpus* num contexto de ensino-aprendizagem de línguas, é necessário fazer a distinção entre aplicações diretas e indiretas. Indiretamente, *corpora* podem auxiliar em decisões quanto ao quê ensinar e quando ensinar, influenciando na confecção de materiais didáticos. Diretamente, *corpora* podem ser acessados por aprendizes e professores dentro da sala de aula, auxiliando no processo de ensino em si. O diagrama abaixo, retirado do artigo de Römer (2011, p. 207), representa as ramificações para os tipos de aplicações pedagógicas de *corpus*:



Figura 1 – O uso de *corpora* no ensino e aprendizagem de segunda língua (L2) – tradução minha
Fonte: RÖMER, 2011, p. 207.

Como observado por Barlow (1996, p. 32), “os resultados de uma investigação baseada em *corpus* servem como uma base firme para a descrição linguística e, do lado aplicado, serve como insumo para o aprendizado de língua.” Logo, para este trabalho, que conta com dados provenientes de aprendizes, o que os resultados oferecerão de contribuição estará vinculado não concretamente a ações pedagógicas, mas ao fornecimento de maiores informações para o professor a respeito da descrição de algumas características da interlíngua dos aprendizes no que se refere ao uso dos quantificadores em estudo. Isso pode auxiliar no rumo a ser tomado durante a elaboração de atividades e aplicação das mesmas em sala de aula. Inclusive, uma sugestão de atividade será apresentada e devidamente discutida posteriormente, no capítulo 4, dentro da seção de implicações dos resultados.

Por hora, é válido ressaltar que as pesquisas com *corpora* de aprendizes, seja com aplicações diretas ou indiretas na instrução pedagógica, ocorrem há aproximadamente 20

anos, destacando-se Johns (1994), Granger e Tribble (1998), Seidlhofer (2002) e Nesselhauf (2005). Johns propôs a abordagem chamada **Aprendizado Movido por Dados - Data-Driven Learning** - (1994), cuja aplicação foi incentivada por meio, por exemplo, do fornecimento de linhas de concordância retiradas de *corpora* em atividades que procuravam induzir os aprendizes a descobrir padrões no uso, por exemplo, de *that-clauses*, *such as*, do artigo definido *the* e de *have to*. Granger e Tribble (1998) apontaram para a queda da acuidade gramatical após o advento do ensino mais focado na comunicação, e compartilham com Johns os pressupostos do **Aprendizado Movido por Dados**. Eles reforçaram que tal abordagem não deve ser utilizada como algo maçante e frustrante para os alunos, mas como algo que os impulse a desenvolver maior autonomia como aprendizes, buscando respostas ao visualizar dados em linhas de concordância. Seidlhofer (2002) relatou a sua experiência em deixar que os aprendizes analisassem suas próprias produções, incentivando-os a utilizarem recursos computacionais para comparar seus dados com os dados de nativos, de forma que eles conseguissem fazer descobertas acerca de sua própria interlíngua. Por fim, Nesselhauf (2005) identificou colocações verbo-nominais na escrita em inglês de falantes nativos do alemão e concluiu que algumas combinações merecem atenção especial nos materiais didáticos preparados para aprendizes com essa língua materna. No Brasil, pesquisas dessa natureza estão também acontecendo, como as pesquisas encabeçadas pelo grupo GELC – Grupo de Estudos em Linguística de Corpus (SHEPHERD, 2009a), por exemplo.

Ainda dentro desse cenário de pesquisas com *corpora* e *corpora* de aprendizes em benefício do ensino, Berber Sardinha (2004, p. 265) declara que “a descrição da linguagem de aprendizes a partir de *corpora* de textos ou transcrições de falas de não-nativos, especialmente de inglês como língua estrangeira, tem um impacto considerável sobre o ensino de línguas.” Considerando que há um tempo as preocupações acerca do ensino de línguas estrangeiras estavam centradas em três principais pilares, quais sejam a descrição da língua-alvo, a dimensão da instrução (materiais, conteúdos, atividades) e a caracterização do aprendiz, é possível concluir que o foco na produção propriamente dita do aprendiz deve ser devidamente incluído como o quarto pilar das preocupações com o ensino (GRANGER, 2002). Uma grande evidência de que tal preocupação com a produção dos aprendizes tem aumentado é a existência dos *corpora* de aprendizes de *Cambridge*⁶ e *Longman*⁷ (CAMBRIDGE, 2014; LONGMAN, 2014a), compilados com o objetivo de auxiliar lexicógrafos e

⁶ Disponível em: http://www.cambridge.org/gb/elt/catalogue/subject/custom/item3646603/Cambridge-English-Corpus-Cambridge-Learner-Corpus/?site_locale=en_GB. Acesso em: 04 abr. 2014.

⁷ Disponível em: <http://www.pearsonlongman.com/Dictionaries/corpus/learners.html>. Acesso em: 04 abr. 2014.

pesquisadores/autores a prover informações relevantes em dicionários e outros materiais confeccionados para cada tipo de aprendizes, com línguas maternas e níveis linguísticos diferentes e, por conseguinte, necessidades diferentes. Os pesquisadores dessas editoras, portando, utilizam os dados provenientes dos *corpora* de aprendizes para identificar erros cometidos e observar quais as palavras e estruturas que despertam maiores dificuldades de uso para os estudantes. Assim, como já mencionado, não só dicionários, mas também materiais didáticos para ensino regular e preparatórios para exames de proficiência obterão o conteúdo necessário para atender ao propósito de melhor informar os alunos e conduzi-los a uma aprendizagem mais plena e a uma melhor desenvoltura na língua. A título de averiguação, na seção de Anexos, foram incluídas duas figuras que mostram o que os sites de *Cambridge* e *Longman* informam a respeito dos propósitos dos *corpora* supracitados (ANEXOS A e B).

1.5 A organização da dissertação

Este estudo foi dividido em cinco capítulos, sendo este primeiro uma reunião das motivações para a condução dessa pesquisa, dos objetivos a serem alcançados e das perguntas a serem respondidas ao longo do trabalho.

O segundo capítulo congrega explicações sobre os fundamentos teóricos que direcionaram os tipos de análises conduzidas nesse estudo, além de apresentar um panorama histórico sobre quantificadores e *corpora* de aprendizes, e fornecer discussões e ponderações acerca dos conceitos de **prosódia semântica**, de erro e de conscientização linguística.

O terceiro capítulo delinea a metodologia empregada, desde a compilação dos *corpora* utilizados até o tipo de análises que foram conduzidas, como a observação das linhas de concordância e do contexto expandido para identificação tanto de padrões léxico-gramaticais quanto da emergência do fenômeno da **prosódia semântica**. A escolha do teste estatístico usado para validar os resultados também foi justificada.

O quarto capítulo descreve os detalhes dos resultados obtidos de ambas as formas, qualitativa e quantitativa. Além disso, nele é desenvolvida uma seção de implicações dos resultados na qual é proposta e teoricamente discutida uma sugestão de atividade pedagógica sobre o tema pesquisado.

O quinto e último capítulo responde às questões de pesquisa e apresenta as considerações finais, que apontam para possíveis desdobramentos desse estudo e para perspectivas de investigações futuras.

2 REVISÃO DE LITERATURA

O capítulo de revisão servirá para melhor situar o leitor quanto aos temas e teorias invocados nesse trabalho. Para satisfazer tal propósito, ele será dedicado a explicações mais detalhadas acerca dos seguintes tópicos, a saber: i. os fundamentos básicos sobre os quais este trabalho se apoia; ii. a apresentação de um breve panorama histórico e gramatical sobre quantificadores; iii. a exposição do conceito de **prosódia semântica** e discussões sobre esse fenômeno; iv. o provimentos de informações relevantes acerca do que caracteriza as noções tanto de erro quanto de conscientização linguística; e v. a menção de exemplos de trabalhos já realizados utilizando *corpora* de aprendizes e de trabalhos, particularmente, sobre quantificadores em *corpora*.

2.1 Fundamentos-base

Tem-se como verdade que “na vida dos indivíduos e das sociedades, a linguagem constitui fator mais importante que qualquer outro.” (SAUSSURE, 2006, p. 14). Levando isso em conta, é possível entender o porquê de Labov dizer que “a teoria da linguagem não pode mais ignorar o comportamento social de falantes de uma língua” (LABOV, 1972, p. 259), e também de se verificar o real uso que os falantes fazem de sua língua.

Discorrendo sucintamente a esse respeito, antes dos anos 1950, uma das tarefas do linguista era a de coletar dados de uma língua, derivar generalizações a partir deles e organizar descrições, realizando comparações e análises translinguísticas. Depois desse período, aconteceu o fenômeno chamado ‘revolução cognitiva’⁸, a partir do qual a língua passou a ser tratada como fenômeno exclusivamente cognitivo, e foram aquecidas as preocupações voltadas para a explicação das capacidades mentais dos falantes, o que fez com que a descrição das línguas fosse, conseqüentemente, relegada a segundo plano. Vale mencionar também que muitos estudos até meados de 1955 foram baseados em *corpora* compilados manualmente, e se concentraram na frequência de palavras, de padrões gramaticais e ortográficos. Entretanto, o linguista gerativista Noam Chomsky criticou duramente os estudos baseados em *corpus* e mudou o rumo das pesquisas, que passaram de empíricas para racionalistas (pautadas na introspecção), dicotomia essa que caracteriza um

⁸ Segundo Gardner (1985), ‘Revolução Cognitiva’ se configura como uma expressão para designar o movimento científico iniciado nos anos 50, o qual deu origem à chamada ciência cognitiva. Trata-se de um movimento interdisciplinar que combina novas ideias da psicologia, da antropologia e da linguística, mesclando-as com os campos da inteligência artificial, ciência da computação e neurociência.

embate antigo (MCENERY, WILSON, 1996, 2001) no método científico. Chomsky estabeleceu a diferença entre **competência** (conhecimento linguístico internalizado) e **desempenho** (o que é realmente falado), e afirmou que o **desempenho** não é um bom espelho para o conhecimento linguístico, mas que, na verdade, a língua deve ser explicada através da **competência** de um falante-ouvinte ideal. Chomsky também apontou para a finitude de um *corpus* e para o fato de ele poder ser tendencioso. (CHOMSKY, 1956)

Não obstante todas essas críticas, a estima em relação à empiria, que é apreciada e defendida nesse trabalho, voltou a ganhar força, – embora na Europa pesquisas com *corpus* nunca tenham deixado de existir, mesmo com as correntes de influências contrárias vindas de adeptos da teoria chomskyana. Mais ainda, vários pontos de refutação das críticas feitas foram levantados, como o fato de os avanços tecnológicos dos dias atuais permitirem a existência de *corpora* de bilhões de palavras. Tais *corpora* são, em grande parte, de acesso público, podendo, portanto, ser verificados pela comunidade científica; além disso, eles subsidiam estudos diversos cujos resultados podem ser replicados. Outra questão a ser colocada é que a introspecção não exclui, necessariamente, a análise baseada em dados; pelo contrário, os dois métodos podem ser combinados, e os *softwares* para leitura e manejo dos dados podem auxiliar muito os linguistas em suas perguntas de pesquisa, além de se ter que considerar que tais *softwares* fornecem informações de estimativas numéricas que nunca estarão precisamente disponíveis à introspecção. Isso é consoante com o que Halliday (1991, 1992) defende, quando afirma que a linguagem é um sistema probabilístico. Logo, é recomendado que possíveis intuições introspectivas sejam testadas a partir de *corpora* e ferramentas disponíveis. Inclusive, como comenta Berber Sardinha (2000, p. 343) “o conhecimento da probabilidade de ocorrência de traços lexicais, estruturais, pragmáticos discursivos, etc. está no cerne da Linguística de Corpus” e deve ser praticado com responsabilidade por pesquisadores da língua.

Tendo em vista, contudo, que não é o objetivo desse trabalho aprofundar nos debates e divergências entre as visões chomskyana e hallidayana, e entendendo que os pressupostos linguísticos necessários e pertinentes aqui já foram expostos, cumpre agora apresentar, nesse instante de forma mais panorâmica, alguns pontos e fundamentos importantes sobre a LC e, mais especificamente, sobre os estudos com *corpora* de aprendizes, bem como fundamentos básicos sobre **prosódia semântica**⁹.

⁹ Esse conceito será amplamente discutido na seção 2.3 deste capítulo.

As aplicações da LC podem ser apreciadas tanto na área da Lexicografia, quanto nos estudos sistemáticos do uso da língua, bem como nos estudos da Tradução, da Linguística Aplicada ao ensino e do Processamento de Linguagem Natural. Isso se deve majoritariamente à proposta de Sinclair (1978, apud BEBER SARDINHA, 2004), ao afirmar que a língua deveria ser estudada a partir de exemplos reais de uso, e não somente a partir de textos criados com propósitos de exemplificação. Em se tratando mais especificamente da LC sendo usada a serviço da Linguística Aplicada ao ensino, Berber Sardinha (2004) fala sobre o possível impacto da compilação de *corpora* de aprendizes sobre o ensino, e a observação desses *corpora* pode levar a questões como as que seguem abaixo, segundo elencadas por Leech (1998, p. xiv):

Quais características linguísticas da língua-alvo são empregadas com mais (sobreuso) ou menos (subuso)¹⁰ frequência em comparação com falantes nativos?;

- i. Qual é a extensão da influência da língua nativa (transferência) na produção dos aprendizes?;
- ii. Em que áreas eles tendem a usar estratégias de ‘evitação’, deixando de explorar a fundo o potencial da língua-alvo?;
- iii. Em que áreas eles tendem a demonstrar desempenho nativo ou não-nativo?;

Quais são as áreas nas quais os aprendizes de um dado país parecem necessitar de mais ajuda para desenvolver sua produção na língua-alvo? (tradução minha)¹¹

Como já mencionado, o presente trabalho se preocupou em tentar responder, mesmo que em meio a algumas limitações, como o tamanho dos *corpora* utilizados aqui, algumas dessas questões no tocante ao objeto de estudo escolhido, qual seja o uso das expressões quantificadoras *a few* e *few* em *corpus* de aprendiz, cujos informantes são brasileiros e aprendizes de inglês em nível de intermediário a avançado.

O conceito relativo à **prosódia semântica** também é invocado nesse trabalho, e por hora basta esclarecer que a criação do termo **prosódia semântica** é atribuída a Louw

¹⁰ Os conceitos de **sobreuso** e **subuso** serão tratados posteriormente nesse capítulo, na seção 2.5.1.

¹¹ “What linguistic features in the target language do the learners in question use significantly more often (‘overuse’) or less often (‘underuse’) than native speakers do?; How far is the target language behavior of the learners influenced by their native language (NL transfer)?; In which areas do they tend to use ‘avoidance strategies’, failing to exploit the full range of the target language’s expressive possibilities?; In which areas do they appear to achieve native-like or non-native-like performance?; What (in order of frequency) are the chief areas of non-native-like linguistic performance which learners in country A suffer from and need particular help with?”

(1993). Sinclair (1991) e Partington (1998) também desenvolveram noções sobre esse conceito, sendo que o último o interpreta como o espalhar de uma coloração de conotação que transcende os limites de uma única palavra. Berber Sardinha (2004, p. 236) comenta sobre o casamento entre **prosódia semântica** e tradução, dizendo que a primeira é importante para o entendimento da segunda, já que faltam indicações mais sistemáticas em dicionários e manuais de tradução acerca de vocabulários tidos como equivalentes. Levando em conta, então, que termos equivalentes podem apresentar prosódias semânticas distintas, “um tradutor pode utilizar a prosódia semântica inadequada, sem saber.” É indubitável, portanto, a necessidade desse conhecimento para tradutores; contudo, não deve se esgotar aí. Aprendizizes de um modo geral, na verdade, podem ser grandemente beneficiados ao adquirirem noções de **prosódia semântica**.

Por isso, esse trabalho também traz a proposta de uma aliança entre a **prosódia semântica** e o ensino. Tal proposta está fundada no que Conrad (2000) acredita ser uma grande mudança incitada por estudos baseados em *corpora* para o ensino, e essa mudança consiste justamente no ensino integrado de vocabulário e gramática, isto é, aderindo a uma perspectiva léxico-gramatical. Sendo assim, ao se ensinar as expressões quantificadoras *a few* e *few*, não basta dizer, por exemplo, que de acordo com a norma gramatical padrão, elas devem ser seguidas de substantivos contáveis, mas deve-se ir além disso, e informar também sobre os colocados mais frequentes à direita e à esquerda, sobre padrões recorrentes nos diferentes registros existentes (acadêmico, conversação, etc.), e sobre as implicações de **prosódia semântica** no uso desses quantificadores. Para tanto, é recomendada a utilização de linhas de concordância em sala de aula, e tal procedimento de exploração dessas linhas, como aponta Kettemann (1995), pode oferecer uma experiência altamente motivadora para o aprendiz. Assim sendo, professores de inglês, além de fazerem uso de dicionários e gramáticas baseados em *corpus* como obras de referência, podem trazer para a sala de aula a proposta de Johns (2002, p. 108), oriunda da **Aprendizagem Movidada por Dados**, que é a de formular atividades com o intuito de despertar nos alunos o gosto por desempenharem “o papel de Sherlock Holmes”, ou seja, de detetives da língua que buscam por evidências de padrões linguísticos.

Tem-se, portanto, que este trabalho se sustenta em análises empíricas de dados de aprendizizes brasileiros de inglês, em nível avançado, análises essas que procuram oferecer algum rumo de respostas para as perguntas de pesquisa citadas na Introdução, que são baseadas naquelas elencadas por Leech (1998), principalmente quanto ao **sobreuso** ou **subuso** de colocações com quantificadores e sobre a extensão da influência da língua materna nas

escolhas léxico-gramaticais feitas pelos aprendizes. Busca-se, também, entender melhor a interação que existe entre as expressões quantificadoras em estudo e a **prosódia semântica**.

2.2 Panorama histórico e gramatical sobre os quantificadores

Desde a década de 70, Hintikka dizia que “a sintaxe e a semântica de quantificadores representam temas de estudo altamente significativos para a teoria linguística” (1973, p. 329). Ele ilustrativamente retoma Lakoff (1971), cujo trabalho se apoiou na gramática de quantificadores como base de teste para analisar o comportamento dos mesmos no que diz respeito às seguintes situações: preservação de sentido em operações como as de transformações (exemplificadas por deslocamentos, topicalizações), análises de correferência, e verificação do papel da estrutura superficial na interpretação semântica.

Desse modo, é possível afirmar que o estudo da quantificação já contribuiu para descobertas linguísticas interessantes e certamente agora, época em que a linguística vislumbra uma relação mais estreita com métodos empíricos e ferramentas propiciadas pela LC, há um vasto terreno a ser percorrido e explorado na esfera da quantificação. Pautando-se em estudos com um viés empírico e probabilístico, muitos degraus já começaram a ser galgados em diversos domínios linguísticos por diferentes pesquisadores e estudiosos. De Bod, Hay, Jannedy (2003), extraem-se exemplos, como os seguintes capítulos: o de Pierrehumbert, que investiga aspectos fonológicos a partir desse olhar; o de Baayen, que traz abordagens probabilísticas para a morfologia; o de Manning, que trata da influência de informações semânticas e discursivas sendo analisadas sob um prisma estocástico ao se entrelaçarem com a sintaxe; e o de Cohen, que comenta especificamente sobre a semântica probabilística.

Assim sendo, é possível entender a relevância de se realizar trabalhos que analisem itens lexicais e colocações, no caso aqui os quantificadores *a few* e *few*, dentro de uma perspectiva que englobe a probabilidade, a verificação de frequências e colocados, e a identificação de padrões recorrentes em determinados contextos e registros, para que tudo isso possa, conjuntamente, contribuir para uma compreensão mais sólida e profunda do objeto sendo estudado.

2.2.1 Um breve histórico

Desde o tempo em que Aristóteles introduziu a lógica, a quantificação se configurava como um tema que incitava curiosidade. Naquela ocasião, o filósofo se

concentrou no sentido de alguns determinantes naturais da linguagem, como ‘todos’, ‘alguns’, ‘nenhum’, ‘nem todos’, cujas relações lógicas recíprocas são graficamente expressas no seguinte quadrado de oposições (D’ALFONSO, 2011, p. 85):

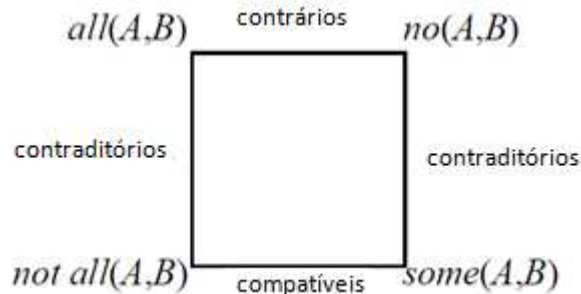


Figura 2 – Quadrado de oposições de quantificadores proposto por Aristóteles
Fonte: D’ALFONSO, 2011, p. 85.

A lógica aristotélica pode ser considerada como um segmento da lógica da quantificação natural, que delibera quantificadores que podem ser expressos por Sintagmas Nominais (SNs). Quando, contudo, Frege (1879, apud PINTO, 2011) inventou a teoria da lógica formal, dois quantificadores lógicos artificiais, \forall e \exists , que são estruturalmente diferentes dos naturais, foram introduzidos como operadores que juntam variáveis. No século XX, ocorreu um divórcio entre a lógica e a linguagem natural (D’ALFONSO, 2011) causado pelo fato de esses quantificadores lógicos de primeira ordem, muito atrelados à argumentação lógica em si, caracterizarem-se, segundo os especialistas da área, como inadequados para lidar com sentenças quantificadas da linguagem natural, por haver um problema de interface e de expressividade nessa transposição. Nesse ínterim, a teoria acerca dos **Quantificadores Generalizados**¹² chega para trazer novos caminhos que aproximam a sintaxe lógica da sintaxe da linguagem natural. Contudo, o presente trabalho não se prestará a explicar em detalhes as questões referentes a essas teorias, pois não é o propósito aqui aprofundar em problemáticas de quantificação ligadas à lógica.

2.2.2 Quantificadores Generalizados

A fim de visitar um pouco a teoria de quantificação e trazer alguns conceitos desse campo de pesquisa, é importante dizer que de acordo com os estudiosos Barwise e Cooper (1981, p. 176), os quantificadores artificiais citados na seção anterior são extremamente atípicos e possuem propriedades especiais que podem causar enganos. Esses mesmos autores declaram que, em parte, a teoria linguística de um modo geral está preocupada com universais de línguas naturais, isto é, fatos que se mantêm para todas as

¹² Esse conceito será brevemente tratado na próxima seção.

línguas humanas de ocorrência natural e que as distinguem de outras línguas que poderiam ser logicamente possíveis. Pelo fato, então, de haver uma preocupação com línguas naturais, os autores declaram que os quantificadores artificiais são protagonistas de uma teoria obsoleta.

Adentrando as elucidações propriamente ditas a respeito de quantificadores generalizados, Keenan (2006) explica que o tipo de quantificação melhor compreendido em línguas naturais é aquele exemplificado pela frase ‘*All poets daydream*’ (‘Todos os poetas divagam’ – tradução minha). O verbo ‘divagar’ é tratado como denotando uma propriedade de indivíduos, representada pelo subconjunto de um domínio E de objetos em discussão. O SN quantificado – ‘todos os poetas’ – é tratado como denotando funções, chamadas **quantificadores generalizados**, que mapeiam ou relacionam propriedades a valores de verdade (verdadeiro ou falso). O que ocorre, portanto, é que a relação entre o SN conectado ao verbo só mapeará a uma verdade caso o conjunto de poetas seja um subconjunto do conjunto de objetos que divagam.

Indo um pouco mais adiante, Szabolcsi (2010) destaca que uma importante inovação trazida por Montague (1974, apud SZABOLCSI, 2010) foi o provimento de um tipo de denotação uniforme para todos os Sintagmas Determinantes (SDs)¹³ na forma de quantificadores generalizados, introduzidos matematicamente nos estudos de Mostowski (1957, apud SZABOLCSI, 2010) baseados na ideia fundamental de Frege. O termo ‘generalizados’ é explicado pelo fato de que se generaliza a partir de quantificadores e na verdade, como já dito, um quantificador generalizado é um conjunto de propriedades que são distinguidas ou não pela semântica extensional¹⁴, dependendo da coincidência de elementos que compõem cada conjunto. Assim, na frase ‘Todo dragão voa’, tem-se que o conjunto de dragões é {a, b, c} e o conjunto de coisas que voam é {a, b, d}; logo, as propriedades **ser um dragão** e **ser uma coisa que voa** podem ser distinguidas. Se, contudo, os conjuntos forem compostos pelos mesmos elementos, a semântica extensional não os consegue distinguir. Sumariamente, Montague supõe que frases quantificadas se combinam com expressões que denotam propriedades, e que o efeito semântico dessa combinação é asseverar que a propriedade se configure como um elemento do quantificador generalizado.

¹³ Cumpre dizer que muitos linguistas, em vez de usar a nomenclatura SN (ou NP, no inglês), preferem usar SD (DP, no inglês) em parte porque propriedades significativas da expressão como um todo são muitas vezes definidas pela escolha do determinante (todavia, há propriedades, como a animacidade, o gênero e restrições seletivas de predicados, que são estipuladas pelo nome, não pelo determinante).

¹⁴ Ou denotacional (que atribui a cada constituinte sintático uma extensão ou denotação).

A discussão se torna mais instigante, especificamente para os propósitos desse trabalho, a partir do momento em que entram em cena os *negative polarity items* (NPIs), ou itens de polaridade negativa, em português. Exemplos deles na língua inglesa são *ever* e *not*, e existem alguns SDs que licenciam tais itens (e.g. ‘*No student here has ever been to Pinsk*’). Tais sintagmas são classificados como quantificadores generalizados decrescentes.¹⁵

2.2.3 Monotonicidade dos quantificadores

Barwise e Cooper (1981) destacam duas subclasses de quantificadores, quais sejam os monotons crescentes (*‘monotone increasing’*) e os monotons decrescentes (*‘monotone decreasing’*). A fim de testar um sintagma nominal quanto a sua monotonicidade, deve-se tomar dois sintagmas verbais (SV1 e SV2) de forma que a denotação de SV1 seja um subconjunto da denotação de SV2, e assim verificar se as proposições seguintes parecem logicamente possíveis:

- Se SN SV1, então SN SV2. (SN é monotom crescente)
- Se SN SV2, então SN SV1. (SN é monotom decrescente)

Por exemplo, se SV1 for **entraram na corrida cedo** e SV2 for **entraram na corrida**, tem-se o seguinte:

- ⇒ Se **muitos homens** entraram na corrida cedo,
então **muitos homens** entraram na corrida.

Todavia, a implicação reversa não se aplica, já que pode haver homens que entraram na corrida, porém não entraram cedo. Dessa forma, o SN ilustrado (**‘muitos homens’**) se trata de um quantificador monotom crescente, e por isso a validade da implicação supramencionada se confirma.

De igual modo, deve-se considerar o seguinte:

- ⇒ Se **poucos homens** entraram na corrida,
então **poucos homens** entraram na corrida cedo.

Quando se tem o exposto acima, é verificada a validação da implicação para o quantificador monotom decrescente.

Na língua inglesa, alguns exemplos de quantificadores monotons crescentes são *some*, *many*, *most*, entre outros, e exemplos de quantificadores monotons decrescentes são *no*, *few*, *neither*. Há também os casos de quantificadores que não são monotons, como

¹⁵ Tal conceito será explicado na próxima seção. Por hora, segue o exemplo adicional: ‘*Either John or Bill ever helped me’ VS. ‘Neither John nor Bill ever helped me’ são frases que mostram que o DP representado na segunda frase é o que se configura como quantificador generalizado decrescente e licencia o NPI *ever*.

exactly. Os mesmos autores comentam, e isso se faz relevante para o presente trabalho, sobre sintagmas nominais que poderiam ser capazes de denotar ambos quantificadores monotons e não-monotons. Nesse tocante, eles dizem que se *a few* for utilizado para dar o sentido de *some but not many* ('alguns, mas não muitos'), logo *a few men* não é um monotom; se, contudo, *a few* for utilizado para dar o sentido de *at least a few* ('pelo menos alguns'), *a few* é um monotom crescente. Conjectura-se, então, que enquanto *few* se configura como um monotom decrescente que licencia itens de polaridade negativa, *a few* pode engendrar mais de uma leitura, e provavelmente, qualquer que seja essa leitura, ela não será tão restritiva quanto a que se faz de *few*.

A suposição descrita, portanto, está ligada às investigações conduzidas nessa pesquisa, que visam à melhor compreensão do uso desses quantificadores. Todos os achados brevemente colocados acima, embora não tenham sido frutos de pesquisas empíricas ou pesquisas que levassem em conta a metodologia de *corpus*, podem ser reavaliados e verificados à luz dessas direções de análise. Além disso, pode-se também levantar a hipótese de que o conceito da **prosódia semântica** está, em alguma medida, relacionado com a teoria dos itens de polaridade negativa. Reforço, novamente, que não há como 'sincronizar' tais teorias, dado que elas estão localizadas em tempos diferentes e possuem origens e fundamentos de análise distintos. Ainda assim, é possível vislumbrar algumas similaridades e pontos em comum por elas compartilhados, e um ponto de encontro aqui em destaque leva em consideração a noção de que algumas palavras ou agrupamentos de palavras carregam um determinado 'tom' ou prosódia.

2.2.4 *Os quantificadores em gramáticas*

Mostrou-se pertinente relatar, com o objetivo de melhor compreender o tratamento que os quantificadores protagonistas desse trabalho recebem em gramáticas, informações obtidas por meio de consultas a elas, tanto gramáticas do inglês como do português. Pretende-se, a partir dessas consultas, averiguar como as gramáticas, especialmente da língua inglesa, tratam os quantificadores no sentido de estabelecer alguma distinção quanto ao uso de *few*, quando este é precedido ou não do artigo *a*, e também se é feita alguma menção acerca do uso pragmático desses quantificadores, isto é, se há favorecimento de uso de um ou outro quando o contexto é negativo ou positivo, se há padrões que geralmente são escolhidos e mais usados em determinados contextos semânticos ou em caso de registro mais ou menos formal.

2.2.4.1 As gramáticas do inglês

Começando com BIBER *et al.*, 1999, que se configura como grande referência de gramática produzida com base em dados retirados de *corpora*, ela trata *a few* e *few* como determinantes que especificam substantivos em termos de quantidade, sendo, por isso, denominados quantificadores, e os inclui no grupo intitulado ‘Moderados ou de pequena quantidade’. Os autores dizem que *a few* e *few* são usados com substantivos contáveis no plural, e que enquanto *a few* se aproxima semanticamente de *some*, *few* sugere uma quantidade menor do que a esperada (BIBER *et al.*, 1999, p. 276). Dentre os quantificadores ‘moderados ou de pequena quantidade’, *some* é o mais recorrente no registro prosa acadêmica. Quanto aos quantificadores em estudo nessa pesquisa, o registro no qual *a few* é encontrado com maior frequência é *Fiction* (Ficção), e o registro com mais ocorrências de *few* é o de *News* (Notícias). Na versão estudantil dessa gramática, (BIBER; CONRAD; LEECH, 2002), informações bem semelhantes são encontradas para *a few* e *few*. O primeiro, juntamente com *several*, está classificado como quantificador de plurais contáveis que expressa um número ou quantidade pequena. Já o segundo, juntamente com as derivações *fewer* e *fewest*, está classificado como quantificador de plurais contáveis que expressa uma quantidade que é menor do que o esperado. Os autores dizem que *a few* é utilizado para indicar quantidade pequena, como no exemplo ‘*There were a few people sitting at the tables in the back*’ (‘Havia algumas pessoas sentadas nas mesas de trás’, tradução minha), fornecido nessa gramática e retirado de um *corpus* cujo registro é Ficção. Por sua vez, *few* (sem ser precedido do artigo *a*) significa ‘não muitos’, e o exemplo dado, ‘*Very few women have hair that’s that short*’ (‘Muito poucas/Não muitas mulheres têm cabelo que é tão curto’, tradução minha), retirado de um *corpus* de registro Conversação, é acompanhado do intensificador *very*. A título de ilustração, a figura (BIBER; CONRAD; LEECH, 2002, p. 75), na qual constam essas informações e os exemplos citados foi incluída na seção de Anexos (ANEXO C).

Em consulta a Swan (2005, p. 309), uma gramática mais pedagógica, há uma seção que aborda especificamente o uso do artigo *a*, evidenciando que sem esse artigo, *few* geralmente possui um sentido negativo que sugere ‘não tanto quanto se gostaria’; em contrapartida, *a few* é mais positivo, e relaciona-se com o sentido que o quantificador *some* denota, perpassando as ideias de que a quantidade referida é ‘melhor do que nada’ ou ‘maior do que o esperado’. Ainda, essa mesma gramática faz menção à questão da adequação aos discursos formal e informal, alertando para o fato de que sem ser precedido do artigo *a*, *few* se

caracteriza como mais formal, enquanto que em estilo informal (fala cotidiana), *only a few* é mais usado. Em tempo, o autor chama atenção para o fato de que *very few* pode também ser utilizado em estilo informal¹⁶. Novamente, para fins ilustrativos, uma figura (SWAN, 2005, p. 309) foi incluída (ANEXO D).

2.2.4.2 As gramáticas do português

Com relação ao tratamento que as gramáticas da língua portuguesa dão aos quantificadores, pode-se dizer que tal termo não é facilmente encontrado nelas. Após consultar Cipro Neto e Infante, 2004 e Cegalla, 2008, por exemplo, não foi possível encontrar uma seção específica que abordasse os quantificadores. Já ao consultar Moura Neves (2000), o vocábulo **pouco** está inserido na seção de pronomes indefinidos, dentro do capítulo de quantificação e indefinição. Em ambas as suas formas, singular e plural, Moura Neves (2000, p. 562) descreve que o uso de **pouco(s)** serve “para indicar algo que existe em pequena quantidade, quase sempre insuficiente (indicação avaliativa).” Quando se trata do plural, ou seja, de **poucos**, a autora também explica que o sintagma com esse vocábulo pode vir determinado por artigo definido ou artigo indefinido, sendo que no último caso ele é geralmente anteposto, como em ‘uns poucos amigos’.

Devido ao fato de não se ter conseguido muitas informações sobre o tema de pesquisa desse trabalho nas gramáticas consultadas, surgiu a necessidade de recorrer a outras fontes a fim de compreender melhor as características dos quantificadores na língua portuguesa. Leal (2009), em sua dissertação sobre a semântica das expressões nominais na língua portuguesa, explica que a quantificação consiste em um fenômeno linguístico que afeta vários domínios, já que “a expressão de quantidade pode se referir a entidades de diferentes estatutos ontológicos” (LEAL, 2009, p. 112). Ele inclusive ressalta que o fenômeno não se limita apenas ao domínio nominal, mas também se expande ao domínio verbal (como no exemplo ‘O aluno nadou durante dez minutos’). Igualmente, ele esclarece que os quantificadores são tradicionalmente designados na gramática como determinantes, adjetivos, pronomes ou advérbios, e sua distribuição não se dá de forma regular.

A proposta de classificação de quantificadores de acordo com Sánchez López, 1999 (apud LEAL, 2009) inclui, por exemplo, ‘algum(ns)’, ‘pouco(s)’, e ‘muito(s)’ no conjunto de quantificadores não-universais (ou indefinidos), os quais medem o tamanho de

¹⁶ Swan (2005, p. ix) afirma que as explicações fornecidas na gramática refletem o inglês britânico moderno de uso corrente, sendo que os exemplos dados se baseiam nas amostras de língua oral e escrita provenientes de grandes bancos eletrônicos de dados autênticos. Entretanto, não há menção específica de quais bancos de dados foram consultados.

uma forma aproximada, de um grupo indeterminado de elementos, ou expressam a parte indeterminada de um grupo definido previamente introduzido ou não no discurso. Especificamente sobre ‘alguns’, que seria o correspondente à *a few* ou *some* no inglês, o autor em questão explica que se trata de um quantificador de tipo existencial (e.g. ‘Há alguns problemas na sala’) que denota uma quantidade não total. Lopes 1972, (apud LEAL, 2009) diz que ‘alguns’ é usado também em acepções que se aproximam de ‘poucos’ ou de ‘não muitos’. Mesmo assim, embora a linha divisora de sentido possa parecer tênue, o raciocínio de que ‘alguns’ e ‘poucos’ se aproximam não deve ser levado a tal ponto de se acreditar que tais quantificadores possam ser usados de forma intercambiável. Sobretudo, isso definitivamente não pode ocorrer ao se pensar na transposição feita entre línguas, quando da correspondência com o inglês ao longo da aprendizagem dessa língua-alvo pelos brasileiros.

2.3 Panorama sobre prosódia semântica

O termo **prosódia semântica** (PS, doravante) foi primeiramente utilizado por Louw (1993), mas amplamente desenvolvido por Sinclair (1991¹⁷), e também explorado por outros linguistas, como Stubbs (1996, 2001), Tognini-Bonelli (2001) e Partington (1998, 2004). A palavra ‘prosódia’ foi emprestada do uso feito por Firth (1957) quando ele se referia à coloração fonológica que se espalha por entre as fronteiras semânticas das palavras. Da mesma forma que a palavra ‘tâmara’ é dotada de uma prosódia nasal tão forte que a vogal inicial *a* adquire uma qualidade nasal por assimilação, itens lexicais também compartilham desse fenômeno prosódico em sua disposição sequencial. É importante ressaltar que os avanços obtidos para os estudos baseados em *corpus* formataram um palco propício para o aquecimento das investigações sobre PS.

2.3.1 Apanhado de conceitos de prosódia semântica

A definição de Louw (1993, p. 157) para PS é que o fenômeno se trata de “uma aura de sentido consistente com a qual uma forma é imbuída pelos seus colocados.” Embora tal definição reflita a posição linguística de que forma e sentido são inseparáveis, não há reverberações quanto às funções da PS ou seu sentido pragmático, sentido este que captura elementos não-linguísticos que impactam a comunicação.

Já Sinclair (1996, p. 86), considerado o pai dos estudos sobre PS, fornece a seguinte definição: “A escolha inicial da **prosódia semântica** é uma escolha funcional que

¹⁷ Sinclair já falava do comportamento colocacional das palavras nessa época, mas não com o termo **prosódia semântica**, o qual foi cunhado por Louw.

conecta sentido a finalidade; todas as escolhas subsequentes dentro do item lexical se referirão à prosódia”. A partir de tal definição, é possível evidenciar três traços importantes, de acordo com Zhang (2010): primeiro, a funcionalidade, já que quando um falante/escritor escolhe um item lexical para formar orações, além de se atentar às regras léxico-sintáticas que governam a gramaticalidade da oração, esse falante também deve considerar a preferência semântica que aponta para as funções; em segundo lugar, a escolha linguística, uma vez que a combinação que se dá em cada colocação não é simplesmente arbitrária, mas todas as palavras se encontram numa relação seletional mútua; e em terceiro lugar, o propósito comunicativo, tendo como verdade o fato de que as prosódias semânticas são atitudinais e se encontram no lado da pragmática em meio ao contínuo semântica-pragmática, de forma que, com a PS adequada, o falante/escritor certamente expressará a atitude pretendida com maior harmonia e clareza. Inclusive, um dos objetivos do presente trabalho é poder observar a interface da PS com o uso dos quantificadores *a few* e *few* e colocados, a fim de melhor entender como os nativos expressam algum tipo de prosódia nesse tocante para satisfazer um propósito comunicativo, e como os aprendizes utilizam tais estruturas em contextos comunicativos semelhantes. Mesmo oferecendo um maior estreitamento com a pragmática, ainda não há na abordagem de Sinclair um esclarecimento mais transparente quanto a quais funções e propósitos as prosódias semânticas exercem no uso da língua.

Para Stubbs (1996), PS é um fenômeno colocacional particular, e colocação é a co-ocorrência habitual de duas ou mais palavras. O autor também oferece três categorizações para a PS, quais sejam a prosódia negativa, a prosódia positiva, e a prosódia neutra. Este trabalho, entretanto, não considera a prosódia neutra como relevante, pela crença de que existe sempre uma manifestação de atitude no discurso dos indivíduos. Vale mencionar que outra contribuição do linguista Stubbs é a de destacar que a PS de modo algum é uma simples colocação, que emerge no nível léxico-gramatical; na verdade, a PS é construída a partir de um nível pragmático. Por isso, definir PS como um fenômeno colocacional traz apenas um quadro parcial do que realmente ocorre, e é insuficiente para congrega a síntese completa de ambos os sentidos semântico e pragmático.

Quanto a Partington (1998), ele insere uma nova explicação para o termo PS, definindo-o como a expansão de coloração conotacional para além das fronteiras de uma palavra isolada. Segundo ele, as prosódias semânticas são interpretáveis com base em conotações, isto é, se o nóculo (palavra central) e seus colocados tendem a co-ocorrer frequentemente, eles irão adquirir os mesmos traços conotacionais, fazendo com que forma e sentido se combinem em uma fusão.

2.3.2 *Problematização dos conceitos*

Cada um dos conceitos apresentados na seção anterior possui validade e apreço na literatura que abrange pesquisas concernentes à colocações e discurso. Contudo, há muito debate no que diz respeito às asserções feitas por linguistas quanto ao que eles agregam ao conceito de PS, o que torna esse objeto de estudo um tanto quanto contencioso. Esse é o motivo pelo qual é necessário ter cautela ao abordar esse tema, e também delimitar uma posição conceitual a ser adotada para que haja coerência e harmonia nas análises.

Recuperando Hunston (2007), ela revisita a PS e alerta o leitor para as armadilhas de se usar linhas de concordância para observar linguagem atitudinal em textos altamente argumentativos. De fato, e especialmente porque textos escritos são desprovidos de informações entonacionais, é minimamente plausível pensar que autores possam lançar mão de estratégias de intertextualidade, ou fazer combinações lexicais inusitadas, ou romper com paralelismos semânticos com o objetivo de imprimir um determinado efeito. A interpretação disso é dada pelo leitor, mas o cuidado que se deve ter reside no fato de que as reações subjetivas podem ser diversas, dependentes de crenças, conhecimentos prévios, e percepção linguística individual, e isso faz com que declarações acerca da PS necessitem de comprovações empíricas robustas.

Um dos críticos de trabalhos sobre PS é Whitsitt (2005), que questiona o quanto ela se prolonga de um contexto para outro, ou seja, se haverá ‘coloração’ na interpretação caso um item lexical seja usado em um contexto tipicamente inesperado para ele. Obviamente, se uma palavra é usada de uma determinada forma na maioria dos contextos, isso não significa que ela será usada dessa mesma forma em todos os contextos. Hunston (2007), portanto, diz que para comprovar com mais segurança a existência da PS¹⁸, deve-se buscar uma consistência na função discursiva de uma sequência de itens – ou unidade de sentido formada por uma série de coocorrências, – e observar com precisão o co-texto e a fraseologia. Essas análises podem ser desenvolvidas a partir de observações de dados provenientes de *corpora* criteriosamente compilados, e vão ao encontro do que Bennett (2010) alega, quando ela diz que estudos com *corpora* (checagem de linhas de concordância, do contexto expandido e de outras informações fornecidas por *corpora*) conseguem responder questões diversas que envolvem a dimensão da fraseologia, a qual engloba colocações e agrupamentos lexicais.

¹⁸ Ou preferência semântica ou atitudinal, que se trata de um termo sugerido por Hunston (2007) para indicar a co-ocorrência frequente de um item lexical com itens que expressam um determinado sentido avaliativo.

2.3.3 Estudos relevantes de PS e o entrelaçar com o ensino

Louw (2000) declarou que as prosódias semânticas negativas são muito mais frequentes do que as positivas, e vários trabalhos já realizados confirmam isso, como, por exemplo, os trabalhos de Sinclair (1991), que investigou, entre tantas estruturas, o verbo frasal *set in*, do próprio Louw (1993), que teve como uma de suas pesquisas a prosódia do advérbio *utterly*, de Stubbs (1995), que observou os colocados de *cause*, como verbo e substantivo, e de Partington (1998), que conduziu diversos estudos, como do adjetivo *rife* – e todos esses itens investigados por esses pesquisadores estavam manifestos em contextos cuja prosódia é mais desfavorável, ou negativa. Ilustrando o exemplo do estudo de Sinclair (1991), ele constatou que colocados que expressam situações desagradáveis são os mais recorrentes em companhia a *set in*, como *decadence*, – ‘decadência’ – *infection*, – ‘infecção’ – e *prejudice* – preconceito.

Importa dizer que mais do que simplesmente definir entre negativo ou positivo, Stubbs (1995) expandiu a noção de PS, explicando que além de ter como colocados agrupamentos de palavras negativas ou positivas, itens lexicais também podem se colocar com grupos semânticos. Hunston (2002) foi nessa mesma direção de análise e, em seus estudos, atestou que a palavra *unemployment* apresenta a tendência de se colocar com o grupo semântico de *statistics*. Mais ainda, Tribble (2000) lançou a noção de que itens lexicais podem ser caracterizados ambos com uma PS ‘global’, em termos da língua como um todo, e ‘local’, quando está atrelada a um determinado contexto ou registro. Um exemplo disso é o estudo feito por esse pesquisador da palavra *experience* em um *corpus* baseado em gênero, – no inglês, *genre-based corpus*, composto por propostas de projetos europeus – por meio do qual ele descobriu que o sentido local expresso pela palavra *experience* nesse determinado gênero é o de ‘capital profissional’, e tal definição não é contemplada dessa forma em dicionários. Outra contribuição para a compreensão da PS foi, novamente, dada por Louw (1993), que evidenciou a associação entre ela e princípios gramaticais ao observar o uso transitivo de *build up*, revelando uma prosódia positiva (*build up confidence*), e o uso intransitivo, que mostra uma prosódia negativa (*resistance builds up*).

Dessa forma, mesmo que as características do que tipifica a PS sejam debatidas por diferentes pesquisadores, é impossível negar que o conceito que subjaz esse fenômeno aumentou em seu escopo, e isso possibilitou abertura para que mais pesquisas fossem realizadas e mais descobertas fossem trilhadas e alcançadas. Desde que resguardados os cuidados pertinentes a qualquer investigação empírica, cuidados tais que incluem a utilização

de dados significativamente robustos em sua arquitetura e devidamente tratados qualitativa e quantitativamente, novas pesquisas podem e precisam continuar emergindo com o intuito de iluminar o entendimento acerca da natureza e manifestações da PS, como fenômeno intrigante e instigante que é. Entendendo que a inabilidade de percepção apropriada da **prosódia semântica** também pode resultar erros na produção de aprendizes, vale dedicar uma parte deste trabalho para a discussão acerca das noções que estudiosos da área levantam sobre erros.

2.4 Noções sobre erro e conscientização linguística

Em se tratando de aprendizagem de Língua Estrangeira (LE), é necessário fazer menção da existência de erros na produção dos aprendizes e de como isso tem sido dimensionado por pesquisadores, além de como a literatura da área aborda esse tema.

Segundo Santos (2011, apud FERNÁNDEZ, 2011), surge nos anos 40, a partir de seus precursores Fries (1945) e Lado (1957), a Análise Contrastiva (AC). Ela surge como uma metodologia que compara a língua materna (LM) do aprendiz e a língua-alvo que o mesmo se propõe a aprender com o intuito de estabelecer as diferenças e as semelhanças entre as duas. Com isso realizado, espera-se chegar a uma predição das possíveis áreas nas quais os estudantes teriam mais dificuldades ao longo de seu processo de aprendizagem. Em sua versão forte, é declarado pelos seguidores da AC que quanto maior a distância entre as línguas sendo contrastadas, e quanto mais numerosos os itens linguísticos diferentes entre elas, maior será a dificuldade na aprendizagem, devido à maior interferência da LM (SANTOS GARGALLO, 1993). Por conseguinte, falantes nativos do português teoricamente teriam um percurso mais tranquilo de aprendizagem, por exemplo, da língua espanhola – também de origem românica – do que da língua inglesa, cuja origem é germânica. Tal visão pode ser considerada um tanto polêmica, pela falta de comprovações sólidas que atestam a real influência da língua materna ao longo do processo de aquisição de uma outra língua. Na versão fraca da AC, contudo, postula-se que é preciso primeiro observar a presença de dificuldades, ou erros, para então lançar mão de estudos contrastivos que expliquem tais desvios.

Aproximadamente a partir da metade da década de 70, por causa da influência do cognitivismo nos estudos linguísticos, surgiu a chamada Análise de Erros (AE), que seria uma nova versão não-estruturalista da AC que deixa de conceber de forma negativa as interferências da LM do aprendiz no processamento e produção da língua-alvo, por entender que tais interferências são mecanismos cognitivos acionados pela mente humana como um passo inicial para se adquirir uma língua (ALVES, 2011). Logo, entende-se na AE que há um

período transitório no processo de aprendizagem dentro do qual se manifesta a interlíngua, cujo conceito já foi mencionado em nota no capítulo 1. Pode-se dizer que a AE tenha contribuído para duas boas conquistas: a primeira é a de conferir ao aluno e à sua produção uma posição de destaque no ensino de línguas, e a segunda é a de evidenciar a existência de outros fatores – extralinguísticos e psicológicos, como a atitude do aprendiz mediante a comunidade falante da língua que está aprendendo – que podem ocasionar erros, além da interferência da LM (BERETTA, 2001). Não obstante essas conquistas, os dados dessa metodologia de análise, a AE, nem sempre “seguiram princípios de compilação e desenho rigoroso” (SHEPHERD, 2009, p. 101). Além disso, segundo Ferris (2004, apud SHEPHERD, 2009), o estudo de erros tinha como foco erros individuais, que eram estudados fora de seu contexto e cotexto¹⁹ originais, e não havia um consenso quanto à gravidade ou seriedade de um erro. Desse modo, questões como essas continuamente adquiriam um tratamento não apenas relativamente superficial, mas também subjetivo, e prova disso é que a própria língua materna do avaliador de um erro poderia exercer influência na decisão acerca de qual tipo de erro era qualificado como mais ou menos sério.

Mesmo com todos esses percalços, pode-se dizer com firme certeza que a AE está à frente da AC (ALVES, 2011, p. 31), uma vez que conseguiu abranger as discussões acerca de erros, ampliando o território de análises ao incluir outras fontes possíveis, além da transferência da LM, para a geração de erros, tais como o contexto sociolinguístico de comunicação em que os aprendizes estão inseridos, as estratégias cognitivas ou psicolinguísticas que eles utilizam e variáveis afetivas.

2.4.1 *Conceitos de erro*

Tradicionalmente na área de metodologias para o ensino de línguas estrangeiras, mais especificamente focando no ensino de inglês, duas correntes antagônicas para o tratamento com erros se opunham. De um lado, estudiosos como Ellis (1985), George (1972) e Ringbom (1986), citados por JIMÉNEZ CATALÁN, 1992, adotavam uma perspectiva um pouco mais rigorosa, mesmo que em graus de intensidade diferentes, ressaltando os aspectos negativos de erros, em geral tidos como indesejáveis ao longo do processo de aprendizagem. Tal visão de erro caminha emparelhada com a tradição behaviorista que prima pelo

¹⁹ Baseio-me em Partington (1998), Fox (1998), entre outros autores, que utilizam o termo cotexto para se referirem ao ambiente linguístico que acompanha o item linguístico em investigação. O termo contexto seria, então, utilizado para se referir ao ambiente situacional no qual o item é utilizado (FOX, 1998).

condicionamento, cujo impacto como teoria de aprendizagem é o de entender que o professor é um conhecedor mais proficiente da língua e, nesta posição, deve atuar realizando correção incisiva de todos os erros cometidos pelos aprendizes. Do outro lado, Corder (1967), Svartvirk (1973), Dulay, Burt & Krashen (1982), citados por JIMÉNEZ CATALÁN, 1992, defendem outro entendimento a respeito da natureza dos erros, alegando que a construção de hipóteses criativas por parte dos aprendizes os auxilia, ativando suas capacidades cognitivas. Sendo assim, eles enxergam os erros não só como necessários ao processo, mas até como positivos, por permitirem que os alunos aprendam através deles, e por eles mesmos, os erros, servirem de informantes aos professores e autores de materiais e livros didáticos quanto a quais são pontos da língua-alvo que geram mais dificuldades para os aprendizes e quais tipos de erros mais comprometem a comunicação. Já esta visão, diferente da explicitada anteriormente, equipara-se ao método comunicativo de ensino, que se fundamenta na cognição e por meio do qual o professor deve agir como um guia facilitador que seleciona o tipo de correção que deve ser feito, considerando o nível dos alunos, e ajuda os aprendizes a alcançar suas metas de maior proficiência linguística progressivamente.

Neste trabalho, a segunda visão acima exposta é adotada e, em consonância com a fala de Berber Sardinha (2004, p. 266), tem-se que a pesquisa em LC com a linguagem de não-nativos “reconhece a inevitabilidade de erros no processo de aquisição e se propõe a descrever tais desvios de modo mais neutro e substanciado, tendo como base coletâneas da produção de aprendizes.” Dessa forma, entende-se que o erro não deve ser avaliado, primeiramente, de forma a conduzir a julgamentos, e em segundo lugar, de forma individual e descontextualizada. Tal asserção está de acordo com o que Shepherd (2009a, p. 103) alega quando diz que

a ênfase dos estudos de corpora eletrônicos no léxico, sua frequência e sua co-ocorrência mudou o foco do estudo sobre a produção de aprendizes do individual para o coletivo, do erro atribuído a uma determinada ‘interferência’ da língua materna, para o uso real da língua, uso esse circunscrito a textos e gêneros textuais produzidos em condições semelhantes, além de, em contrapartida, focar nos acertos, também.

Mais ainda, é válido mencionar que, nos dias de hoje, com a presença poderosamente retumbante de tecnologias computacionais, é possível compilar *corpora* de milhões de palavras, que podem passar por processo de anotação²⁰ e ser etiquetados com o

²⁰ Para mais informações sobre anotação de *corpora*, recomenda-se consultar um bom guia de LC, como o de O’KEEFEE e MCCARTHY (2010), *The Routledge Handbook of Corpus Linguistics*.

propósito de servir a diferentes fins de pesquisa. Essa observação se faz relevante justamente para comprovar que os tipos de dados com os quais pesquisadores podem lidar na atualidade são diferentes daqueles que eram utilizados nos anos 80, quando da fase prolífera da AE. Nessa época, muitas coletâneas de dados não passavam de 2 mil palavras (BEBER SARDINHA, 2004), provenientes de poucos alunos e que eram compiladas de forma randômica, talvez sem uma proposta comunicativa estabelecida para a produção textual.

Dessa forma, ainda que o erro não fosse tratado, em seu conceito, como algo fundamentalmente negativo em sua essência dentro da AE, ele não era tratado como algo que poderia refletir uma característica de uma coletividade, mas recebia uma análise individual. Tal posicionamento não endossa ou não contribui com uma avaliação da língua como um marcador de identidade étnica ou racial, e como um marcador de perfil coletivo. Muito embora exista a noção de idioleto, que remete ao sistema linguístico por meio do qual as características pessoais de um indivíduo são refletidas, Fairclough (2003, apud SHEPHERD, 2009a, p. 103) aponta que “em suas interações sociais, os seres humanos compartilham modos de falar que os identificam como grupo, isto é, o compartilhamento do ‘como’ pode identificar um grupo coletivamente.”

A realização de estudos como este tem o objetivo de descrever, a partir de uma quantidade de dados significativamente mais robusta, análises acerca de unidades lexicais usadas com maior ou menor frequência por um determinado grupo, seja por influências de sua LM, ou mais provavelmente por uma junção de diversos fatores que vão além da linguística, como fatores sociais, históricos ideológicos e culturais. Desvios, então, daquilo que é previsto na língua-alvo não devem ser tratados negativamente e simplesmente como erros quaisquer, mas como inadequações que naturalmente fazem parte da interlíngua e que podem ser descritas e abordadas em materiais didáticos e em sala de aula de uma maneira a ativar a conscientização linguística dos aprendizes.

2.4.2 *O papel da conscientização linguística*

Schmidt e Frota (1986, apud DUTRA, 2009) alegam que há itens linguísticos, gramaticais ou sociolinguísticos, de percepção obscura por parte dos aprendizes, ou seja, que não são percebidos com facilidade. Por essa razão, os autores argumentam que é necessário promover oportunidades no ensino de línguas estrangeiras para *noticing*²¹, ou percepção consciente.

²¹ Termo cunhado por Schmidt e Frota (1986).

Tomlin & Villa, 1994 e Schmidt, 2001 (apud SCHMIDT, 2010) explicam que a percepção como atenção, seja ela intencional ou não, trata-se não apenas de um mecanismo único, mas de uma série de mecanismos ou subsistemas que incluem a capacidade de se manter alerta, de orientação, de detecção em atenção seletiva, de facilitação e de inibição. Todos esses fatores controlam o processamento de informações, porém ainda não há um consenso quanto à noção de que todo tipo de aprendizagem requer atenção propriamente dita. Os autores também dizem que ao mesmo tempo em que percepção e atenção estão intimamente ligadas, já que geralmente percebemos em decorrência de termos depositado nossa atenção em algo, não se pode afirmar que a percepção de regras abstratas de gramática seja um pré-requisito para a aprendizagem. Tendo em vista que nativos de uma determinada língua possuem entendimento intuitivo de pontos gramaticais sutis que eles mesmos não conseguem verbalizar, e que aprendizes de língua estrangeira em nível avançado também possuem tal conhecimento intuitivo, – próximo ao dos nativos – não se pode afirmar que as regras geralmente expostas de forma ‘simplificada’ em sala de aula sejam o suficiente para compor esse tipo de conhecimento implícito.

Ainda assim, Schmidt (2010) declara que a maior parte dos estudos empíricos comprovou que a percepção exerce um papel de grande importância na aprendizagem. Mackey (2006, apud SCHMIDT, 2010), por exemplo, utilizou medidas de percepção e desenvolvimento para investigar se fornecer *feedback* promove uma melhor percepção de formas da língua-alvo dentro de sala de aula e se existe uma relação entre relatórios de percepção feito pelos aprendizes e resultados em sua aprendizagem. No Brasil, alguns estudos já foram realizados com o objetivo de aferir o papel da percepção consciente, como o de Marques (2007). Essa autora implementou um experimento que visou verificar se a interlíngua dos participantes, falantes avançados da língua inglesa, mostraria alterações no nível de proficiência após atividades que procurassem ativar sua percepção consciente quanto ao uso de conectivos concessivos em sua produção escrita. Ela verificou que após a realização das atividades propostas, as quais serviram para intensificar a atenção dos participantes quanto às estruturas alvo, houve reformulação linguística por parte deles no sentido de um aumento expressivo do uso dos conectivos em sua produção, surtindo, então, o efeito desejado, que era o da complexificação do discurso na escrita de artigos. Desse modo, foi levada em consideração e empregada com êxito a sugestão de Schmidt (2001, apud MARQUES, 2007) de que professores devem maximizar oportunidades para que

determinadas estruturas sejam utilizadas em situações e contextos significativos de modo a conduzir os aprendizes a prestarem atenção em sua forma. Outro exemplo de trabalho realizado em prol da conscientização linguística é o de Dutra (2009). Tal trabalho, que inclusive agrega a utilização da LC como subsídio metodológico, e que foi conduzido ao longo de um curso semipresencial com foco em *corpora*, mostrou ter havido aumento da conscientização de aspectos linguísticos por parte dos participantes que “passaram a ter uma melhor compreensão de conceitos linguísticos, além de terem incorporado novos conceitos sobre o que são colocações, porções (*chunks*), expressões idiomáticas, materiais autênticos, entre outros.” (DUTRA, 2009, p. 95).

2.5 Panorama de trabalhos com *corpora* de aprendizes e sobre quantificadores em *corpora* e *corpora* de aprendizes

2.5.1 Um micro retrato dos estudos com corpora de aprendizes

Na obra seminal de Granger (1998), há a reunião de diversos trabalhos sobre *corpora* de aprendizes, e dentre eles há o de Meunier (1998 apud GRANGER, 1998), que apresenta e discute uma gama de ferramentas computacionais úteis para a análise de dados de aprendizes, e o de Aarts e Granger (1998 apud GRANGER, 1998), que analisaram sequências etiquetadas em *corpora* de aprendizes de três diferentes nacionalidades. Trabalhos como esse último, de Aarts e Granger (1998), revelam um viés investigativo usualmente adotado em pesquisas com *corpora* de aprendizes, que é o de detectar, em suas produções, **sobreuso** e **subuso** de palavras e/ou estruturas, em comparação com produções de nativos, resguardados os cuidados com o tipo e tamanhos dos dados sendo comparados. Em seu trabalho, Delegá-Lúcio (2006, p. 2) define que “o sobreuso de palavras diz respeito à utilização de uma mesma palavra mais vezes do que um falante nativo normalmente o faria”. Já o subuso é o oposto, isto é, um uso menos frequente de palavras que são comumente utilizadas pelos falantes nativos. Aarts e Granger (1998) detectaram, na produção dos aprendizes avaliados em sua pesquisa, sobreuso de sequências posicionadas em início de sentença compostas por conectivos, advérbios e auxiliares e pronomes, e subuso de substantivos, conjunções seguidas de substantivos e preposições seguidas de verbos no gerúndio.

Seguindo essa linha de investigação, o presente trabalho, que se ocupa da descrição da interlândia de aprendizes brasileiros de inglês em nível de intermediário a avançado, estuda o uso dos quantificadores *a few* e *few* com fins de averiguar possível

sobreuso ou subuso de agrupamentos ou colocações com esses quantificadores em particular na produção dos estudantes.

Ainda sobre estudos de consolidada importância que levam em conta *corpora* de aprendizes, vale citar o de Römer e Wulff (2010), que faz uma exploração de dados do MICUSP²², e ressalta o quão profundamente as ferramentas de *software* de tratamento de dados são úteis para acessar *corpora* e chegar a descrições de padrões de uso em uma língua que nunca seriam percebidos ‘a olhos nus’. Cumpre citar também a conferência recente sobre *corpus* de aprendizes, *Learner Corpus Research 2013*, sediada na Noruega, cujos resumos dos trabalhos apresentados podem ser encontrados *online* no Livro de Resumos²³ desse evento.

No Brasil, destacam-se trabalhos interessantes também, como o de Delegá-Lúcio (2006), que trata da relexicalização de adjetivos, de Pinto (2008), que analisa o uso de *things*, *thing*, *anything*, *something* e *everything*, de Dutra e Silero (2009), que investiga agrupamentos com a preposição *for* na escrita de aprendizes, de Almeida (2012), que faz um levantamento do uso de colocações com os verbos *get*, *give*, *make*, e *take*, e de Tenuta, Oliveira e Orfanó (2012), que trata da expressão de modalidade feita por aprendizes.

2.5.2 Trabalhos sobre quantificadores em corpora

Em relação a trabalhos realizados especificamente sobre quantificadores em *corpora*, eles não são tão numerosos na literatura da área, mas é possível citar alguns exemplos, como os de Wang e Piao (2007), Ruzaité (2009), AnderBois, Brasoveanu e Henderson (2014), e Tottie e Neukom-Hermann (2010). O primeiro deles analisou a tradução de quantificadores que expressam ideias vagas, como *many*, *some*, *a few* em um *corpus* paralelo de inglês-chinês; o segundo mostrou as diferentes características do sistema de quantificadores em inglês e em lituano utilizando, também, um *corpus* paralelo como base; o terceiro apresentou um estudo de *corpus* sobre a pragmática do escopo de quantificadores; e o último revelou traços típicos da interação entre negação e quantificação no inglês utilizando o *British National Corpus* (BNC) como referência.

²² *Michigan Corpus of Upper-Level Student Papers*, que possui um total de 2,6 milhões de palavras (RÖMER; WULFF, 2010).

²³ *Book of Abstracts*, encontrado no endereço eletrônico Disponível em: <<http://lcr2013.b.uib.no/files/2014/09/abstracts-book.pdf>>. Acesso em: 24 abr. 2014.

Se trabalhos sobre quantificadores usando a metodologia da LC já são mais raros, mais difícil ainda é encontrar tal temática em trabalhos especificamente realizados a partir de *corpora* de aprendizes. Todavia, Rundell e Granger (2007) apontam para a importância de se atentar aos quantificadores na produção de aprendizes. Ao explicarem sobre o projeto colaborativo que reuniu esforços dos autores do dicionário da editora *Macmillan* e dos pesquisadores do Centro Inglês de Linguística de *Corpus* da Universidade Católica de Louvain, eles narram que foram feitas comparações entre as produções escritas – pertencentes ao mesmo gênero textual – de aprendizes e de nativos com o objetivo de identificar diferenças sistemáticas entre os textos de ambas as fontes. Os autores esclarecem que, com isso, não estavam simplesmente procurando por erros dos aprendizes, mas buscando evidências de disparidades significativas nas estratégias de expressão usadas por cada grupo. Dentre os achados relatados, eles destacam duas grandes áreas de dificuldade: a primeira é a acuidade lexical e gramatical, e a segunda é a fluência na escrita. Dentro da primeira área citada, os dados pesquisados mostraram, dentre outros, problemas quanto à estruturação sintática, escolha lexical, e quanto ao uso de artigos e quantificadores (RUNDELL, GRANGER, 2007, p. 16). Sendo assim, faz-se de fato pertinente investigar os quantificadores, principalmente pela comprovação de que seu uso suscita dificuldades.

3 METODOLOGIA

De acordo com Peters e Westerstahl (2006), o estudo linguístico de quantificadores originalmente era mais focado na expressão gramatical da quantificação do que no seu sentido. Entendendo que a forma e o sentido caminham juntos, uma das maneiras de estudar o entrelaçamento léxico-gramatical com a produção de sentido é por meio da LC. Já é sabido que esta é uma área que está atrelada a uma série de procedimentos – ou métodos²⁴ – utilizados para o estudo da língua, procedimentos tais ainda em desenvolvimento, embora alguns deles, como as linhas de concordância, já estejam bem estabelecidos e sejam vistos como centrais para a abordagem com *corpus* (MCENERY; HARDIE, 2012).

O presente trabalho não será diferente, e terá como base análises feitas a partir da geração de linhas de concordância pelo *software* AntConc 3.2.1²⁵, procedimento que será demonstrado posteriormente, ainda no capítulo de metodologia. Por agora, cumpre iniciar o detalhamento das fontes de dados utilizados para essa pesquisa.

3.1 Dos corpora utilizados

Três *corpora* de tamanhos semelhantes serviram como fonte de dados para o presente estudo, a saber:

a) O *corpus* escrito em inglês por aprendizes brasileiros de nível intermediário a avançado, intitulado *Brazilian Subcorpus of the International Corpus of Learner English* (Br-ICLE), que possui pouco mais de 200.000 palavras (precisamente 206.075 palavras) de redações argumentativas de alunos de Letras, habilitação Inglês. Tais redações são provenientes de diferentes universidades que fizeram suas contribuições, como a Pontifícia Universidade Católica de São Paulo (PUC- SP), a Metodista de São Paulo, a Universidade Estadual do Rio de Janeiro (UERJ), a Universidade Federal de Minas Gerais (UFMG), entre outras. Este *corpus* foi compilado primariamente sob a supervisão do Professor Tony Berber Sardinha, da PUC-SP, além de outros supervisores colaboradores²⁶. Cada redação possui uma

²⁴ Biber *et al.* (1998, apud SHEPHERD, 2009b, p.152) “confirmam o status de metodologia da LC dizendo que uma abordagem que parte do *corpus* pode ser aplicada a praticamente qualquer área de investigação linguística, algo semelhante ao que McEnery e Wilson (1996) alegam quando dizem ser possível atrelar a expressão ‘de corpus’ a qualquer ramo da Linguística.”

²⁵ Versões mais atualizadas desse software estão disponíveis para download para diferentes sistemas operacionais, disponível em: < http://www.antlab.sci.waseda.ac.jp/antconc_index.html>. Acesso em: 24 abr. 2014.

²⁶ Dentre os outros colaboradores de supervisão do Br-ICLE estão a Professora Deise Prina Dutra (UFMG) e a Professora Tania Shepherd (UERJ).

média de 500 a 1.000 palavras, obedecendo aos critérios estabelecidos pelo *International Corpus of Learner English* (ICLE). Faz-se necessário, então, esclarecimento sobre o ICLE, que se trata de um *corpus* de 3,7 milhões de palavras (GRANGER *et al.*, 2009) que recebeu contribuições de textos argumentativos produzidos por aprendizes de inglês com nível avançado de diferentes origens, como russos, finlandeses, franceses, alemães, italianos, japoneses, holandeses, chineses, entre outros. O ICLE possui critérios fixos que devem ser obedecidos, como o tamanho dos textos – mínimo de 500 palavras e máximo de 1.000, – e o tipo de textos – em sua maioria, redações argumentativas. Segundo Granger *et al.* (2009), 91 % dos textos do ICLE são redações desse tipo textual, e 13 sugestões²⁷ de tópicos foram dadas aos participantes, como listadas a seguir:

(1) *Crime does not pay*

(2) *The prison system is outdated. No civilised society should punish its criminals: it should rehabilitate them*

(3) *Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value*

(4) *A man/woman's financial reward should be commensurate with their contribution to the society they live in*

(5) *The role of censorship in Western society*

(6) *Marx once said that religion was the opium of the masses. If he was alive at the end of the 20th century, he would replace religion with television*

(7) *All armies should consist entirely of professional soldiers: there is no value in a system of military service*

(8) *The Gulf War has shown us that it is still a great thing to fight for one's country*

(9) *Feminists have done more harm to the cause of women than good.*

(10) *In his novel Animal Farm, George Orwell wrote "All men are equal: but some are more equal than others" How true is this today?*

(11) *In the words of the old song "Money is the root of all evil"*

(12) *In the 19th century, Victor Hugo said : "How sad it is to think that nature is calling out but humanity refuses to pay heed. "Do you think it is still true nowadays?"*

²⁷ Elas foram retiradas do *site* oficial do ICLE, (GRANGER, 2014).

(13) *Some people say that in our modern world, dominated by science technology and industrialisation, there is no longer a place for dreaming and imagination. What is your opinion?*

Vale dizer que, embora já tenha atingido o número mínimo de palavras que é exigido – 200.000 – para se incorporar ao ICLE, o Br-ICLE ainda está em processo final de organização e, portanto, ainda não faz parte oficialmente do ICLE. Outro fator importante a ser mencionado é que todos os participantes do ICLE (e Br-ICLE) tiveram que preencher um formulário chamado *Learner Profile* – LP (ANEXO E)²⁸, no qual são solicitadas dos participantes informações de gênero, idade, escolaridade, país de origem/língua materna, entre outras. Além de satisfazer um propósito ético, tendo em vista que no final do LP o estudante assina o nome, dando sua permissão para que a redação seja utilizada para fins de pesquisa, os LPs são documentos que servem para a contabilização das características dos informantes de cada *subcorpus* montado, já que pode haver pesquisas para as quais tais informações se fazem relevantes. Logo, ao serem convidados a participar da compilação do Br-ICLE, além das diretrizes dadas quanto aos critérios já mencionados (tamanho, tópicos), foi pedido que os estudantes preenchessem o LP, e eles também foram informados de que suas redações receberiam um número e, assim sendo, a identidade de cada um ficaria inteiramente preservada. Mais detalhes sobre o ICLE podem ser encontrados por meio do *link* (GRANGER, 2014) como citado acima, e sobre o Br-ICLE, por meio do *link* (BR-ICLE, 2014).

b) Uma parte do *corpus* escrito em inglês por nativos, intitulado *Louvain Corpus of Native English Essays* (LOCNESS) (GRANGER; SANDERS; CONNOR, 2014), cuja compilação se deu com o objetivo de ele se tornar o equivalente nativo do ICLE. A parte usada nesse trabalho é composta pelas redações argumentativas e não-literárias de alunos universitários americanos e britânicos, perfazendo um total de 209.783 palavras. Por completo, o LOCNESS possui 324.304 palavras, e é assim subdividido: 60.209 palavras de redações argumentativas de alunos britânicos com conceito A (alguns dos tópicos são “transporte” e “o sistema parlamentar”); 95.695 palavras de redações literárias e não-literárias de estudantes universitários britânicos (um dos tópicos foi “uma Europa única: a perda da soberania para a Grã-Bretanha”); e 168.400 palavras de redações argumentativas literárias e não-literárias de estudantes americanos de universidades como a de Michigan, da Carolina do Sul, de Indianópolis, entre outras (alguns dos vários tópicos são “o dinheiro como raiz de todo

²⁸ Em português, chamamos esse formulário de Termo de Consentimento, o qual foi aprovado no Comitê de Ética da UFMG através de projeto registrado pela minha orientadora

o mal”, “a igualdade dos sexos”, e “a legalização da maconha”). É possível ter acesso ao LOCNESS por meio do *link*: <http://www.learnercorpusassociation.org/resources/corpora/locness-corpus/> Neste trabalho, as redações literárias de ambos os estudantes americanos e britânicos não foram incluídas. Dessa maneira, como dito acima, apenas uma parte do LOCNESS foi utilizada, e favoravelmente o tamanho selecionado (209.783 palavras) se assemelhou mais ao tamanho total do Br-ICLE (206.075 palavras).

c) O *corpus* escrito em português, de nativos do português brasileiro, aqui nesse trabalho intitulado Co-RAP (Corpus de Redações Acadêmicas em Português), o qual foi obtido por meio da permissão de acesso às redações escritas por alunos de diversos cursos como de Engenharia, Administração, Matemática, Ciência da Computação, entre outros, oferecidos pela UFMG. Tais alunos cursaram a disciplina Escrita Acadêmica – oferecida pela Faculdade de Letras da UFMG – como disciplina eletiva, e as redações estão disponibilizadas no site: <http://textolivre.pro.br/blog/?p=735> (MATTE; COSCARELLI, 2013), organizado pelas Professoras Ana Cristina Fricke Matte e Carla Coscarelli, que ministravam as aulas, em sua maioria não-presenciais. Vale dizer que as redações faziam parte do quadro avaliativo da disciplina, não tinham limite de palavras, e possuíam tópicos pré-definidos, como “educação à distância”, “o papel da internet em períodos de eleição”, e “o conhecimento e uso de *softwares* livres”. Foram selecionadas aleatoriamente 122 redações que possuem uma média de 1.000 a 1.100 palavras, e o *corpus* possui um total de 153.766 palavras.

Além dos três *corpora* de estudo já detalhados, faz-se necessária a inclusão do seguinte item na listagem de fontes de dados, que informa acerca dos *corpora* de referência utilizados:

d) Para o enriquecimento e complementação das análises com informações adicionais, principalmente sobre colocados das expressões quantificadoras pesquisadas, a seção acadêmica do *corpus* geral disponibilizado na plataforma criada por Mark Davies²⁹ *Corpus of Contemporary American English* (COCA) e o Corpus do Português, também da plataforma de Davies – foram utilizados. Em termos de tamanho, a seção acadêmica do COCA possui 81 milhões de palavras e o Corpus do Português possui 45 milhões de palavras. Ambos serviram de *corpora* de referência, e embora esse trabalho esteja voltado para buscas mais pontuais nas seções acadêmicas, é válido mencionar que a plataforma de Davies para esses e outros *corpora*, como o *Corpus of Historical American English* (COHA) e o *British*

²⁹ Linguista e professor da Brigham Young University (BYU) no estado e Utah (EUA).

National Corpus (BNC), é riquíssima e oferece não só o acadêmico, como também outros registros (ficção, revista, jornal, conversação), além de inúmeros recursos (buscas por frequência, relevância, buscas com etiquetas, buscas de colocados, resultados de frequência graficamente representados, etc.) de extrema utilidade metodológica para pesquisas com *corpus*. Particularmente nesse trabalho, a obtenção de quadros com listas de colocados mais frequentes em relação a um nóculo foi o recurso mais utilizado.

3.2 Das análises realizadas

Como um primeiro comentário, faz-se importante esclarecer que as análises e comparações feitas entre os *corpora* são justas por três principais características de sua arquitetura, quais sejam: i. o tamanho deles, que é equiparado, o que reduz a possibilidade de haver margem para discrepâncias na quantificação dos dados; ii. o fato de se tratarem de *corpora* da mesma modalidade, a escrita; e iii. o fato de eles serem *corpora* do mesmo registro, o acadêmico, e terem sido compilados sob critérios e condições semelhantes. Outra observação que merece atenção se trata do cuidado mantido ao longo das análises das linhas de concordância com os quantificadores *a few* e *few*, as quais não foram provenientes de apenas um aprendiz, o que garante a eliminação de conclusões tendenciosas, baseadas em apenas um informante como parâmetro.

Em se tratando dos tipos de análises feitas, tem-se a seguinte subdivisão:

3.2.1 O apanhado geral

As análises transcorreram da seguinte forma: em primeiro lugar, analisando as linhas de concordância do Br-ICLE com *few*, foi feito um apanhado qualitativo geral com os dois seguintes objetivos: o primeiro foi separar as ocorrências que possuíam expressões de tempo – considerando que são numerosas (como ‘*in a few decades*’, ‘*for a few weeks*’, ‘*over the past few years*’, etc.) – seguindo o nóculo em questão a fim de verificar o uso dos colocados à direita relacionados a tempo e fazer uma comparação com a seção acadêmica do COCA (gerando o quadro de colocados mais frequentes à direita), e também com a parte do LOCNESS (após separar as linhas manualmente, conforme feito com o Br-ICLE); e o segundo foi investigar os colocados à esquerda de *few*, além do artigo *a*, na seção acadêmica do COCA com o intuito de obter um parâmetro vindo de um *corpus* de referência e averiguar, comparativamente, os resultados que seriam gerados desses colocados para o Br-ICLE e também para o LOCNESS.

3.2.2 O apanhado com nuances de prosódia semântica

Tendo em vista que um dos objetivos desse trabalho era verificar se há alguma relação entre o uso de quantificadores e seus colocados com a manifestação de uma **prosódia semântica** característica na expressividade do discurso acadêmico, foi necessário fazer uma checagem de como os nativos fizeram isso no LOCNESS a fim de comparar com o que sucedeu, também nessa direção de análise, no Br-ICLE.

Para tanto, foram separadas manualmente, e observando o contexto expandido, as linhas de concordância que satisfaziam esse propósito de pesquisa para tentar buscar por um padrão seguido pelos nativos e identificar as inadequações cometidas pelos aprendizes. Nesse momento, foi necessário contar com o respaldo estatístico do *Fisher's Exact Test*,³⁰ para uma melhor interpretação quantitativa dos dados no sentido de testar sua significância estatística e obter a validação devida dos resultados.

Sabe-se que a maioria do que se deseja medir geralmente está sujeita a um grau de flutuação randômica, e por isso é recomendado se valer de testes de significância a fim verificar se um determinado resultado é uma mera coincidência ou se é fruto de um fenômeno significativo, a partir do qual vale a pena investir em mais pesquisas.

Dunning (1993) destaca diferentes testes que podem ser usados por pesquisadores, dependendo do tipo de dados que possuem, como o teste chi-quadrado, o teste t e o teste *log-likelihood*. Para o presente trabalho, como já dito, nenhum desses três testes foi usado, devido ao tamanho dos *corpora* sendo comparados. Cumpre informar que o teste chi-quadrado, por exemplo, costuma presumir uma distribuição normal para os dados, e tal distribuição não é típica para dados linguísticos. Enquanto a distribuição normal é graficamente representada por uma curva senoidal, o gráfico que é comumente plotado para a distribuição de palavras é o de uma curva decrescente, revelando a característica de haver poucas palavras – em sua grande maioria, as palavras gramaticais – que são muito frequentes, um número significativo de palavras de frequência média – em geral, as palavras funcionais, – e muitas palavras de frequência baixa, das quais as várias listadas nos diversos últimos lugares do ranqueamento por frequência são as denominadas *hapax legomena* (as que possuem apenas uma ocorrência no *corpus*).

³⁰ Teste de significância estatística geralmente aplicado em amostras de tamanho pequeno. Em português, Teste Exato de Fisher. Há páginas online que disponibilizam a aplicação desse teste, e o endereço visitado quando da realização do teste foi <http://www.physics.csbsju.edu/stats/fisher.form.html>.

Assim, pode-se dizer que o chavão “poucos gigantes e muitos anões” reflete essa configuração, que é, inclusive, estatisticamente prevista pela Lei de Zipf³¹, e na sessão de Anexos (ANEXO F) há um exemplo de um gráfico com a curva típica para a distribuição de palavras em um *corpus*. Desse modo, por presumir uma distribuição normal, e também por ser recomendada a sua aplicação em *corpora* maiores, o teste chi-quadrado, bem como os demais mencionados no parágrafo anterior, não seriam indicados para aferir quantitativamente a significância do fenômeno da **prosódia semântica** sendo aqui tratado. Para evitar, portanto, uma interpretação distorcida ou tendenciosa dos dados, a melhor opção de teste estatístico a ser empregado recaiu sobre o Teste Exato de Fisher, que obtém preferência no caso de amostras pequenas. Mais ainda, nesse momento, o Co-RAP também se fez relevante para informar acerca de características dos agrupamentos com quantificadores produzidos pelos nativos de português que também possuíam nuances de **prosódia semântica**. A partir disso, então, tornou-se possível adquirir uma noção, ainda que um pouco restrita, da extensão da influência da interlíngua na escrita de aprendizes avançados de inglês, e lançar uma discussão a esse respeito.

3.2.3 *O detalhamento dos dados dos aprendizes*

Após os passos descritos nos dois itens acima, partiu-se para a explanação mais individualizada das ocorrências com *few* provenientes dos dados dos aprendizes. Assim sendo, foram reunidas em quadros todas as linhas de concordância que estavam adequadas e todas as que não estavam.

A partir daí, foram feitos comentários gerais sobre os acertos dos aprendizes e, para o tratamento com as inadequações, foram estabelecidas cinco categorias de erros. Feito isso, foi possível organizar melhor os tipos de inadequações, agrupando-as dentro das categorias e detalhando-as mais criteriosamente.

Nesse capítulo, a arquitetura de cada *corpus* utilizado como fonte de dados para essa pesquisa foi descrita, bem como os procedimentos metodológicos empregados ao longo do trabalho a fim de alcançar os objetivos traçados e obter respostas para as perguntas de pesquisa delineadas no capítulo 1. No capítulo seguinte, os resultados das análises conduzidas serão apresentados e discutidos.

³¹ Mais detalhes sobre a descrição dessa lei podem ser encontrados na obra de George Zipf: *Human Behavior and the Principle of Least Effort* (1949).

4 RESULTADOS

O capítulo de resultados terá a seguinte subdivisão: i. A primeira seção apresentará um apanhado geral de análises do *corpus* de aprendizes Br-ICLE; ii. A segunda abordará o fenômeno da **prosódia semântica**, mostrando resultados e conclusões acerca da manifestação de tal fenômeno nos *corpora* escritos – o Br-ICLE e a parte do LOCNESS – que serviram de fonte para esse trabalho; iii. A terceira seção lidará com a discussão detalhada das linhas de concordância oriundas das produções dos aprendizes que contribuíram para o Br-ICLE quanto ao uso adequado e inadequado dos quantificadores em questão; iv. A quarta seção será dedicada às considerações sobre as implicações dos resultados; e v. A quinta e última seção trará um modelo de atividade sobre padrões e quantificadores, elaborada segundo a proposta da AMD – **Aprendizagem Movidada por Dados**.

4.1 Análises: o apanhado geral

O total de ocorrências no Br-ICLE com *few* é de 71, sendo que 26 delas (36,6%) são com *few* precedido do artigo indefinido *a*, isto é, *a few*, e o restante, 45 (63,4%), com *few*, conforme segue abaixo:

Quadro 1 – Porcentagem de ocorrências com *a few* e *few* no Br-ICLE

Total de ocorrências com <i>few</i> no Br-ICLE	71 (corresponde a 100%)
Ocorrências com <i>a few</i>	36,6% (26)
Ocorrências com <i>few</i>	63,4% (45)

Fonte: Elaborado pela autora/Adaptado: Br-ICLE.

4.1.1 Os colocados à direita

Investigando, inicialmente, os colocados à direita, os dados mostram que 22% (10 das 45 ocorrências com *few*) têm como colocados à direita palavras/expressões temporais, tais como *years*, *centuries* e *decades*. Para *a few*, tem-se que 46% (12 de 26 ocorrências) carregam colocados semelhantes, que também incluem *days*, *moments*, *minutes*. (Os quadros com as linhas de concordância especificamente com colocados de sentido temporal se encontram na seção de ANEXOS - G e H) Contudo, é necessário mencionar a existência dos outros tipos de colocados à direita, em sua maioria, relacionados a pessoas, como *people* e *women*. Há, ainda,

colocados que representam palavras referenciais, como *topics* e *questions*, e outros pertencentes a domínios semânticos distintos, como *wars* e *dreams*. O quadro a seguir mostra as porcentagens dos colocados à direita de *a few* e de *few*, de acordo com os domínios semânticos descritos:

Quadro 2 – Porcentagens no Br-ICLE dos colocados à direita de *a few* e de *few*, separados por domínio semântico

Colocados à direita no Br-ICLE	<i>a few</i> + ____	<i>few</i> + ____
Domínio semântico de tempo (<i>years, centuries, days, minutes, etc.</i>)	46,15% (12)	22,22% (10)
Domínio semântico de pessoas (<i>people, women, students, owners, etc.</i>)	23,09% (6)	44,44% (20)
Outros domínios semânticos (<i>examples, courses, universities, opportunities, etc.</i>)	30,76% (8)	33,34% (15)

Fonte: Elaborado pela autora/Adaptado: Br-ICLE

Parte-se, agora, para uma comparação com o LOCNESS e com o outro *corpus* de referência adotado neste trabalho, qual seja a seção acadêmica do COCA. No LOCNESS, as ocorrências com *few* perfazem o total de 100 (46 com *a few*, e 54 com *few*). 39% – 18 de 46 – das ocorrências com *a few* possuem os mesmos colocados relacionados a tempo à direita, como verificado no Br-ICLE, e 27% – 15 de 54 – das ocorrências com *few* apresentam, também, o mesmo tipo de colocados em semelhança ao Br-ICLE. Uma diferença, contudo, que pode ser destacada é que há, no LOCNESS, menos colocados à direita de *few* relacionados a pessoas (ver quadro 3 abaixo). Enquanto no Br-ICLE esses colocados chegam a quase metade do número de ocorrências com *few*, no LOCNESS eles compõem pouco mais de 1/5 do número de ocorrências com esse mesmo nóculo.

Quadro 3 – Porcentagens na parte do LOCNESS dos colocados à direita de *a few* e de *few*, separados por domínio semântico

Colocados à direita na parte do LOCNESS	<i>a few</i> + ____	<i>few</i> + ____
Domínio semântico de tempo (<i>years, decades, months, weeks, etc.</i>)	39,13% (18)	27,77% (15)
Domínio semântico de pessoas (<i>people, women, boxers, police, etc.</i>)	17,39% (8)	27,77% (15)
Outros domínios semânticos (<i>facilities, cars, suggestions, mathematical puzzles, calls,</i>	43,48% (20)	44,46% (24)

Fonte: Elaborado pela autora/ Adaptado: LOCNESS.

Mesmo com essas diferenças de porcentagem dos colocados em cada domínio semântico, não se pode afirmar que, distributivamente, a proporção dos colocados utilizados pelos aprendizes esteja muito distante do que é verificado no *corpus* dos falantes nativos. De igual modo, ainda não é possível declarar que houve, por parte dos aprendizes, uma tendência clara a **sobreuso** ou **subuso**³² de algum determinado agrupamento ou colocação com os quantificadores em estudo.

Já tomando o COCA acadêmico como base, foi feita uma busca dos colocados à direita tanto de *few*, como de *a few*, e os quadros 4 e 5 mostram os resultados:

Quadro 4 – 12 primeiros colocados à direita de *few* na seção acadêmica do COCA

	■	CONTEXT	FREQ		ALL	%	MI
1	■	YEARS	2695		79093	3.41	6.61
2	■	DAYS	932		16280	5.72	7.36
3	■	MONTHS	799		13857	5.77	7.37
4	■	STUDIES	712		50343	1.41	5.34
5	■	WEEKS	542		7062	7.67	7.78
6	■	DECADES	505		9426	5.36	7.26
7	■	MINUTES	498		6250	7.97	7.83
8	■	PEOPLE	484		90397	0.54	3.94
9	■	EXCEPTIONS	335		1948	17.20	8.94
10	■	HOURS	318		10114	3.14	6.49
11	■	HUNDRED	298		5330	5.59	7.32
12	■	EXAMPLES	265		10893	2.43	6.12

Fonte: DAVIES, 2014.

³² Conceitos já apresentados no capítulo 2.

Quadro 5 – 12 primeiros colocados à direita de *a few* na seção acadêmica do COCA

	■	CONTEXT	FREQ		ALL	%	MI
1	▣	YEARS	1455		79093	1.84	6.67
2	▣	DAYS	766		16280	4.71	8.02
3	▣	MONTHS	614		13857	4.43	7.94
4	▣	MINUTES	444		6250	7.10	8.62
5	▣	WEEKS	438		7062	6.20	8.42
6	▣	HOURS	283		10114	2.80	7.28
7	▣	HUNDRED	271		5330	5.08	8.14
8	▣	EXAMPLES	203		10893	1.86	6.69
9	▣	STUDIES	177		50343	0.35	4.28
10	▣	CASES	150		25331	0.59	5.03
11	▣	WORDS	138		24188	0.57	4.98
12	▣	MILES	129		4844	2.66	7.20

Fonte: DAVIES, 2014.

Como se pode atestar por meio dos quadros, as palavras/expressões temporais são, de fato, muito frequentes no uso que os nativos de língua inglesa fazem dos quantificadores em estudo. Vale reparar também os altos valores da **medida de informação mútua** (*MI score* – *Mutual Information Score*)³³, que mostram o quão recorrentemente associadas essas palavras aparecem no discurso dos falantes nativos.

Observando, portanto, as linhas de concordância dos textos produzidos pelos aprendizes, foi possível perceber que os mesmos parecem ir ao encontro do que se espera nesse tocante, isto é, eles também fazem uso amplo dos quantificadores com palavras/expressões relacionadas a tempo.

4.1.2 Os colocados à esquerda

Especificamente, agora, em relação aos colocados à esquerda, tomou-se a decisão de iniciar a pesquisa pelos dados de referência do COCA acadêmico. Logo, foi feita uma busca de colocados nesse *corpus*, gerando o quadro 6 a seguir:

³³ Medida que informa o quanto os itens lexicais se atraem mutuamente. Segundo Anderson e Corbett (2009), é comumente aceito que um valor de MI maior do que 3 (>3) sugere uma ligação forte entre o nóculo pesquisado e o seu colocado.

Quadro 6 – 12 primeiros colocados à esquerda de *few* na seção acadêmica do COCA, destacando-se *very*, *past* e *last*

	CONTEXT	FREQ	ALL	%	MI
1	A	16450	1747801	0.94	4.75
2	VERY	1410	54143	2.60	6.22
3	PAST	676	29904	2.26	6.02
4	LAST	572	23729	2.41	6.11
5	RELATIVELY	556	13935	3.99	6.84
6	NEXT	504	19318	2.61	6.22
7	FIRST	415	97216	0.43	3.61
8	TOO	298	29127	1.02	4.87
9	EVERY	155	26226	0.59	4.08
10	SELECT	55	3715	1.48	5.41
11	SURPRISINGLY	30	2429	1.24	5.14
12	PRIVILEGED	21	1560	1.35	5.27

Fonte: DAVIES, 2014.

Até o quarto colocado listado, os aprendizes parecem manter a tendência de uso dos nativos, embora haja no Br-ICLE uma ocorrência a mais com *past few* em comparação com *very few*, e o quadro acima mostra que *very few* é mais frequente. Os quadros 7 e 8 trazem as linhas de concordância com as 4 ocorrências de *past few* e as 3 de *very few* no Br-ICLE:

Quadro 7 – Ocorrências de *past few* no Br-ICLE

Concordance	Concordance Plot	File View	Clusters	Collocates	Word List	Keyword List
Hit	KWIC					File
1	le to realize by studies on past few years is the increase of cancer on p					BrICLEall12009.txt
2	a has been altered over the past few years and now religion is not societ					BrICLEall12009.txt
3	of poor responses over the past few years. In summary, the relationship					cabrill1088.txt
4	a high tech world> Over the past few decades, technology and the science					cabrill1090.txt

Fonte: Gerado pelo programa AntConc 3.2.1.

Quadro 8 – Ocorrências de *very few* no Br-ICLE

Concordance	Concordance Plot	File View	Clusters	Collocates	Word List	Keyword List
Hit	KWIC					File
1	ever, this is a challenge which very few people are able to overcome on t					BrICLEall12009.txt
2	t people would treat seriously. Very few women would have their lives apa					BrICLEall12009.txt
3	cerns about this limitation are very few . Why? Because we are not aware c					cabrill1066.txt

Fonte: Gerado pelo programa AntConc 3.2.1.

Dando seguimento às análises, o LOCNESS também apresenta colocados à esquerda em concordância com a lista gerada na consulta ao COCA acadêmico (ver quadro 6 acima) e, proporcionalmente, pode-se afirmar que o Br-ICLE se assemelha ao LOCNESS quando se compara o uso de *past few* e de *last few*. Na parte do LOCNESS, totalizando 9

ocorrências, *past few* ocorre três vezes mais do que *last few*, que possui, por sua vez, 3 ocorrências. No Br-ICLE, com 4 ocorrências no total (ver quadro 7 acima), *past few* ocorre o dobro de vezes em comparação a *last few*, como está demonstrado no gráfico 1 a seguir:

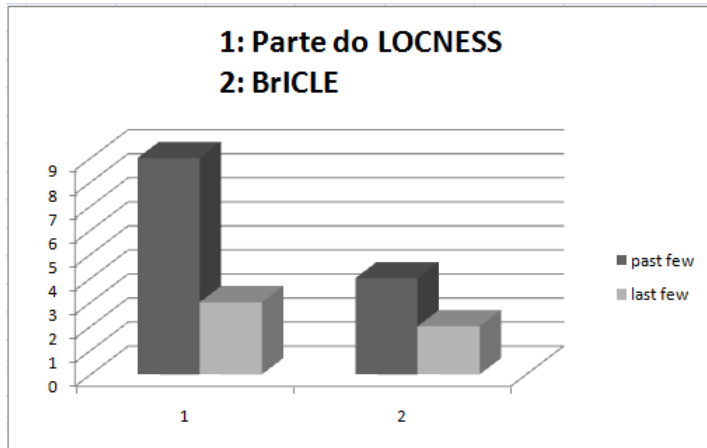


Gráfico 1 – Proporção das ocorrências de *past few* e *last few* na parte do LOCNESS e no Br-ICLE
Fonte: Elaborado pela autora, 2014.

Passando para o vocábulo *very*, que é o segundo colocado à esquerda de *few* mais frequente de acordo com a lista gerada a partir do COCA acadêmico (ver quadro 6), as linhas de concordância mostram que nem na parte do LOCNESS e nem no Br-ICLE há colocados à direita de *very few* relacionados a tempo.

Quadro 9 – Ocorrências de *very few* no Br-ICLE, com os colocados à direita em destaque

Hit	KWIC	File
1	ever, this is a challenge which very few <u>people</u> are able to overcome on t	BrICLEall12009.txt
2	t people would treat seriously. Very few <u>women</u> would have their lives spa	BrICLEall12009.txt
3	cerns about this limitation are very few . Why? Because we are not aware c	cabrill1066.txt

Fonte: Gerado pelo programa AntConc 3.2.1.

Quadro 10 – Ocorrências de *very few* na parte do LOCNESS, com os colocados à direita em destaque

Hit	KWIC	File
1	s was introduced but is used by very few <u>cars</u> , amounting to another failure. Alo	alevels1.txt
2	ajor issue to most citizens but very few of them are prepared to give up their c	alevels1.txt
3	ns can become very cramped with very few <u>facilities</u> and often dirty carriages. T	alevels1.txt
4	unt. In a time when Britain has very few <u>surviving wild animals</u> surely something	alevels3.txt
5	cted over a year ago there were very few <u>groups</u> which were against it. The gener	alevels5.txt
6	ossible by genetic advancements, very few <u>people</u> would argue that there is a mora	alevels8.txt
7	cts of genetic manipulation are very few in number, and most people need not con	alevels8.txt
8	would certainly apply. However, very few <u>scientists</u> would allow their "creation"	alevels8.txt
9	omeone from ever working again. Very few <u>criminals</u> have worker's compensation. T	USARG.txt
10	tragedy is that there have been very few <u>efforts</u> to do so. The lack of sight fr	USARG.txt
11	atisfied with their lives. The very few <u>women</u> who went to college ended up drop	USARG.txt
12	ncerning the matter. There are very few <u>legal objections</u> to this form of euthan	USARG.txt
13	zens do not need handguns since very few use them to defend themselves. More co	USARG.txt

Fonte: Gerado pelo programa AntConc 3.2.1.

Tal constatação, isto é, de não haver colocados à direita de *very few* relacionados a tempo, e de uma parte considerável desses colocados estar ligada a pessoas, também está de acordo com a lista de colocados à direita de *very few* gerada a partir de consulta à seção acadêmica do COCA, conforme o quadro 11 a seguir, que mostra itens lexicais como *people, women, children, teachers*:

Quadro 11 – 12 primeiros colocados à direita de *very few* na seção acadêmica do COCA

	■	CONTEXT	FREQ		ALL	%	MI
1	■	PEOPLE	65		90397	0.07	5.54
2	■	STUDIES	50		50343	0.10	6.00
3	■	EXCEPTIONS	25		1948	1.28	9.70
4	■	WOMEN	19		63266	0.03	4.28
5	■	CASES	16		25331	0.06	5.35
6	■	CHILDREN	10		76375	0.01	3.08
7	■	EXAMPLES	8		10893	0.07	5.57
8	■	SCHOOLS	8		31298	0.03	4.05
9	■	TEACHERS	8		56792	0.01	3.19
10	■	INSTANCES	7		3522	0.20	7.00
11	■	RESPONDENTS	7		12846	0.05	5.14
12	■	PUBLISHED	7		13845	0.05	5.03

Fonte: DAVIES, 2014.

Nos *corpora* de estudo, tem-se que há 13 ocorrências de *very few* na parte do LOCNESS e apenas 3 no Br-ICLE, ou seja, *very few* ocorre 4 vezes mais no LOCNESS, como mostra o gráfico 2 a seguir, e isso pode ser desse modo afirmado pelo tamanho semelhante desses *corpora* – o Br-ICLE com 206.075 palavras, e a parte do LOCNESS com 209.783.



Gráfico 2 – Ocorrências de *very few* na parte do LOCNESS e no Br-ICLE
Fonte: Elaborado pela autora, 2014.

Em primeira conclusão, é possível afirmar que os aprendizes parecem demonstrar resistência ao uso de *very* como um intensificador e colocado à esquerda de *few*, apontando, possivelmente, para um **subuso** da colocação *very few*, a qual é consideravelmente usada pelos nativos, como comprova o quadro 6. Tal alegação será melhor discutida na seção que aborda a **prosódia semântica**.

4.1.3 Os padrões léxico-gramaticais relacionados a tempo

Faz-se pertinente reparar a formação de padrões léxico-gramaticais temporais recorrentes a partir da junção do nódulo, nesse caso o quantificador *few*, com os colocados à esquerda e à direita. A recorrência desses padrões pode ser evidenciada numericamente, já que ao tomarmos *past / last few* como exemplo, todos os colocados à direita dessas combinações de palavras, tanto na parte do LOCNESS quanto no Br-ICLE, são palavras temporais.

Para melhor visualização desses padrões observados, eles podem ser assim transcritos e exemplificados:

a) Os padrões observados compostos de 4 palavras são:

- (preposição) + (artigo indefinido) + (quantificador) + (substantivo), como *in a few decades* e *for a few moments*
- (artigo indefinido) + (quantificador) + (substantivo) + (advérbio), como *a few years ago*

b) O padrão observado composto de 5 palavras é:

- (preposição) + (artigo definido) + (advérbio) + (quantificador) + (substantivo), como *over the last few years*, *over the past few months* e *in the past few years*

Fornecendo, ainda, mais detalhes sobre esses padrões, do total de 13 ocorrências de *past few* nos dois *corpora*, 10 são de *past few years* (4 delas formando *in the past few years*, e outras 4 formando *over the past few years*). Para *last few*, do total de 5 ocorrências nos dois *corpora*, 2 formam a colocação *over the last few years*, e outras 2 formam *in the last few centuries* e *in the last few years*.

4.2 Análises: a interface com a prosódia semântica

No grupo de linhas de concordância, (QUAD. 12), nas quais não há colocados relacionados a tempo, é possível, por meio de averiguação do contexto expandido, destacar algumas ocorrências em que se verifica a manifestação do fenômeno da **prosódia semântica**, especificamente aqui da **prosódia semântica negativa**. Tomando a parte do LOCNESS como fonte de dados nesse momento, a fim de realizar uma descrição do que acontece no discurso dos falantes nativos, eis alguns exemplos disso:

Quadro 12 – Grupo de linhas de concordância provenientes da parte do LOCNESS que apresentam **prosódia semântica** negativa

1. Along with the traffic congestion, comes pollution. The green house effect is a major issue to most citizens but very **few** of them are prepared to give up their car to show
2. reliable service even developing countries. **Few** efforts have been made to make the fares cheaper and a more desirable form of family travel, with the Family railcard
3. dissatisfied with their lives. The very **few** women who went to college ended up dropping out to get married for fear of too much education being a bar to marriage
4. The saddest part of this tragedy is that there have been very **few** efforts to do so. The lack of sight from our leaders is almost appalling. With the lack of respect for
5. an opportunity. There are **few** alternatives for low-income youths, and fewer obstacles in the quest for cash. With violent crime rising and financial support for more police
6. Again at peak times, the trains can become very cramped with very **few** facilities and often dirty carriages. This isn't the fault of British rail but the last passenger on
7. walking is faster at peak times of the day (ie 08:00 and 17:00). Resulting from this failure, a bypass was introduced but is used by very **few** cars, amounting to another failure.
8. In a time when Britain has very **few** surviving wild animals surely something should be done to protect the fox. The British government is always among the first countries to
9. Law-abiding citizens do not need handguns since very **few** use them to defend themselves. More commonly, Guns lead to violence within families or cause tragic accidents when

Fonte: Elaborado pela autora, por meio do programa AntConc 3.2.1.

A partir desses exemplos, nota-se com considerável clareza a presença de uma coloração negativa que é adicionada pelo uso de *few*, não precedido do artigo *a*, e especialmente em combinação com o intensificador *very* como colocado à esquerda. Certamente, a **prosódia semântica** negativa não teria sido expressa de forma tão contundente caso outro quantificador, supostamente mais neutro, como o quantificador *some*, tivesse sido empregado. Como já ressaltado, o contexto expandido revela bastante sobre a identificação da prosódia, pois o tipo de palavras/expressões que circundam o nóculo compõe a formação de sentido, fornecendo pinceladas de conotação, nesse caso desfavorável, como verificado nas linhas de concordância aqui apresentadas. Em breve discussão das mesmas, pode-se destacar

as palavras/expressões negativas ‘*give up*’, linha 1; ‘*fear*’, linha 3; ‘*tragedy*’, ‘*lack of sight*’ e ‘*appalling*’, linha 4; ‘*cramped with*’, linha 6; ‘*failure*’, linha 7; ‘*violence*’, linha 9, e outras mais, como responsáveis pela composição da formatação negativa de sentido.

Não obstante o fenômeno da **prosódia semântica** possa ser percebido nas línguas de um modo geral, trata-se de algo que suscita fervorosos debates entre os estudiosos (Hunston, 2007). Isso, conforme já discutido no capítulo 2, pode ser explicado pelo fato de tal fenômeno estar intimamente ligado à análise de atitude discursiva, que pode gerar interpretações subjetivas diferentes e, por conseguinte, estabelecer um grau de dificuldade na formalização de conclusões mais concretas a respeito de determinadas estruturas apresentarem tendências de prosódia positiva ou negativa.

Ainda com toda essa problemática, vale reforçar que é impossível negar a existência e importância da **prosódia semântica**, haja vista a recorrência de uso de determinadas estruturas em contextos positivos ou negativos. Vários estudos dessa natureza, como demonstrado, já foram documentados (SINCLAIR, 1991; LOUW, 1993; PARTINGTON, 1998). Mais ainda, segundo observado por Zhang (2009), diferentes estudos foram realizados com o objetivo de explorar a questão da **prosódia semântica** na interlíngua de aprendizes de segunda língua (SL) e língua estrangeira (LE). Tais estudos revelaram que, em diversos momentos, os aprendizes se encontram suscetíveis a erros pragmáticos decorrentes da falta de percepção e/ou compreensão da **prosódia semântica** de itens lexicais da língua-alvo.

O presente trabalho, portanto, confirma a observação feita por Zhang justamente por também apresentar exemplos que apontam para o fato de que a expressão de uma **prosódia semântica** (positiva ou negativa), ao se utilizar uma determinada estrutura, nem sempre se dá de forma tão natural e fluente. No caso aqui sendo analisado, a investigação está debruçada sobre a **prosódia semântica negativa** que emerge do uso de *few*, principalmente quando agrupado com *very* à esquerda, uso tal que pode ser verificado na escrita acadêmica de nativos de língua inglesa, como mostrado nos exemplos supracitados da parte do LOCNESS. Parte-se, então, para as linhas de concordância retiradas do Br-ICLE, (QUAD. 13) e subsequente análise das mesmas.

Quadro 13 – Grupo de linhas de concordância provenientes do Br-ICLE que apresentam **prosódia semântica** negativa

1. behalf of God, tricking poor people and profiting from this deplorable attitude. Financial practices of many churches and ministries are unknown and **few** people have access to them. There are no
2. seriously. Very **few** women would have their lives apart from a mans financial support, either the fathers or the husbands. Even their sexual activity would exist to satisfy the husband. This,
3. and find it out. By that time people were not instructed at all. In fact those who knew how to read were **few**, and fewer were those who knew how to write (yes, at that time this two almost unsplitable
4. A great number of citizens is aware of the emergencial situation the civilization is going through, however, **few** of these citizens demonstrate genuine concern with the problem. The major of the
5. criminals, and there isnt way to rehabilitate them. In fact, in general there are a **few** opportunities for everybody in our society. Moreover, if you are an ex-criminal your chances are lows, despite of
6. The concerns about this limitation are very **few**. Why? Because we are not aware of it. We live in our own bubbles and forget about the world. I think this is crucial in our indifference to pure dreams and
7. another person. The absence this cause **few** dreams because without imagination and little relationship leave the mind less active.
8. has been used, most of the times, to turn peoples misery into show, offer **few** poor cultural programs, establish standards to achieve happiness and the perfect body, just to mention some things from this
9. their lives. However, this is a challenge which very **few** people are able to overcome on the grounds that it is difficult to balance both things at the same time.

Fonte: Elaborado pela autora, por meio do programa AntConc 3.2.1.

Os exemplos de linhas de concordância acima mostram que os aprendizes utilizaram *few* em contexto negativo, e isso é verificado pelas palavras e expressões do contexto expandido, como ‘*deplorable*’, linha 1; ‘*apart from*’, linha 2; ‘*concerns*’ e ‘*limitation*’, linha 6; ‘*absence*’, linha 7; ‘*overcome*’ e ‘*difficult*’, linha 9, entre outras. A única exceção se encontra na quinta linha de concordância, na qual há o uso pragmaticamente inadequado de *a few* em contexto negativo. O uso inadequado se dá pelo fato de que a intenção do escritor era a de dizer que há poucas oportunidades para todos na sociedade, e não algumas. Utilizando *a few*, geralmente mais associado ao quantificador *some* (BIBER *et al.*, 1999), a interpretação se torna mais neutra, e menos restritiva. Logo, o uso de *few*, ou mais negativamente, *very few*, teria sido mais adequado. Além disso, enquanto na parte do LOCNESS houve muitas ocorrências de *very few*, gerando o padrão léxico-gramatical (intensificador) + (quantificador) + (substantivo), o mesmo não foi verdade para o Br-ICLE. Sendo assim, conjectura-se que os aprendizes subusaram tal padrão, e três argumentos podem ser invocados para confirmar a suposição levantada.

4.2.1 O primeiro argumento

Conforme citado anteriormente na primeira seção desse capítulo (ver quadro 6), o item lexical *very* está classificado na seção acadêmica do COCA como segundo mais frequente colocado à esquerda de *few*, o que confere à colocação *very few* alta relevância. Era de se esperar, por conseguinte, que dentre os usos de *few* na escrita acadêmica dos aprendizes brasileiros houvesse maior frequência desse padrão, ou seja, mais linhas de concordância com (intensificador = *very*) + (quantificador = *few*) + (substantivo), porém isso não aconteceu.

4.2.2 O segundo argumento

Segundo McEnery e Hardie (2012), para melhor entender e caracterizar os dados de frequência que emergem de um *corpus*, medidas estatísticas vêm sendo usadas por linguistas de *corpus* com o intuito de permitir que uma simples descrição do que está sendo observado seja averiguada por um teste de significância. Desse modo, o segundo argumento é respaldado estatisticamente por meio da aplicação do Teste Exato de Fisher (*Fisher's Exact Test*).

Como mencionado no capítulo 3, o cálculo do teste foi feito *online*, em visita ao endereço eletrônico <http://www.physics.csbsju.edu/stats/fisher.form.html>. Os dados foram compostos da seguinte maneira: Para 100 ocorrências do quantificador *few* na parte do LOCNESS, 13 são da colocação *very few*, por meio da qual as análises qualitativas mostraram a presença da **prosódia semântica negativa**. Já para o Br-ICLE, do total de 71 ocorrências com *few*, 3 formam a colocação *very few*. A figura abaixo mostra a página quando da preparação para o cálculo:

← → ↻ 🏠 www.physics.csbsju.edu/stats/fisher.form.html

Fisher Exact Test

Consider two dichotomous tests: [1 | 2] and [A | B] applied to a sample of size N . Display the number of (A1), (B1), (A2), (B2) cases in a 2x2 contingency table:

	A	B
1	a	b
2	c	d

If the tests are independent, ratios should be equal (e.g., $a/c = b/d$)... They rarely are equal, so the question is: are the ratios different enough to reject the hypothesis of independence.

In the below table fill in the required values. All entries should be whole numbers!

a= b=
 c= d=

number of (A1) cases= a number of (B1) cases= b
 number of (A2) cases= c number of (B2) cases= d

Figura 3 – Composição dos dados da parte do LOCNESS e do Br-ICLE para a aplicação do Teste Exato de Fisher

Fonte: FISHER EXACT TEST, 2014.

Considerando que o nível de significância de 5% (valor $p < 0.05$) determina uma direção para a interpretação dos resultados, tem-se que os *valores p* abaixo de 0.05 conferem significância ao fenômeno em estudo, desviando a hipótese de que ele se tenha dado pelo acaso. Assim, ao solicitar o cálculo, a página gerada mostrou o seguinte resultado, conforme a figura 4 que segue:

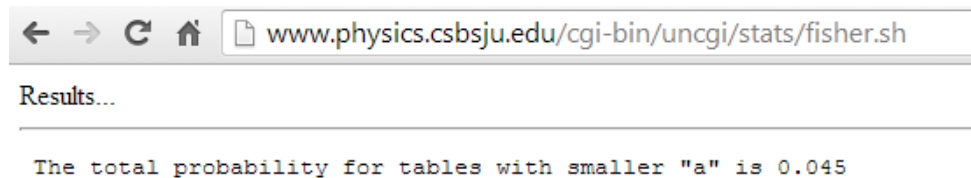


Figura 4 – Resultado do Teste Exato de Fisher para os dados fornecidos
Fonte: FISHER EXACT TEST, 2014.

A partir desse resultado (0.045), no qual o *valor p* foi abaixo de 0.05, fica corroborada a suposição de que os aprendizes de fato subusaram o agrupamento em questão, e que tal evento não se deu aleatoriamente. Isso deixa margens para o reforço da alegação de Xiao e McEnery (2006), quando eles dizem que a intuição dos aprendizes em relação à língua-alvo – diferentemente do que acontece com os falantes nativos – é tipicamente menos confiável e, como consequência, não os ajuda muito a detectar o uso de um item lexical em termos de sua **prosódia semântica**.

A baixa acuidade da intuição dos aprendizes poderia ser, em alguma medida, compensada por uma melhor instrução pedagógica, por meio da qual professores sejam conscientizados da necessidade de abordar questões de **prosódia semântica** em sala de aula e de confeccionar atividades que possam trazer à tona tais discussões. Uma maneira de fazer isso é utilizando as próprias produções dos aprendizes, como já o fez Seidlhofer (2002), que encorajou os aprendizes a fazerem uso dos recursos e ferramentas da LC para contrastar seus dados com os dados de nativos, com o objetivo de depreender padrões linguísticos e identificar possíveis inadequações cometidas.

Dentro das discussões acerca da prosódia é válido abordar a conexão dela com o registro que se pretende empregar nas produções textuais que são elaboradas pelos aprendizes. A observação do registro se torna importante porque além de a **prosódia semântica** poder ocorrer ‘globalmente’ no uso da língua em geral, há casos de **prosódia semântica** ‘local’, como atestados por Nelson (2000, 2006) e Tribble (2000). Nelson, por exemplo, realizou trabalhos analisando o comportamento dos itens lexicais *manage* e *package* em contexto de

negócios, comparando-os com o seu uso geral na língua. A **medida de informação mútua** (*MI score*) também pode contribuir para informar acerca do comportamento de colocações em determinados registros. Ao realizar uma busca de *very few* no COCA, sem pré-definir a seção acadêmica, o seguinte quadro 14 foi obtido:

Quadro 14 – 12 primeiros colocados à esquerda de *few* no COCA

	■	CONTEXT	FREQ		ALL	%	MI
1	■	A	151511		10456394	1.45	4.90
2	■	LAST	6342		356212	1.78	5.20
3	■	VERY	5720		480412	1.19	4.61
4	■	NEXT	5520		234111	2.36	5.60
5	■	PAST	4922		151236	3.25	6.07
6	■	FIRST	2847		550261	0.52	3.41
7	■	EVERY	2039		245961	0.83	4.09
8	■	RELATIVELY	1169		29826	3.92	6.33
9	■	SELECT	242		10533	2.30	5.56
10	■	PRECIOUS	202		8162	2.47	5.67
11	■	LUCKY	119		18415	0.65	3.73
12	■	SURPRISINGLY	87		9948	0.87	4.17

Fonte: DAVIES, 2014.

O valor MI para *very few* nesse quadro é de 4.61, enquanto que no quadro 6, além de *very* ocupar o segundo lugar – não o terceiro – entre os colocados, o valor MI é 6.22. Isso mostra que no registro acadêmico a atração entre o intensificador *very* e o quantificador *few* é maior, e tal fato pode ser agregado à manifestação do fenômeno da **prosódia semântica**, em especial no registro em questão, o acadêmico.

4.2.3 O terceiro argumento

Ao fazer uma busca no Co-RAP, foram encontradas 22 ocorrências de **poucos** e 7 de **poucas**, totalizando 29 ocorrências de **poucos(as)**. Entretanto, não houve ocorrências de **muito poucos** nem de **muito poucas**, que seriam as expressões correspondentes no português brasileiro para *very few* na língua inglesa. Tal resultado de busca pode apontar para a suposição de que a combinação **muito poucos(as)** talvez não seja vastamente usada pelos nativos do português brasileiro em registro acadêmico.

Partindo para o *corpus* de referência Corpus do Português³⁴, a mesma busca foi realizada, isto é, foram encontradas 3.983 ocorrências de **poucos** e 1.954 de **poucas**, totalizando 5.937 ocorrências de **poucos(as)**. Para a colocação **muito poucos**, foram 56

³⁴ CORPUS DO PORTUGUÊS, 2014.

ocorrências, e para **muito poucas** foram 60, somando 116 no total. Vale dizer que buscas com **muito pouco(a)** foram excluídas porque o quantificador *few* exige substantivos contáveis como complementos, logo não cabe fazer uma comparação com a colocação no singular, mas sim no plural.

De acordo com o cálculo da porcentagem, tem-se que as 116 ocorrências de **muito poucos(as)** no Corpus do Português correspondem a apenas 1,9% do total das ocorrências de **poucos(as)** – 5.937 – nesse corpus. A baixa porcentagem obtida atesta um uso pouco frequente do padrão em questão. Uma hipótese que pode contemplar uma explicação para esse resultado é a de que como o português brasileiro é uma língua na qual a concordância nominal leva em conta o número, podem parecer contra-intuitivas construções como ‘muito poucos estudos’ e ‘muito poucas pessoas’, por exemplo, embora as mesmas estejam gramaticalmente corretas.

Ainda em consulta ao Corpus do Português, foi feita uma busca dos colocados à esquerda de **poucos** e de **poucas**, e os seguintes quadros, 15 e 16, foram gerados:

Quadro 15 – 12 primeiros colocados à esquerda de **poucos** no Corpus do Português

	■	CONTEXTO	FREQ	ALL	%	MI
1	<input type="checkbox"/>	AOS	577	34911	1.65	6.96
2	<input type="checkbox"/>	HÁ	210	41868	0.50	5.24
3	<input type="checkbox"/>	UNS	107	8064	1.33	6.64
4	<input type="checkbox"/>	SÃO	84	59285	0.14	3.42
5	<input type="checkbox"/>	ERAM	53	14186	0.37	4.81
6	<input type="checkbox"/>	BEM	49	37904	0.13	3.28
7	<input type="checkbox"/>	ALGUNS	46	15482	0.30	4.48
8	<input type="checkbox"/>	HAVIA	35	20394	0.17	3.69
9	<input type="checkbox"/>	TÃO	35	26916	0.13	3.29
10	<input type="checkbox"/>	PASSADOS	12	863	1.39	6.71
11	<input type="checkbox"/>	HA	12	989	1.21	6.51
12	<input type="checkbox"/>	ÃOS	11	546	2.01	7.25

Fonte: CORPUS DO PORTUGUÊS, 2014.

Quadro 16 – 12 primeiros colocados à esquerda de **poucas** no Corpus do Português

	■	CONTEXTO	FREQ	ALL	%	MI
1	<input type="checkbox"/>	AS	122	220690	0.06	3.09
2	<input type="checkbox"/>	DAS	100	94964	0.11	4.02
3	<input type="checkbox"/>	UMAS	89	4612	1.93	8.21
4	<input type="checkbox"/>	MUITO	60	57069	0.11	4.01
5	<input type="checkbox"/>	SÃO	57	59285	0.10	3.88
6	<input type="checkbox"/>	NAS	40	34931	0.11	4.14
7	<input type="checkbox"/>	HÁ	40	41868	0.10	3.87
8	<input type="checkbox"/>	TÃO	26	26916	0.10	3.89
9	<input type="checkbox"/>	ERAM	24	14186	0.17	4.70
10	<input type="checkbox"/>	ESTAS	19	8165	0.23	5.16
11	<input type="checkbox"/>	ALGUMAS	18	10221	0.18	4.76
12	<input type="checkbox"/>	RESTAM	7	258	2.71	8.70

Fonte: CORPUS DO PORTUGUÊS, 2014.

Enquanto no quadro 15 é possível perceber que o item lexical **muito** nem consta na lista dos 12 primeiros colocados à esquerda de **poucos**, o quadro 16 apresenta esse item na quarta posição. Contudo, nesse mesmo quadro 16, verifica-se que o valor MI para **muito poucas** (4.01) é menor se comparado ao valor MI de *very few* (6.22) no quadro 6, o que reduz o grau de atração entre as duas palavras dessa colocação no português brasileiro.

Presume-se, portanto, que a interlíngua dos aprendizes quanto ao uso do quantificador *few* no que concerne a interação com a **prosódia semântica negativa** – gerada na língua inglesa pelo padrão *very few* – pode ser afetada pela influência do uso, ou pouco uso, de **muito poucos(as)** na língua materna. Desse modo, além de a captação do fenômeno da **prosódia semântica** na língua-alvo não se dar de forma tão natural ou transparente pelos aprendizes, o uso que os falantes brasileiros fazem de **muito poucos(as)** parece não estar pragmaticamente ou numericamente (em termos de frequência) emparelhado com o uso que os falantes nativos do inglês fazem do padrão correspondente em sua língua. Isso, portanto, contribui para fomentar a postura de ‘evitação’ que os brasileiros adotam frente ao uso dessa colocação.

Tendo em vista que no português brasileiro existe a forma adjetiva **pouquíssimos(as)**, cujo sentido se assemelha a **muito poucos(as)**, tomou-se a decisão de averiguar qual seria a manifestação da **prosódia semântica** em seu uso. Para tanto, seguem os dois quadros 17 e 18 que mostram ocorrências aleatórias de **pouquíssimos** e de **pouquíssimas**, retiradas do Corpus do Português.

Quadro 17 – Ocorrências aleatórias de **pouquíssimos** no Corpus de Português

6	19: Fic: Br: Carvalho: Suómi	A B C	couro cru, roupas de carrapicho.. Os urubus debandaram para outras regiões e os pouquíssimos que permaneceram, raquíticos, passivos e desmoralizados, emagre
7	19: Fic: Br: Comparato: Guerra	A B C	em água fervente a fim de que ressuscitassem no dia seguinte. Procedimento que rendia pouquíssimos êxitos e, mesmo quando o indivíduo ressuscitava, acordava t
8	19: Fic: Br: Olinto: Trono	A B C	que eu saiba distinguir entre o bem para o grande número e o bem para pouquíssimos que é o mal, que eu saiba ser dura como o aço e branda
9	19: Fic: Br: Paiva: Brasil	A B C	multos riachos de águas transparentes e geladas. Meses de treinamento. Poucos desistiram, pouquíssimos eram mandados embora por inadaptação. No comando, o
10	19: Fic: Br: Verissimo: Resto	A B C	desprezível: o Cel. Quim e seu feudo. Assim, aos trinta e pouquíssimos anos chegava Aristides Barreiro a uma posição que outros só atingiam depois' de entrados
11	19: Fic: Br: Lopes: Donas	A B C	fazem testamentos em que concentram todos os haveres nos seus herdeiros da lei ou em pouquíssimos mais. Assim, ninguém que as não tivesse conhecido em vida;
12	19N: Br: Bahia	A B C	recém-implantado, entre outros tópicos. # O professor lamenta que o assunto tenha ainda pouquíssimos estudos sistemáticos. " Temos que buscar e reconhecer o B

Fonte: CORPUS DE PORTUGUES, 2014

Quadro 18 – Ocorrências aleatórias de **pouquíssimas** no Corpus de Português

11	19: Fic: Pt: Lisboa: Pouco	A B C	misturada de risos acanhados. Ele de ordinário é homem de poucas falas e de pouquíssimas ideias. Cava. A terra pesada mais o enxatão não lhe terão dado grar
12	19N: Pt: Expr	A B C	Amália viveu sempre tão fora dele que nem o seu nome fez moda. Há pouquíssimas Amálias em Portugal. Como se fosse proibido imitar-lhe a música, copiar-lhe
13	19Ac: Pt: Enc	A B C	foram lançados, na década de 1960, os robôs eram dispendiosos e acessíveis a pouquíssimas empresas existentes em países mais desenvolvidos, principalmente
14	18: Branco: Casa	A B C	. Depois de haver bandarreado vida de fausto, com muitas ilusões perdidas, mas pouquíssimas lágrimas, porque a desgraça lhe anda sempre a morder os tãçoes
15	18: Caminha: País	A B C	, exceder de vinte e cinco por dia. Pois bem, o comandante Saldanha pouquíssimas vezes castigava conforme a lei. Colocava acima dela seus caprichos inexplicá
16	18: Machado: Pareceres	A B C	Constantino do Amaral Tavares, é mais uma composição que vem tomar lugar entre as pouquíssimas que conta o teatro nacional. Esta qualidade impõe à crítica r
17	18: Machado: Polêmicas	A B C	, e os que o exercem são o vício e a ignorância, feitas as pouquíssimas e honrosas exceções. Não serei exagerado se disser que o altar tornou-se balcão e

Fonte: CORPUS DE PORTUGUES, 2014.

Por meio da observação de tais exemplos, é possível depreender a emergência de uma prosódia negativa imbuída no uso desses vocábulos. No quadro 17, alguns exemplos de itens lexicais que contribuem para essa prosódia são ‘mal’, linha 8, ‘desistiram’, linha 9, e ‘lamenta’, linha 12. Para o quadro 18, podem ser destacadas palavras como ‘ordinário’, linha 11, ‘dispendiosos’, linha 13, e ‘castigava’, linha 15.

No total para o Corpus do Português, são 32 ocorrências de **pouquíssimos** e 17 de **pouquíssimas**. Já no Co-RAP, não houve ocorrências para **pouquíssimos**, e apenas uma de **pouquíssimas**, como mostra o quadro 19, a seguir:

Quadro 19 – Imagem gerada pelo programa AntConc da única ocorrência de **pouquíssimas** no Co-RAP

Concordance			Concordance Plot	File View	Clusters	Collocates	Word List	Keyword List
Hit	KWIC							File
1	vel em algum nível (salvo pouquíssimas exceções). Cabe ao admin:							#35.txt

Fonte: Gerado pelo programa AntConc 3.2.1.

É possível estimar, embora ainda seja necessária a realização de muitas outras pesquisas para melhor validar tal estimativa, que as variantes **muito poucos(as)** e **pouquíssimos(as)** parecem competir, no português brasileiro, para expressar a **prosódia semântica negativa**.

Após a apresentação e explanação de cada argumento, portanto, é possível defender a hipótese de que o padrão *very few* se encontra em **subuso** na produção escrita dos aprendizes participantes do Br-ICLE muito provavelmente pela dificuldade de transpor o fenômeno da **prosódia semântica** entre as línguas e também pelo fato de o padrão equivalente na língua materna – **muito poucos(as)** – não ser amplamente usado pelos falantes nativos do português. A língua materna, então, parece influenciar as decisões quanto a escolhas de palavras e expressões feitas pelos aprendizes, e a maneira de atenuar tal influência é fornecendo instruções mais pontuais para os aprendizes, através das quais eles possam visualizar exemplos reais de uso que os nativos fazem de uma determinada estrutura, e serem estimulados a colocarem os ensinamentos em prática em suas produções.

Além de se referendar a língua materna para a melhor compreensão dos fenômenos de interlíngua, Römer (2011) também reforça que estudos com *corpora* de

aprendizes são de extrema relevância para a confecção de materiais didáticos e para o provimento de instruções pedagógicas mais conscientes das “necessidades específicas de cada população de aprendizes” (MEUNIER, 2002 apud RÖMER, 2011, p. 209). Mais ainda, *corpora* de aprendizes ajudam a testar a intuição de professores quanto ao nível de dificuldade que os alunos têm na absorção de determinados fenômenos e padrões linguísticos da língua-alvo (GRANGER, 2002).

4.3 Análises: o detalhamento dos dados dos aprendizes

Nesse momento, as ocorrências provenientes do Br-ICLE com os quantificadores em estudo serão discutidas detalhadamente. Para tanto, a presente seção contará com quatro quadros nos quais estão dispostas todas as linhas de concordância com o nódulo *few* desse *corpus*.

Os três primeiros quadros, cada um deles contendo 16 linhas de concordância, mostram ocorrências em que não foram identificadas inadequações. Já o quarto, com 23 linhas de concordância, apresenta amostras de inadequações cometidas. Em termos numéricos, portanto, pode-se dizer que 67,6% das ocorrências com *a few / few* obtiveram uma estruturação léxico-gramatical adequada, e em 32,4% houve inadequações, como descrito no quadro 20, a seguir:

Quadro 20 – Porcentagens das linhas de concordância adequadas e inadequadas do Br-ICLE

Linhas de concordância adequadas no Br-ICLE	Linhas de concordância inadequadas no Br-ICLE
48 (= 67,6%)	23 (= 32,4%)

Fonte: Elaborado pela autora, 2014.

Isso mostra que em sua maioria os dados provenientes das produções dos aprendizes de nível intermediário a avançado de inglês revelaram um bom uso dos quantificadores, condizente com aquilo que é produzido por nativos do inglês. Passa-se, então, para a discussão tanto das adequações quanto das inadequações.

4.3.1 Os três quadros com as adequações

Quadro 21 – 1º quadro com 16 linhas de concordância com *a few* e *few* no Br-ICLE cuja avaliação léxico-gramatical é adequada

3 ||: special student. In conclusion, there are few universities that can really prepare students for
 5 ||al world in their paintings, sometimes spending a few days to complete their work. With the invention o
 7 ||ces and things that wouldve not been acessible a few decades ago are now available to us through the to
 9 || everything we may become a society controlled by few people, with no opinion and no power and this is
 10 ||ir lives. However, this is a challenge which very few people are able to overcome on the grounds that i
 12 || sit and study for hours, other must stop after a few hours, or minutes, have a break and continue. I'v
 13 ||te moment of human history we've got many doubts, few foresights and only one assurance: job's future,
 14 ||ion is wealthy but that wealth is divided amongst few people who are usually not productive. Life cond
 16 ||ge this situation. It is unacceptable that, while few people detain the power and as a consequence the
 17 || our world is divided into two unequal parts: the few owners of capital and the huge mass of hired men.
 18 ||en that are all the time being manipulated by the few existing 'pigs', the ones that detain the real mo
 19 ||ism and Communism, advertising us to the danger f few in command declaring themselves equal... but more
 20 ||agacin, discussing who was pretty and selecting a few to take away. Dr. Femi Bojaxhiu, a gynecologist a
 21 ||reated as slave. In the 19th century there were a few reform movements and they invented a new system:
 22 ||ngs, because we are loosing the power of words. A few moments after the birth, the child is already int
 23 ||be one of the most important conquest in the last few centuries. fileseparator <BRDEN004> <Feminism: th

Fonte: Gerado pelo programa AntConc 3.2.1.

Com relação aos colocados à direita de *few*, todos eles se configuram como substantivos contáveis (*'universities'*, *'people'*, *'decades'*, *'days'*, etc.), conforme é verificado nos *corpora* de referência e como rege a norma gramatical para esse quantificador. Há também os casos das linhas 19 e 20, em que *few* retoma implicitamente um substantivo contável, que em ambas é *people*.

As expressões ligadas a tempo foram construídas com *few* precedido de *a*, como nas linhas 5, 7, 12 e 22, e também obedecendo aos padrões de estruturas utilizados por nativos, como na linha 23 (*'in the last few centuries'*). Pode-se verificar esse padrão no LOCNESS (*"has expanded rapidly in the last few years and has raised many moral questions"*), e atestam-se 226 ocorrências na seção acadêmica do COCA com tal agrupamento de 4 palavras.

Mais ainda, a **prosódia semântica** negativa foi devidamente expressa pelo uso de *few* – sem o artigo *a* à esquerda – em linhas como a 9, 10, 13, 16, 17, 18 e 19. O quadro 22 a

seguir mostra os itens lexicais de cada uma dessas linhas que contribuíram para a conotação desfavorável nelas impressa:

Quadro 22 – Itens lexicais do contexto expandido (referentes às linhas de concordância do quadro 21) que contribuíram para a emergência da **prosódia semântica** negativa no Br-ICLE

9	10	13	16	17	18	19
<i>controlled</i>	<i>challenge</i>				<i>manipulated</i>	
<i>no opinion</i>	<i>overcome</i>	<i>doubts</i>	<i>unacceptable</i>	<i>unequal</i>	<i>'pigs'</i>	<i>danger</i>
<i>no power</i>						

Fonte: Elaborado pela autora/Adaptado do *corpus* Br-ICLE.

Quadro 23 – 2º quadro com 16 linhas de concordância com *a few* e *few* no Br-ICLE cuja avaliação léxico-gramatical é adequada

24 ||e times, to turn peoples misery into show, offer **few** poor cultural programs, establish standards to ac
 25 ||osite programs, problems are forgotten even for a **few** minutes. Moreover, today the programs and TV seri
 26 ||s of many churches and ministries are unknown and **few** people have access to them. There are no supervis
 29 ||ve all kinds of content, appropriate or not. In a **few** cases, which is better depending on the point of
 31 ||t watched ones. But they are not. In fact, just a **few** people watch them. So, what can one conclude from
 32 ||ers. fileseparator <ICLE-BR-CAT-0021.1> There are **few** topics so much linked to modernity and our contem
 33 ||e means which were so enthusiastically welcomed a **few** decades ago. As a result, it is common to see you
 36 ||r neighborhood. These citizens have their already **few** and hard earned belongings such as watches, mobil
 38 ||y to clarify the problem, we must ask ourselves a **few** questions: is a university degree really an essen
 39 ||o, a nationwide exam created by the government a **few** years ago with the objective of evaluating the qu
 40 || in professions that require a technical level. A **few** years ago the job market favored those who had go
 41 ||pprove measures that represent the interests of a **few** men if their monthly bribe (Mensalão) is payed.
 43 ||a subject that people would treat seriously. Very **few** women would have their lives apart from a mans f
 44 ||d throughout History, tending to be accepted by a **few** nations in the world. These perspectives tend to
 45 ||tal problems. Therefore the reader is left with a **few** crucial questions: can we still believe theres a
 47 ||d at all. In fact those who knew how to read were **few**, and fewer were thos who knew how to write (yes,

Fonte: Gerado pelo programa AntConc 3.2.1.

De modo similar ao quadro anterior, os colocados à direita são todos substantivos contáveis, explicitamente ou implicitamente expressos. Em sua maioria, as expressões ligadas a tempo são construídas com *a few* (*'for a few minutes'*, linha 25; *'a few decades/years ago'*, linhas 33, 39 e 40). Há uso adequado do padrão *just a few* na linha 31, agrupamento esse que é encontrado em 3 linhas de concordância no LOCNESS (*'bring joy to many and misery to just a few'*; *'The ideas above are just a few of the topics we are now concerned with'*; *'These are just a few common misconceptions'*) e conta com 801 ocorrências (dentre elas: *"and*

movie theaters were *just a few* of the injustices that African-Americans suffered throughout the South”) na seção acadêmica do COCA. Pode-se dizer também que a **prosódia semântica** negativa foi identificada nas linhas 24, 26, 36, 43 e 47, ratificada pelos itens lexicais listados no quadro 24, a seguir:

Quadro 24 – Itens lexicais do contexto expandido (referentes às linhas de concordância do quadro 23) que contribuíram para a emergência da **prosódia semântica** negativa no Br-ICLE

24	26	36	43	47
<i>misery , poor</i>	<i>unknown</i>	<i>hard</i>	<i>apart from a man's financial support</i>	<i>not instructed at all</i>

Fonte: Elaborado pela autora/Adaptado do corpus Br-ICLE.

Vale mencionar que o intensificador *very*, inclusive, foi utilizado como colocado à esquerda na linha 43, corroborando, também, para a expressão da conotação desfavorável.

Quadro 25 – 3º quadro com 16 linhas de concordância com *a few* e *few* no Br-ICLE cuja avaliação léxico-gramatical é adequada

49 ||e same time we see people making deals with other few in order to take advantage and profit. Corruption
 52 ||eligion. This idea has been altered over the past few years and now religion is not societys aim. Rel
 53 ||however this is not the reality faced by the last few generations. It was somewhat impossible for most
 54 ||turn has changed little, still maintaining only a few as legitimate owners of free will. fileseparator
 55 ||ation the civilization is going through, however, few of these citizens demonstrate genuine concern wit
 56 ||ntly together for a programs last chapter. Those few examples reveal how television has become an inst
 58 ||this occurs, as it is not to the advantage of the few who are on top, the government and corporate head
 62 ||agination to live happy or escape the world for a few moments. Keep dreaming and using imagination, it
 63 ||r anything with a broken card. Usually it takes a few days to replace a card after a phone call or a pe
 64 ||ming increasingly popular thanks to the work of a few energetic reformers. Many of these ideas were rel
 65 ||hing. The concerns about this limitation are very few. Why? Because we are not aware of it. We live in
 67 ||nship with another person. The absence this cause few dreams because without imagination and little rel
 68 ||r concern because of poor responses over the past few years. In summary, the relationship between AIDS
 69 || <Living in a high tech world> Over the past few decades, technology and the science have been sev
 70 ||ndling money and credit cards wisely is a talent few of us are born with. But it is a skill that can e
 71 ||cts through rehabilitation. Both are necessary. A few years ago, the Brazilian TV show Fantástico sho

Fonte: Gerado pelo programa AntConc 3.2.1, 2013.

Atentando-se ao uso dos quantificadores nesse último quadro, pode-se dizer que, novamente, os colocados à direita são adequados e há exemplos de padrões recorrentemente usados por nativos, como *over the past few years/decades*, das linhas 52, 68 e 69, *by the last few generations*, linha 53, *for a few moments*, linha 62, *it takes a few days*, linha 63, e *a few years ago*, linha 71. No que tange aos exemplos da manifestação da **prosódia semântica** negativa, o quadro 26 que segue mostra as linhas com suas respectivas palavras de sentido negativo:

Quadro 26 – Itens lexicais do contexto expandido (referentes às linhas de concordância do quadro 25) que contribuíram para a emergência da **prosódia semântica** negativa no Br-ICLE

49	55	56	65	67
<i>take advantage</i>	<i>going through</i>	<i>manipulating</i>	<i>limitations</i>	<i>absence</i>

Fonte: Elaborado pela autora/Adaptado do corpus Br-ICLE.

4.3.2 O quadro com as inadequações

Quadro 27 – 4º quadro com 23 linhas de concordância com *a few* e *few* no Br-ICLE cuja avaliação léxico-gramatical é inadequada

1 urgent for the mankind do not face consequences in few years.
 2 bjective, something that is accessible just for a few people. In sum, what Orwell actually said is that
 4 ge, but is necessary we will continue to learn a few more every day and we will acquire the experienc
 6 a licenses degree to become a teacher, have only few courses during the four or five years of study to
 8 ars, we have great furnitures, papers... all this few things that made our life so easy. And we still d
 11 urselves: do you think that these politician earn few money and it is really necessary to have even mor
 15 an states. Unfortunately, in Brazil, we have just few examples of prisons with agriculture work. We are
 17 information, which probably they will forget in a few time, and they do not practice, so they do not kn
 20 heir name to sell courses in a capitalist way. In few years people can get a diploma in degrees called
 30 e is still getting a shape. That is why we see so few wars in the Americas. Maybe for a political postu
 34 ll, recover of mens extractions as there were only few people living in the Earth. Centuries later, the
 35 criminals to live in society. Actually, there are few of them, but they are truly engaged in the aim of
 37 ave it. Drug dealing, sexual abuse and murder are few of the possibilities they elaborate in their mind
 42 ations this certainly would not be possible in so few time. All things considered we should not say the
 46 ut that was possible to realize by studies on past few years is the increase of cancer on people who liv
 48 an I believe that we are going to over pass it in few years. fileseparator <ICLE-BR-CAT-0094.1> Regrett
 50 area of interest. In Germany, for instance, only few students attend a university and the requirement
 51 sculous life will be in a hundred years from now. few years ago, society dreamt that by the year two th
 57 to the supposed underqualification, Brazil has a few percentage of citizens studying at universities.
 59 niversities were able to prepare students in just few areas, such as medical and law. Eventually they s
 60 s a university degree is not a privilege for just few people anymore: it is the first step to get a res
 61 ehabilitate them. In fact, in general there are a few opportunities for everybody in our society. Moreo
 66 ll improve quality of live. Nowadays, diseases that few decades ago were considered fatal can be quickly

Fonte: Gerado pelo programa AntConc 3.2.1.

O tratamento das inadequações representadas pelas linhas de concordância do quadro 27 foi feito a partir da escolha de dividir os tipos de erros em categorias nas quais as linhas foram encaixadas a fim de agrupá-las de acordo com a divisão proposta.

Logo, analisando os dados qualitativamente à luz de noções semântica, pragmática e sintaticamente direcionadas, as categorias que emergiram são as seguintes: 1. Colocações improváveis (de baixa frequência ou quase inexistentes); 2. Interpolação indevida de preposição; 3. Inadequação de uso de *few* e *little* (como quantificadores ou não); 4. Uso equivocado de pronome demonstrativo e de palavras de ênfase com *few*; e 5. Problemas na expressão da **prosódia semântica**.

O quadro 28 a seguir mostra as categorias e as linhas de concordância que nelas se enquadram:

Quadro 28 – Divisão em categorias das inadequações cometidas pelos participantes do Br-ICLE

Categorias	Linhas de Concordância
1) Colocações improváveis	1, 6, 15, 28, 34, 37, 46, 48, 50, 51, 59, 60, 66
2) Interpolação indevida de preposição	2
3) Inadequação de uso de <i>few</i> e <i>little</i>	4, 8, 11, 27, 42
4) Uso equivocado de pronome demonstrativo, de adjetivo e de palavras de ênfase com <i>few</i>	8, 30, 42, 57
5) Problemas na expressão da prosódia semântica	28, 35, 61

Fonte: Elaborado pela autora/Adaptado do *corpus* BR-ICLE.

No que diz respeito à primeira categoria, pode-se dizer que *in few years*, agrupamento que aparece nas linhas 1, 28 e 48, é praticamente inexistente. Nem no LOCNESS nem no COCA há ocorrência de tal agrupamento; em contrapartida, *in a few years* possui uma única ocorrência no LOCNESS e 75 no COCA acadêmico. Igualmente, o agrupamento *on the past few years*, que aparece na linha 46, pode ser considerado quase inexistente. Não há ocorrências dele nem na parte do LOCNESS e nem no COCA, mas sim de *over the past few years* – 4 no LOCNESS e 146 ocorrências no COCA acadêmico. Outros dois agrupamentos problemáticos são *few years ago* e *few decades ago*, das linhas 51 e 66, respectivamente, ambos não verificados no LOCNESS. De um lado, no COCA acadêmico, *few decades ago* não é encontrado, e *few years ago* possui apenas 1 ocorrência; do outro lado, *a few decades ago* possui 42 ocorrências e *a few years ago* totaliza 458 nesse mesmo *corpus* de referência. No LOCNESS, não há ocorrências de *a few decades ago*, mas há 2 ocorrências de *a few years ago*. Dando seguimento à análise de colocações improváveis, *just few examples*, que aparece na linha 15, bem como *just few areas* e *just few people*, das linhas 59 e 60, são, também, praticamente inexistentes. Nem no LOCNESS nem no COCA há ocorrências de *just few*, ao passo que há 3 ocorrências no LOCNESS de *just a few* e 801 na seção acadêmica do COCA. Nas linhas 6, 34 e 50 aparecem as colocações *have only few*

courses, there were only few people e only few students. Embora *only few* seja possível, pode-se dizer que é uma combinação pouco frequente, pois não aparece no LOCNESS e conta com apenas 27 ocorrências no COCA acadêmico. *Only a few*, contudo, possui 5 ocorrências no LOCNESS e 1.911 no COCA. A linha 37 apresenta um agrupamento que teria sido mais bem estruturado com a inclusão do determinante *a*, pois observando o contexto expandido é possível notar que a substituição pelo vocábulo *some* pode ser feita, e isso é indício de que *a few*, não *few*, deveria ter sido empregado – ‘*sexual abuse and murder are a few (some) of the possibilities they elaborate...*’.

Para a segunda categoria, tem-se que a interpolação da preposição ocorrida na linha 2 foi indevida. No LOCNESS, não há ocorrências de *just for a few* nem de *just to a few*. Considerando que o vocábulo *accessible* é o quarto colocado à esquerda do nódulo *few*, a preposição utilizada deve obedecer à sua regência. Consultando o COCA acadêmico, foram encontradas 769 ocorrências de *accessible to*, contra 82 de *accessible for*, o que indica que o primeiro agrupamento seria o mais recomendado. Ainda assim, a sequência sem a interpolação – *accessible to just a few people* – se comparada a *accessible just to a few people*, é a melhor opção, pois segue o padrão apresentado nas linhas de concordância geradas pelo COCA, como mostra o quadro 29, a seguir:

Quadro 29 – Linhas de concordância com o padrão *to just a few*, aleatoriamente retiradas do COCA acadêmico

CLICK FOR MORE CONTEXT		[?]	SAVE LIST	CHOOSE LIST	----- ▾	CREATE NEW LIST	[?]
1	2011	ACAD	Archaeology	A	B	C	. # Most of our understanding of this period has been limited to just a few written sources. Almost all of the
2	2009	ACAD	Bioscience	A	B	C	contact rates, the impact of these spillover events can be limited to just a few individuals, or may lead to le
3	2005	ACAD	ABAJournal	A	B	C	says there is more pressure on existing partners to generate business compared to just a few years ago. #
4	2005	ACAD	Environment	A	B	C	competitive economic system. These global results, however, are limited to just a few variables and gloss (
5	2001	ACAD	IntlAffairs	A	B	C	# Increasingly, the presence of child prostitution is no longer contained to just a few places. As we have ex
6	2000	ACAD	Mercury	A	B	C	von Hippel states, in research scientists tend to communicate new ideas to just a few colleagues, whereas i
7	1999	ACAD	Mercury	A	B	C	path lengths from the outer mirrors to the eyepiece must be maintained to just a few wavelengths of light.
8	1998	ACAD	MusicEduc	A	B	C	software programs. Then they can try introducing a computer software program to just a few students. Aft
9	1993	ACAD	WorldAffairs	A	B	C	in the midst of a declining membership (from 5,000 in 1991 to just a few hundred). (n40) Some said that th

Fonte: DAVIES, 2014.

A terceira categoria lida como a distinção entre *few* e *little*, vocábulos que podem ser comumente confundidos por aprendizes. Considerando que na linha 4 não há substantivo contável explícito após *few*, e nem que possa ser recuperado no texto, *little* é o item lexical correto a ser usado. Na linha 8, *little* também deveria ter sido empregado (ou o adjetivo *small*), por serem, nesse caso, adjetivos – não quantificadores – que, de acordo com o contexto, estão em contraste à *big*. Em se tratando da linha 11, pelo fato de *money* ser um substantivo incontável, o vocábulo *little* também deveria ter sido escolhido, em detrimento de *few*. Do mesmo modo, nas linhas 27 e 42, *few* não poderia quantificar *time*, por este último se

tratar de um substantivo incontável. Novamente, portanto, seria o caso de se empregar *little*, muito embora, especificamente para a linha 27, o agrupamento *in a little time* se configure como de baixíssima frequência – apenas 1 ocorrência no COCA acadêmico.

Já a quarta categoria está atrelada ao uso de algumas palavras na circunvizinhança do nóculo *few*. Na linha 8, tendo em vista a pluralidade de *few things*, o pronome demonstrativo *these*, e não o singular *this*, deveria ter sido utilizado. Nas linhas 30 e 42, devido à presença dos substantivos *wars* e *time*, respectivamente, o vocábulo gerador de ênfase *such*, e não *so*, teria sido a escolha correta, como em ‘*such few wars*’ e ‘*such little time*’. Vale recuperar, do parágrafo anterior, que além de estar presente nessa categoria, a linha 42 foi discutida dentro da categoria 3. Por último, na linha 57, o escritor deseja qualificar a porcentagem de alunos brasileiros em universidades, atribuindo-lhe a circunstância de que é pequena. Logo, por ser uma atribuição de característica, o item lexical adequado, no lugar do quantificador *few*, seria o adjetivo *small*, que nesse contexto se opõe a *great / large*.

A quinta e última categoria está ligada aos efeitos de **prosódia semântica** e ao melhor cumprimento da expressão de conotação negativa a partir do que se percebe pelo contexto expandido. Assim, pode-se dizer que o item lexical *just* ou o item *very* deveriam ter sido empregados na linha 35 (‘*Actually, there are just a few of them / very few of them, but they are...*’). O escritor está se referindo a pouca quantidade de organizações não-governamentais que auxiliam ex-criminosos a serem inseridos ou re-inseridos na sociedade. De modo semelhante, na linha 28 (também encaixada na categoria 1), o vocábulo *very* poderia ter sido empregado, o que faria com que o agrupamento em questão – *in few years* – se tornasse melhor formulado (*in very few years*), auxiliando na transmissão do argumento crítico do escritor, o qual passa a ideia de que ao realizar determinados cursos técnicos, em apenas ‘**alguns**’ anos, ou em ‘**muito poucos**’ anos, os alunos já conseguem adquirir um diploma. Por fim, na linha 61, a intenção do escritor era também a de imprimir uma prosódia negativa afirmando que há ‘**poucas**’ oportunidades para muitas pessoas na sociedade, e não ‘**algumas**’. Logo, ao invés de *there are a few opportunities*, que conduz ao entendimento de ‘**algumas**’ oportunidades, os agrupamentos *there are few opportunities* ou *there are very few opportunities* teriam sido mais apropriados, por implicarem o sentido de ‘**poucas**’ ou ‘**muito poucas**’ oportunidades.

4.4 Implicações dos resultados

Conforme reportado na Introdução desse trabalho, Leech (1997) distingue entre as aplicações pedagógicas de *corpus* que são diretas e as que são indiretas. Indiretamente, os resultados do presente trabalho apontam para a necessidade de revisitação e aprimoramento das informações contidas em livros didáticos a respeito dos quantificadores *a few* e *few*. O quadro 23 a seguir, retirado do *Longman Online Dictionary*³⁵, além de explicar o sentido dos quantificadores em questão, inclui considerações sobre os registros nos quais eles ocorrem. Nas partes sublinhadas é explicado que ‘*a few* significa um número pequeno’, enquanto ‘*few* significa não muitos, e é principalmente usado na escrita ou fala formal’ (tradução minha).

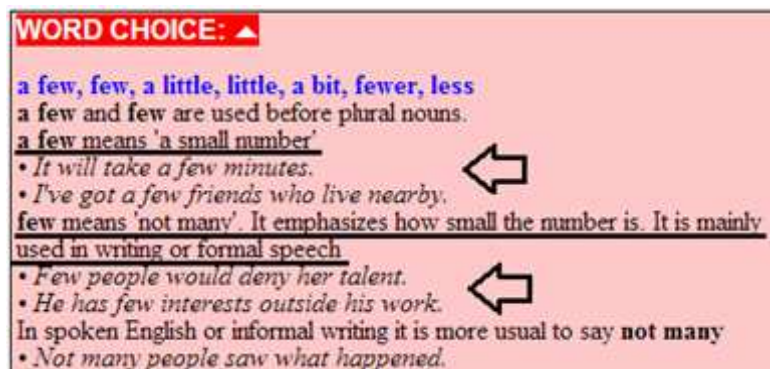


Figura 5 – Obtida pela busca do item lexical *few* no dicionário *online Longman*
 Fonte: LONGMAN, 2014b.

De modo similar a essas definições, os livros didáticos deveriam trazer discussões que levem em conta o *framework* tridimensional de Larsen-Freeman (2003), o qual inclui **forma, sentido e uso**. Na verdade, seria ideal que as discussões dos livros didáticos fossem ainda mais extensas e detalhadas, enriquecidas com exemplos retirados de *corpora*, abordando não somente a definição dos quantificadores e seus requisitos gramaticais, mas também, com base nos moldes da montagem de um perfil léxico-gramatical³⁶, segundo O’Keeffe, McCarthy e Carter (2007, p. 14), as seguintes características:

- a. seus colocados mais frequentes (como os itens lexicais relacionados a tempo, majoritariamente ligados a *a few* → *a few years/minutes/weeks*, etc.; e os itens lexicais relacionados a pessoas, muito ligados a *few* → *few people/women*, etc.);

³⁵ LONGMAN, 2014b.

³⁶ De acordo com O’Keeffe, McCarthy e Carter (2007, p. 14), uma estratégia permitida pela metodologia de *corpus* é a criação de um perfil léxico-gramatical de uma palavra e de seus contextos de uso. Os autores dizem que tal perfil é traçado a partir da descrição de colocados, de agrupamentos e expressões idiomáticas, de restrições sintáticas e semânticas e da **prosódia semântica**.

- b. as colocações que são tipicamente formadas (como *a few years ago/ over the past few years/ very few people*, etc.) e o uso adequado das mesmas nos diferentes registros (conversaço, acadêmico, por exemplo);
- c. a **prosódia semântica** de determinadas colocações, em consonância com a sua função discursiva em algum registro.

Tendo em vista que, ao contrário do abundante uso que *very few* recebe por parte dos nativos do inglês, as combinações **muito pouco(s)** e **muito pouca(s)** não sejam vastamente empregadas pelos falantes do português brasileiro, talvez se faça necessário lançar mão de exercícios direcionados, à luz dos princípios da **Aprendizagem Movidada por Dados** – AMD (JOHNS, 1994), para mais diretamente influenciar o ensino dos quantificadores e minimizar o impacto da transferência da língua materna nesse processo. Muito embora a AMD já tenha influenciado diversas pesquisas nessa direção (BRAUN; KOHN; MUKHERJEE, 2006), Braun (2007) aponta para alguns desafios quanto a dificuldades metodológicas de se usar *corpora* pedagogicamente. Como colocado por Widdowson (2003 apud BRAUN, 2007), é necessária uma mediação pedagógica no intuito de (re)contextualizar os dados provenientes de um *corpus* para fazê-los adaptáveis à natureza discursiva e comunicativa inerente à aprendizagem de uma língua em sala de aula. Outro ponto a ser considerado se fundamenta na alegação de Cook (1998 apud BRAUN, 2007), quando ele argumenta que pelo fato de os aprendizes em geral não serem usuários eficientes de ferramentas para a exploração de *corpora*, e lembrando que *corpora* são, em sua grande maioria, compilados para auxiliar na descrição de línguas, o salto da linguística até a pedagogia está longe de ser direto. Mais ainda, incluir atividades com *corpus* em programas curriculares não se constitui como uma tarefa fácil, principalmente devido a restrições de tempo. Ainda com todos esses desafios, Braun (2007, p. 308) afirma que

a integração (bem-sucedida) de *corpora* no contexto escolar depende de dois requisitos pedagógicos. O primeiro está ligado aos *corpora per se*. [...] Dados os requisitos curriculares no contexto escolar, postula-se que *corpora* para fins educativos não devem apenas ser mais coerentes do que *corpora* tradicionais; eles devem, na medida do possível, também ser complementares ao currículo escolar, para facilitar ambos o processo de contextualização e os problemas práticos de integração. O outro requisito é concernente à mediação pedagógica de *corpora*. [...] uma forma de alcançar isso é pelo ‘enriquecimento’ dos *corpora* com materiais de suporte, como atividades de aprendizagem baseadas em *corpus* e exercícios que

fundamentem a construção do conhecimento e forneçam aos aprendizes oportunidades de praticar e testar seu conhecimento. (tradução minha)³⁷

Para tanto, a seção 4.5 deste capítulo traz uma proposta de atividade que pode ser aplicada a alunos cujo nível de proficiência na língua-alvo é de intermediário a avançado, com o objetivo de propiciar a eles uma oportunidade mais contundente de melhor absorver e praticar o conhecimento acerca dos quantificadores. Ressalta-se que essa é apenas uma sugestão de atividade, que serve de ilustração para auxiliar professores interessados na aplicação direta – em sala de aula – e métodos de *corpus*.

Logo, a atividade sugerida se divide em duas partes, parte A e parte B (ver seção 4.5). A primeira delas serve o propósito de introduzir os alunos à realidade da existência de padrões linguísticos e também da importância de analisar produções de aprendizes em si. Assim sendo, em A.1 é solicitado que eles leiam três linhas de concordância retiradas do Br-ICLE e tentem descobrir qual palavra se encaixa em todas elas. Essas três linhas de concordância trazem um padrão recorrente; por isso, na parte A.2, os alunos são estimulados a tentar reconhecer esse padrão e discriminá-lo usando um tipo de fórmula linguística. Para tanto, além das linhas de concordância já listadas – do Br-ICLE – são fornecidas linhas de concordância retiradas do COCA – *corpus* de referência de falantes nativos – a fim de, com mais exemplos, facilitar a identificação do padrão que se procura que os aprendizes reconheçam. Após terminada a parte A, inicia-se a B, que foi elaborada com o objetivo de convidar os alunos a, por meio da apresentação de linhas de concordância provenientes do COCA, realizarem uma análise dos quantificadores *a few* e *few*, análise essa que fosse capaz de cobrir não só o escopo de padrões gramaticais, mas também o de padrões semânticos. Esse é o objetivo que se procura alcançar por meio da questão de múltipla escolha colocada em B.1, na qual se espera que os alunos marquem a letra b como resposta, pois é a que indica a existência da prosódia desfavorável na colocação *very few*, distinguindo os quantificadores em termos semânticos e pragmáticos, conforme foi verificado a partir dos resultados obtidos nesse trabalho. Tal questão, B.1, possui uma natureza um pouco mais guiada, podendo assim, de algum modo, ferir os princípios de descoberta da AMD; todavia, ela pode ser útil para perfis de alunos que necessitam de maior direcionamento. Quanto aos itens subsequentes da

³⁷ “the (successful) integration of corpora into a school context hinges on at least two specific pedagogical requirements. The first of these relates to the corpora themselves. [...] Given the curricular requirements in the school context, one can argue that corpora for secondary education should not only be more coherent than traditional corpora; they should, as far as possible, also be complementary to school curricula, to facilitate both the contextualisation process and the practical problems of integration. Another requirement concerns the pedagogical mediation of corpora. [...] one way of achieving this is through the ‘enrichment’ of corpora with support materials such as corpus-based learning activities and exercises to support the construction of knowledge and provide learners with opportunities to practise and test their knowledge.”

parte B (B.2.1 até B.2.5), os mesmos propõem a observação crítica das produções dos próprios aprendizes, no que concerne o uso adequado e inadequado dos quantificadores em questão. A ideia é a de que os alunos tenham em mente as conclusões já traçadas na seção anterior a fim de julgarem o quanto as sentenças produzidas por aprendizes se desviaram do que tenham possivelmente descoberto como padrão. Assim, eles são estimulados a comentar e propor melhorias para as frases do Br-ICLE. Vale mencionar que houve, também, a preocupação de incluir uma pergunta (em B.2.5) que sondasse a relevância da exposição das linhas de concordância retiradas do *corpus* de referência COCA para que os alunos chegassem às deduções esperadas em relação ao uso dos quantificadores e respondessem às questões colocadas.

4.5 Activity - Patterns & Quantifiers

A. Let's try recognizing patterns:

A.1. The following sentences were taken from Br-ICLE, the Brazilian *subcorpus* of the ICLE (*International Corpus of Learner English*). They should be filled in with the SAME word, in different forms. Try to figure out which word it is and complete the sentences.

1. Entertainment has _____ a very big industry.
2. Close relationships, big families, free time to rest are _____ very rare nowadays.
3. Because of this, luck _____ a very important component in the achieving professional success process.

A.2. After you have discovered the word, read the sentences carefully one more time. Also read the concordance lines below - taken from the academic section of COCA (*Corpus of Contemporary American English*) and find a recurrent pattern in them. Write the pattern you observed:

62 2004 ACAD MechanicalEng That **becomes** very important at high-frequency observations, where the telescope will have its biggest impact.

63 2004 ACAD Adolescence they had **become** very independent -- out of necessity rather than desire.

67 2004 ACAD TeachLibrar Time will **become** very precious in the last few weeks before opening, and the last thing you

68 2004 ACAD AfricanHist the situation would have **become** very difficult indeed. " 73 # Qadi " Umar admitted that the girl had

69 2003 ACAD AmerStudies Biographies have **become** very popular. Books devoted to Wilson, Roosevelt, Kennedy, and other prominent

Pattern: _____ (____) _____

B. Take a look at the following sets of concordance lines taken from COCA and pay attention to the uses of the underlined words. What are these words?

1st set:

10 2009 ACAD Adolescence However, few of them recognize how feedback can motivate classroom teachers to continue their difficult tasks

83 2009 ACAD Archaeology Few ancient Maya remains have been as thoroughly studied as Pakal's.

97 2009 ACAD HealthSocialW The participants made very few references to the state of their mind or mental health, but they made

98 2009 ACAD ABAJournal " Very few cases actually go to trial today. It is a disturbing and unhealthy trend

99 2009 ACAD MiddleEastQ Pakistan was among the first countries to recognize the new Islamic Republic and was among very few countries in the region that refrained from supporting Iraq during the Iran-Iraq war.

2nd set:

76 2009 ACAD AnthropolQ Iran now has a few centers of excellence. These centers function, to some extent, autonomously of

77 2009 ACAD Archaeology and important people (perhaps a few from distant lands). Scarlet macaws from the Maya jungles and copper artifacts from

78 2009 ACAD Archaeology so I set out a big map of the region and dotted in a few key sites. I knew Paquim would come up (one old theory held that

79 2009 ACAD Archaeology fair." # Of course I took a bit of guff from a few archaeologists who objected to Chaco Meridian's argument (fair enough). My Paquim

80 2009 ACAD Archaeology Well, not precisely due north--Sacred Ridge is actually a few miles east of Pueblo Bonito, Chaco's biggest great house.

B.1. As you could see, the concordance lines carry the uses of (*very*) *few* and *a few*. Considering the concept behind the term *semantic prosody* (according to John Sinclair, it relates to the favorable or unfavorable connotation contained not in a single item, but in this item associated with others) you might infer that:

- a) (*very*) *few* and *a few* can be used interchangeably and share the same prosody.
- b) (*very*) *few* seems to be more restrictive and assertive than *a few*, and this can be at times linked to a negative prosody.
- c) (*very*) *few* and *a few* differ in meaning, but have the same prosody.
- d) Since there isn't a clear tendency, (*very*) *few* and *a few* may have both positive and negative prosodies, depending on the context.

B.2. Now read the following sentences taken from Br-ICLE:

1. In conclusion, there are few universities that can really prepare students for the real life of society needs.

2. ...a person pursuing a licenses degree to became a teacher, have only **few** courses during the four or five years of study...
3. If we continue accepting everything we may become a society controlled by **few** people, with no opinion and no power and this is certainly not Marx's idea about a society.
4. In the past, painters used to reproduce a real scene from the real world in their paintings, sometimes spending **a few** days to complete their work.
5. Places and things that would've not been acessible **a few** decades ago are now avaiable to us...
6. In fact, in general there are **a few** opportunities for everybody in our society. Moreover, if you are an ex-criminal your chances are lows...
7. I will put forward **a few** arguments for why university degrees are based in theory.
8. First of all we have to consider the fact that television is one of **a few** numbers of "cultural goods"

B.2.1. Write the number(s) of the sentences in which you believe *a few/few* were used inadequately:

Numbers: _____

B.2.2. If it were the case to make changes in sentences #1 and #3, which would you use before *few*, *a* or *very*? Justify.

#1: _____

#3: _____

B.2.3. Comparing the clusters *only few* and *very few*, the former generated 27 concordance lines in the academic section of COCA, whereas the latter generated 1.410. In sentence #2, is it possible to replace *only few* with *very few*? In your opinion, would this alteration improve the meaning conveyed in the sentence or change it?

B.2.4 In sentence #6, would you make any changes? Explain.

B.2.5. Taking into account that you read the concordance lines taken from COCA before answering the questions related to the sentences from Br-ICLE, can you say whether the concordance lines were helpful to you in your reasoning? If they were, say how.

5 CONCLUSÃO

Nesse capítulo, procuro delinear as considerações finais e conclusivas acerca dessa pesquisa, e apontar para desdobramentos futuros que podem abrir portas para outras buscas linguísticas dos quantificadores estudados. Biber, Conrad e Reppen (1998) afirmam que uma investigação acerca dos desvios encontrados nas produções de aprendizes aumenta nossa compreensão sobre a aquisição de uma língua estrangeira, além de fornecer dados para outras perspectivas relacionadas às inadequações cometidas – em grande parte provenientes de questões de interlíngua e variedades não padrão da língua-alvo – e informar sobre evidências para decisões instrucionais. Assim sendo, fazer comparações entre *corpora* representativamente equivalentes – como foi feito nesse trabalho – em muito pode contribuir para a melhor compreensão do que ocorre nas produções dos aprendizes.

Retomo, portanto, os meus objetivos de pesquisa e as perguntas feitas no capítulo introdutório. Quanto à descrição de como os aprendizes de inglês em nível intermediário-avanzado utilizam os quantificadores *a few* e *few*, foi observado que há mais acertos do que erros – 67,6% de acertos e 32,4% de erros – o que aponta para um uso satisfatório dos itens linguísticos em questão não somente em nível sintático, mas também em nível colocacional e pragmático, uma vez que a averiguação das inadequações levou em conta todas essas noções. Um exemplo disso é que a colocação *in few years*, aparentemente de acordo com o que se espera em termos sintáticos, não ocorre no uso corrente da linguagem dos nativos. Logo, *in few years* foi uma colocação classificada como inadequada, e tal classificação é respaldada pela análise dos colocados no contexto expandido e pela frequência dessa combinação de palavras nos *corpora* de nativos utilizados como referência nessa pesquisa. Outro exemplo foi o uso inadequado de *a few* em um contexto que, pragmaticamente, deveria ser restritivo e colorido de uma **prosódia semântica negativa**, requerendo, portanto, o uso de *few* ou *very few*. Tais achados nas produções dos aprendizes apontam para falhas no uso dos quantificadores que podem ser trabalhadas, e até mesmo sanadas, por meio de intervenções pedagógicas baseadas na AMD de Johns, e concretizadas em meio ao auxílio dos recursos da LC. Em relação à descrição do uso dos quantificadores feito pelos nativos, o ponto de mais alta relevância foi conseguir identificar as colocações mais comuns das quais eles lançam mão em registro acadêmico e dividi-las em domínios semânticos, separando, por exemplo, os colocados relacionados a tempo e os colocados relacionados a pessoas. Foi válido perceber que existe uma predominância de uso de *a few* em contextos temporais, e de *few* para quantificar pessoas. Com todos esses achados em mão, foi possível realizar comparações bem

fundamentadas e enxergar mais claramente os pontos de melhorias que precisam ser refinados na produção dos aprendizes. Ao mesmo tempo, foi possível perceber que os aprendizes possuem um bom domínio de certas colocações, principalmente temporais, como *over the past few years*. Desse modo, cumprindo com o que foi nomeado como objetivos, a descrição de como ambos os aprendizes e os nativos utilizam os quantificadores *a few* e *few* foi feita, as inadequações – e também as adequações – foram comentadas, levando em conta traços gramaticais e de **prosódia semântica**, e foi proposta uma atividade pedagógica baseada em *corpus*, direcionada à promover uma melhor compreensão do uso dos quantificadores pela observação de linhas de concordância (ver seção 4.5 do cap. 4).

No tocante às perguntas de pesquisa em si, foi observado que, respondendo à primeira, há sim semelhanças entre o uso que os aprendizes intermediário-avançados e os nativos fazem dos quantificadores *a few* e *few*. Como já afirmado, os aprendizes obtiveram um bom índice de acertos, e esses acertos estão baseados na comparação com a produção dos nativos. Todavia, foi identificada resistência de uso, por parte dos aprendizes, da colocação, *very few*, sendo constatado, portanto, **subuso** dessa colocação. Prováveis explicações para esse comportamento que aponta para **subuso** estão ligadas à segunda pergunta de pesquisa. Foi observado na produção dos nativos que há emergência de **prosódia semântica negativa** no uso da colocação *very few*, e isso não foi satisfatoriamente capturado pelos aprendizes. Tal colocação obteve uso mais expressivo por parte dos nativos, e em contextos semelhantes aos verificados no LOCNESS, os participantes do Br-ICLE não utilizaram tal padrão linguístico. Isso pode estar relacionado aos padrões costumeiros na língua materna, que possivelmente exerce influência nas escolhas lexicais dos aprendizes, já que, nela, não se verifica um vasto uso da colocação correspondente **muito poucos(as)**, e também pode estar ligado ao fato de que, conforme Zhang (2010, p. 193),

o conhecimento da prosódia semântica deve ser transferido aos aprendizes de segunda língua. Devido a sua natureza de sutilezas, a prosódia semântica fica normalmente escondida da intuição humana, só podendo ser explorada pelos poderosos meios da Linguística de Corpus. Portanto, a AMD (Aprendizagem Movida por Dados) como abordagem para aprender e ensinar itens lexicais dentro de um contexto deve ser proposta para que aprendizes possam se tornar cientes dos traços de prosódia semântica de uma palavra. Assim, haverá menos ocorrências de prosódia semântica inapropriada e de colocações pouco frequentes em suas produções em inglês. (tradução minha)³⁸

³⁸ “the knowledge for SP should be transferred to second language learners. Due to its nature of subtleness, SP is often hidden from human intuition and so can only be explored by the powerful means of corpus linguistics. Therefore, a DDL (data-driven learning) approach to learn or teach lexis in context should be proposed so that the learners could be aware of the semantic prosodic features of a word. Then there will be fewer occurrences of inappropriate SP and unusual collocation in their English.”

Por isso, na tentativa de levar adiante a observação acima feita por Zhang, esse trabalho trouxe a proposta de atividade que recupera os pressupostos da AMD. Tal atividade pode ser futuramente aplicada e testada, abrindo caminhos para que uma próxima pesquisa seja conduzida com o intuito de avaliar os resultados de sua aplicação.

Outras buscas quanto ao uso de quantificadores na linguagem dos nativos e na interlíngua dos aprendizes de inglês também podem ser motivadas por meio dessa pesquisa, como, entre outras, a comparação entre o uso dos quantificadores *some* e *a few*, geralmente tidos como similares ou quase-sinônimos, e a confecção de uma listagem mais completa e robustamente analisada de expressões com quantificadores ligadas a tempo, usadas pelos dois grupos, nativos e aprendizes.

Para concluir, menciono que embora esse trabalho tenha primariamente sido voltado para aspectos de descrição linguística, fica o apelo para pesquisadores se voltarem também para uma maior preocupação com a faceta pedagógica. Shepherd (2009a, p. 112), por exemplo, destaca a existência do grupo GELC (Grupo de Estudos em Linguística de Corpus) que foi responsável pela produção de trabalhos acadêmicos que promovem a interface LC e prática pedagógica. Entretanto, “a despeito desses esforços, os trabalhos com *corpus* de aprendiz e *corpora* em geral permanecem na área descritiva e fazem pouco investimento no pedagógico.” (SHEPHERD, 2009a, p. 112). A ideia, portanto, é que esse trabalho sirva de estímulo para o aquecimento de pesquisas baseadas em *corpus* e que sejam pedagogicamente motivadas e orientadas.

REFERÊNCIAS

- AARTS, J.; GRANGER, S. Tag sequences in learner corpora: a key to interlanguage grammar and discourse. In: GRANGER, S. (Ed.) *Learner English on Computer*. London, New York: Addison Wesley Longman, 1998, p. 132-142.
- ALMEIDA, V. C. *Investigando colocações em corpus de aprendizes*. III Congresso Internacional da ABRAPUI, anais. Universidade Federal de Santa Catarina, Florianópolis, 2012, p. 57-69. Disponível em: <<http://www.abrapui.org/anais/ComunicacoesCoordenadasLingua/25.pdf>> Acesso em: 16 abr. 2014
- ALVES, C. F. A contribuição da análise contrastiva e de um corpus de erros para o desenho de um dicionário de espanhol para aprendizes brasileiros. *Cadernos de Letras (UFRJ)*, n.28, jul., p. 89 – 98, 2011.
- ALVES, D.; TAGNIN, S. E. O. Pela visibilidade da Linguística de Corpus em trabalhos acadêmicos. In: X ENCONTRO DE LINGUÍSTICA DE CORPUS: ASPECTOS METODOLÓGICOS DOS ESTUDOS DE CORPORA. DUTRA, D. P.; MELLO, H. R. (Org.). *Anais do X Encontro de Linguística de Corpus: aspectos metodológicos dos estudos de corpora*. Belo Horizonte: Faculdade de Letras da UFMG, 2012.
- ANDERBOIS, S.; BRASOVEANU, A.; HENDERSON, R. *The pragmatics of quantifier scope: A corpus study*. The Programatics of Quantifier Scope. University of Connecticut/UC Santa Cruz. s.d. Disponível em: <http://people.ucsc.edu/~abrs/vn/SUB16_quant_scope.pdf> Acesso em: 14 abr. 2014.
- ANDERSON, W.; CORBETT, J.; *Exploring English with Online Corpora*. Palgrave Macmillan. 2009.
- BARLOW, M. Corpora for theory and practice. *International Journal of Corpus Linguistics*, v.1, n. 1, p. 1-37, 1996.
- BARWISE, J.; COOPER, R. Generalized Quantifiers and Natural Language. *Linguistics and Philosophy*, v. 4, n. 2, p. 159-219, 1981.
- BENNETT, G. R. *Using corpora in the language learning classroom: corpus linguistics for teachers*. Michigan: The University of Michigan Press. 2010.
- BERETTA, J. M. *A correção de erros: inimiga ou aliada? A reescrita como estratégia de correção e de conscientização na produção escrita de aprendizes de inglês como língua estrangeira*. 2001, 175f. Dissertação (Mestrado). Universidade Federal do Rio Grande do Sul. Universidade de Caxias do Sul. Mestrado Interinstitucional em estudos da linguagem. 2001.
- BERBER SARDINHA, T. Linguística de corpus: histórico e problemática. Retrospectiva. *D.E.L.T.A.*, v.16, n. 2, p. 323-367, 2000.
- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri, SP: Manole, 2004.
- BIBER, D.; JOHANSSON, S.; LEECH, G.; CONRAD, S.; FINEGAN, E. *Longman Grammar of Written and Spoken English*. London: Pearson Education Limited. 1999.

BIBER, D.; CONRAD, S.; LEECH, G. *Longman Student Grammar of Spoken and Written English*. Essex: Pearson Education Limited. 2002.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press, 1998, 300 p.

BR-ICLE. *The Brazilian Portuguese sub-corpus of ICLE, the International Corpus of Learner English*. PUC/SP. Disponível em: < <http://www2.lael.pucsp.br/corpora/bricle/>>. Acesso em: 7 de abril de 2014, 2014.

BRAUN, S.; KOHN, K.; MUKHERJEE, J. (Ed.) *Corpus technology and language pedagogy: new resources, new tools, new methods*. Frankfurt/M: Peter Lang, 2006, p. 25–47.

BRAUN, S. *Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora*. ReCALL19(3). UK: Cambridge University Press. 2007.

BOD, R.; HAY, J.; JANNEDY, S. (Ed.) *Probabilistic Linguistics*. Washington: MIT Press. 2003.

CAMBRIDGE, University Press. *Cambridge Learner Corpus*. Disponível em: http://www.cambridge.org/gb/elt/catalogue/subject/custom/item3646603/Cambridge-English-Corpus-Cambridge-Learner-Corpus/?site_locale=en_GB>. Acesso em: 04 abr. 2014.

CEGALLA, Domingos Paschoal. *Novíssima Gramática da Língua Portuguesa*. 48. ed. São Paulo: Companhia Editora Nacional, 2008.

CIPRO NETO, P.; INFANTE, U. *Gramática da Língua Portuguesa*. 2. ed. São Paulo: Editora Scipione, 2004.

CHOMSKY, Noam. Three models for the description of language. *IRE Transactions on Information Theory*, v. 2, n. 3, p. 113-124, 1956.

CONRAD, S. Will Corpus Linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, v. 34, n. 3, p. 548-560, 2000.

CORPUS DO PORTUGUÊS DA. *O corpus do português*. Disponível em: < <http://www.corpusdoportugues.org/>>. Acesso em: 7 jan. 2014.

COWIE, A. P. Phraseology. In: ASHER R.; SIMPSON, J. (Org.) *The encyclopedia of language and linguistics*. Oxford: Pergamon Press, 1994, p.168-171.

D'ALFONSO, D. Generalized Quantifiers: logic and language. *L&PS - Logic and Philosophy of Science*. v. IX, n. 1, p. 85-94. 2011.

DAVIES, Mark. *Corpus of Contemporary American English (COCA)*. Disponível em: < www.americancorpus.org>. Acesso em: 7 jan. 2014.

DELEGÁ-LUCIO, D. *A relexicalização de adjetivos nas redações de alunos de inglês: um estudo baseado em corpus de aprendiz*. 2006, 105f. Dissertação (Mestrado). Pontifícia Universidade Católica de São Paulo. Linguística Aplicada e Estudos da Linguagem. 2006.

- DUNNING, T. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, v. 19, n. 1, p.61–74, 1993.
- DUTRA, D. P. Conscientização Linguística com base em *corpora* online. *Revista Intercâmbio*. São Paulo: LAEL/PUC-SP, v. XX, p. 79-98, 2009.
- DUTRA, D. P.; SILERO, R. P. O uso de *for*: uma análise de itens linguísticos em *corpus* de aprendizes brasileiros. In: SHEPHERD, T. M. G.; SARDINHA, T. B.; PINTO, M. V. (Org.) *Caminhos da Linguística de Corpus*. São Paulo: Mercado de Letras, 2009, p. 325-342.
- FERNÁNDEZ, Y. R. A presença de erros na interlíngua de estudantes brasileiros aprendizes de espanhol. *Revista Desempenho*, v. 12, n. 1, jun., p. 1-27, 2011.
- FIRTH, J. R. A Synopsis of Linguistic Theory, 1930-1955. In: LIU RUNQING *et al.* (Eds). *Readings in Linguistics: Seventy-Five Years since Saussure*. Beijing: Cehui Press, v. 1, p. 328-359, 1957.
- FISHER EXACT TEST. Disponível em: < <http://www.physics.csbsju.edu/stats/fisher.form.html>>. Acesso em: 07 jan. 2014.
- FOX, G. Using corpus data in the classroom. In: TOMLINSON, B. *Materials development in language teaching*. Cambridge: Cambridge, 1998, p. p. 25-43.
- FRIES, Charles C. *Teaching and Learning English as a Foreign Language*. Ann Arbor, Mich.: The University of Michigan Press. 1945.
- GARDNER, H. E. *The mind's new science: A history of the cognitive revolution*. New York: Basic Books. 1985.
- GRANGER, S. *Learner English on Computer*. London: Longman. 1998.
- GRANGER, S.; TRIBBLE, C. Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In: GRANGER, S. (Ed.) *Learner English on Computer*. London, New York: Addison Wesley Longman, 1998, p. 199-209.
- GRANGER, S. A bird's-eye view of learner *corpus* research. In: GRANGER, S.; HUNG, J.; PETCH-TYSON, S. (Ed.). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam, The Netherlands: John Benjamins. 2002, p. 3–33.
- GRANGER, S.; DAGNEAUX, E.. MEUNIER, F.; PAQUOT, M. (Ed.). *International Corpus of Learner English*. Louvain-la-Neuve: Presses Universitaires de Louvain. 2009, v. 2.
- GRANGER, S., SANDERS, C. & CONNOR, U. LOCNESS: *Louvain corpus of native English essays*. Disponível em: < <https://www.uclouvain.be/en-cecl-locness.html>> Acesso em: 7 jan. 2014
- GRANGER, S. *Internacional Corpus of Learner English – ICLE*. Centre for English Corpus Linguistics. Université Catholique de Louvain. Disponível em: <<http://www.uclouvain.be/en-cecl-icle.html>>. Acesso em: 7 de janeiro de 2014.

HALLIDAY, M. A. K. Corpus studies and probabilistic grammar. In: AIJMER, K.; ALTENBERG, B. (Org.). *English corpus linguistics: Studies in honour of Jan Svartvik*. London: Longman. 1991, p. 30-43.

HALLIDAY, M. A. K. Language as system and language as instance: the corpus as a theoretical construct. In: SVARTVIK, J. (Org.). *Directions in Corpus Linguistics: proceedings of Nobel Symposium 82*. Stockholm, v.4, n.8 Aug. 1991. Berlin, New York: De Gruyter. 1992.

HARMER, J. *How to teach English: an introduction to the practice of English language teaching*. Edingburgh: Longman, 1998.

HINTIKKA, J. Quantifiers vs. Quantification Theory. *Dialectica*, v. 27, n. 3-4, Dec., p. 329-358, 1973.

HUNSTON, S. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press. 2002.

HUNSTON, S. Semantic prosody revisited. In: MOON, Rosamund (Ed.). Words, grammar, text: revisiting the work of John Sinclair. *Special issue of International Journal of Corpus Linguistics*, v.12, n. 2, p. 249–268, 2007.

JIMÉNEZ CATALÁN, R. M. *Errores en la producción escrita del inglés y posibles factores condicionantes*. Madrid: Universidad Complutense. 1992.

JOHNS, T. From printout to handout: grammar and vocabulary teaching in the context of Data-Driven Learning. In: ODLIN, T. (Ed.). *Perspectives on Pedagogical Grammar*. Cambridge: Cambridge University Press, 1994, p. 293-313.

JOHNS, T. Data-driven learning: The perpetual challenge. In: KETTEMANN, B.; MARKO, G. (Ed.), *Teaching and learning by doing corpus analysis*. Amsterdam, The Netherlands: Rodopi, 2002, p. 107–117.

KEENAN, E.L. *Quantifiers: Semantics*. 2. ed. Encyclopedia of Language and Linguistics. Oxford: Elsevier, v. 10, p. 302 - 308. 2006.

KETTEMANN, B. On the use of concordancing in ELT. *Arbeiten aus Anglistik und Amerikanistik*, v.20, p. 29–41, 1995.

LABOV, W. The study of language in its social context. In: LABOV, W. *Sociolinguistic Patterns*. Philadelphia: University of Philadelphia Press, Capítulo 8, 1972, p. 183- 259.

LADO, R. *Linguistics across cultures: applied linguistics for language teachers*. Ann Arbor: University of Michigan Press, 1957.

LAKOFF, G. On Generative Semantics. In: STEINBERG, Danny D.; JAKOBOVITS, Leon A. (Org.). *Semantics: an interdisciplinary reader*. Cambridge University Press, Cambridge. 1971.

LEARNER CORPUS ASSOCIATION. *The Louvain corpus of native English essays (LOCNESS)*. Disponível em: < <http://www.learnercorpusassociation.org/resources/corpora/locness-corpus/>>. Acesso em: 7 jan. 2014

LARSEN-FREEMAN, D. *Teaching language: from grammar to grammaring*. Massachusetts: Thomson Heinle. 2003.

LEAL, António José Rodrigues. *Semântica aspectual e nominal: contributo das expressões nominais para a construção aspectual das frases*. 2009, 465f. Dissertação. Faculdade de Letras da Universidade do Porto. 2009.

LEECH, G. N. Students' grammar: teachers' grammar, learners' grammar. In: BYGATE, M. *et al.* (Eds.), *Grammar and the Language Teacher*. London: Prentice Hall. 1994, p. 17-29.

LEECH, G. N. Teaching and language corpora: A convergence. In: WICHMANN, A.; FLIGELSTONE, S.; MCENERY, T.; KNOWLES, G. (Ed.). *Teaching and language corpora*. London, UK: Longman. 1997, p. 1-23.

LEECH, G. N. Preface - Learner corpora: what they are and what can be done with them. In: GRANGER, S. *Learner English on Computer*. London and New York: Longman, 1998, p. xiv – xx.

LONGMAN, dictionary of contemporary English. *Few determiner, pronoun, adjective*. Disponível em: <<http://www.ldoceonline.com/dictionary/few>>. Acesso em: 9 jan. 2014b.

LONGMAN. *Longman corpus network*. Disponível em: <<http://www.pearsonlongman.com/Dictionaries/corpus/learners.html>>. Acesso em: 04 abr. 2014a.

LOUW, W. E. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In: BAKER, M.; FRANCIS, G.; TOGNINI-BONELLI, E. (Ed.). *Text and technology: in honour of John Sinclair*. Amsterdam: John Benjamins. 1993, p. 157-176.

LOUW, B. Contextual prosodic theory: Bringing semantic prosodies to life. In: HEFFER, C., SAUNTON, H., FOX, G. (Eds.), *Words in Context: A Tribute to John Sinclair on his Retirement*. Birmingham: University of Birmingham. 2000.

MARQUES, A. L. S. P. *O papel da percepção consciente de conectivos concessivos na produção escrita e na reformulação da interlíngua*. 2007, 174f. Dissertação. (Mestrado) Universidade Federal de Minas Gerais. Faculdade de Letras da Universidade Federal de Minas Gerais. 2007.

MATTE, A. C. F.; COSCARELLI, C. *UEADSL 2013 Liberdade na Internet*. Universidade, EAD e software livre. Disponível em: <<http://textolivre.pro.br/blog/?p=735>>. Acesso em: 10 mai. 2013.

MCENERY, T.; WILSON, A. *Corpus Linguistics*. 1. ed. Edinburgh: Edinburgh University Press, 1996.

MCENERY, T.; WILSON, A. *Corpus Linguistics: an introduction*. 2. ed. Edinburgh: Edinburgh University Press, 2001.

MCENERY, T.; HARDIE, A. *Corpus linguistics: method, theory and practice*. Cambridge, New York: Cambridge University Press. 2012.

MEUNIER, F. The pedagogical value of native and learner corpora in EFL grammar teaching. In: GRANGER, S.; HUNG, J.; PETCH-TYSON, S. (Ed.). *Computer learner corpora, second*

language acquisition and foreign language teaching. Amsterdam, The Netherlands: John Benjamins. 2002, p. 119–141.

NELSON, M. *A corpus-based study of business English and business English teaching materials*. Unpublished doctoral dissertation, University of Manchester. 2000.

NELSON, M. Semantic associations in business english: A corpus-based analysis. *English for Specific Purposes*, v.25, p. 217-234, 2006.

NESSELHAUF, N. Collocations in a learner corpus. *Studies Corpus Linguistics*. The Netherlands: John Benjamins, v. 14, n. 12, 2005.

NEVES, M. H. M. *Gramática de Usos do Português*. São Paulo: Editora Unesp. 2000.

O'KEEFFE, A.; MCCARTHY, M. (Ed.) *The Routledge Handbook of Corpus Linguistics*. Routledge. 2010.

O'KEEFFE, A.; MCCARTHY, M.; CARTER, R. *From corpus to classroom: language use and language teaching*. Cambridge: Cambridge University Press. 2007.

PARDO, Thiago A. S. *Introdução ao processamento de línguas naturais*. 2012. Disponível em: <<http://wiki.icmc.usp.br/images/1/18/Aula3-5908t-2012.pdf>>. Acesso em: 24 abr. 2014.

PARTINGTON, A. *Patterns and meanings: using corpora for English language research and teaching*. Philadelphia, PA: John Benjamins. 1998.

PARTINGTON, A. Utterly content in each other's company: semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, v.9, n. 1, p. 131-156, 2004.

PETERS, Stanley; WESTERSTAHL, Dag. *Quantifiers in Language and Logic*. Oxford: Clarendon Press. 2006.

PINTO, M. V. *O uso de things, anything, something e everything em corpora de aprendiz*. 2008, 221 f. Dissertação (Mestrado). Pontifícia Universidade Católica de São Paulo. Programa de Pós-Graduação em Linguística Aplicada ao Ensino de Línguas, PUC-SP, São Paulo. 2008.

PINTO, J. A. Algumas concepções de Lógica (Aristóteles, Kant e Frege): verdade, auto-conhecimento e justificação. In: PINTO, J. A. (Ed.) *Investigações em Filosofia da Lógica*. Coleção MLAG Discussion Papers. Porto, FL-UP, v.5, 2011.

RÖMER, U. Corpus Research Applications in Second Language Teaching. *Annual Review of Applied Linguistics*, v. 31, Mar., p. 205-225, 2011.

RÖMER, U.; WULFF, S. Applying corpus methods to writing research: Explorations of MICUSP. *Journal of Writing Research*, v.2, n. 2, p. 99-127, 2010.

RUNDELL, M.; GRANGER, S. *Form corpora to confidence*. English Teaching Professional Magazine. Issue 50. 2007. Disponível em: <<http://www.macmillandictionaries.com/wp-content/uploads/2011/04/Fromcorporatoconfidence-EnglishTeachingProfessional-May2007.pdf>> Acesso em: 10 abr. 2014.

RUZAITÉ, J. Parallel Corpora for applied purposes: a case study of quantifiers in english and lithuanian. *EESTI Rakenduslingvistika Ühingu Aastaraamat*, v. 5, p. 239–250, 2009.
Disponível em: < <http://rakenduslingvistika.ee/ajakirjad/index.php/Aastaraamat/article/view/ERYa5.16/100>> Acesso em: 14 abr. 2014.

SANTOS GARGALLO, I. *Análisis contrastivo, análisis de errores e interlengua en el marco de la lingüística contrastiva*. Madrid: Editorial Síntesis, 1993.

SAUSSURE, Ferdinand de. Matérias e tarefas da lingüística; suas relações com as ciências conexas. In: SAUSSURE, Ferdinand de. *Curso de Linguística Geral*. São Paulo: Editora Cultrix, Capítulo II, 2006, p. 13-14.

SCHMIDT, R.; FROTA, S. Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese. In: DAY, R. R. (Ed.) *Talking to learn: conversation in second language acquisition*. Rowley: Newbury House, 1986, p. 237-326.

SCHMIDT, R. Attention, awareness, and individual differences in language learning. In: CHAN, W. M.; CHI, S.; CIN, K. N.; INSTANTO, J.; NAGAMI, M.; SEW, J. W.; SUTHIWAN, T.; WALKER, I. *Proceedings of CLaSIC, Singapore*. Singapore: National University of Singapore, v. 2, n. 4, dec., p. 721-737, 2010.

SEIDLHOFER, B. Pedagogy and local learner corpora: working with learning-driven data. IN: GRANGER, S.; HUNG, J.; PETCH-TYSON, S. (Ed.). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins, 2002, p. 213–234.

SHEPHERD, T. M. G. O Estatuto da linguística de corpus: metodologia ou área da linguística? *Matraga*. Rio de Janeiro, v.16, n. 24, jan./jun., p. 150–172, 2009b.

SHEPHERD, T. M. G. Corpora de aprendiz de língua estrangeira: um estudo contrastivo de n-gramas. Veredas On-Line – *Linguística de Corpus e Computacional*. PPG Linguística/UFJF, v. 2, p. 100-116, 2009a.

SINCLAIR, J. M. *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press. 1991.

SINCLAIR, J. M. The Search for Units of Meaning. *TEXTUS*, v. 9, n. 1, p. 75-106. 1996.

SONG, L. On the Variability of Interlanguage. *Theory and Practice in Language Studies*, v. 2, n. 4, Apr., p. 778-783, 2012.

STUBBS, M. Collocations and semantic profiles: on the cause of trouble with quantitative studies. *Functions of Language*, v.2, n. 1, p. 23-55. 1995.

STUBBS, M. *Text and corpus analysis: computer-assisted studies of language and culture*. Oxford: Blackwell. 1996.

SWAN, Michael. *Practical English Usage*. 3. ed. New York: Oxford University Press. 2005.

SZABOLCSI, A. *Quantification*. New York: Cambridge University Press. 2010.

TENUTA, A. M.; OLIVEIRA, A. L. A. M.; ORFANÓ, B. M. How Brazilian learners express modality in their writing: a corpus-based study on lexical bundles. *Revista Intercâmbio*. São Paulo: LAEL/PUCSP, v. XXVI, p.1-15, 2012.

TOGNINI-BONELLI, E. *Corpus linguistics at work*. Amsterdam: Benjamins. 2001.
TOTTIE, G; NEUKOM-HERMANN, A. Quantifier-negation interaction in English: a corpus linguistic study of all...not constructions. In: HORN, L. R. *The Expression of Negation*. Berlin, New York, 2010, p.149-185.

TRIBBLE, C. Genres, keywords, teaching: towards a pedagogic account of the language of project proposals. In: BURNARD, L.; MCENERY, T. (Ed.). *Rethinking language pedagogy from a corpus perspective: papers from the third international conference on teaching and language corpora*. New York: Peter Lang. 2000, p. 74-90.

XIAO, Z.; MCENERY, A. Collocation, semantic prosody and near synonymy: a cross-linguistic perspective. *Applied Linguistics*, v. 27, n. 1, Mar., p. 103-129, 2006.

WANG, A. Y.; PIAO, S. Translating vagueness? A study on translations of vague quantifiers in an english-chinese parallel corpus. *In the Proceedings of the Corpus Linguistics Conference*. Birmingham, UK. 2007. Disponível em: < http://ucrel.lancs.ac.uk/publications/CL2007/paper/278_Paper.pdf > Acesso em: 14 abr. 2014.

WHITSITT, S. A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics*, v. 10, n. 3, p. 283–305, 2005.

ZHANG, C. An Overview of Corpus-based Studies of Semantic Prosody. *Asian Social Science*, v. 6, n. 6, Jun., p. 190- 194, 2010.

ZHANG, W. Semantic Prosody and ESL/EFL Vocabulary Pedagogy. *TESL. Canada Journal/Revue TESL du Canada*, v. 29, n. 2, Spring., p. 1- 12, 2009.

ZIPF, G. K. *Human Behavior and the Principle of Least Effort*. Addison-Wesley, 1949.

ANEXOS A – ABOUT THE CAMBRIDGE LEARNER CORPUS

What is unique about the Cambridge Learner Corpus?

Specialists at Cambridge University Press carefully check each exam script and **highlight all errors** made by students. We can then use this information to see which words or structures are difficult for learners of English.

The Cambridge Learner Corpus also contains information about the student's first language, nationality, level of English, grade, age, gender, and date of exam. Along with the error information, this means that we can:

- Focus on certain groups of learners and see what they find easy or hard.
- Make sure our materials contain appropriate content for a particular level or exam.
- Find mistakes which universal to English language learning, and those which are a result of first-language interference.
- Find plenty of examples of language used by students and use this to help other students.

Figura 6 – About the Cambridge Learner Corpus
Fonte: CAMBRIDGE, 2014

ANEXO B – ABOUT THE CAMBRIDGE LEARNER CORPUS

● HOW DO WE USE THE INFORMATION?

The Longman Learners' Corpus offers so much invaluable information about the mistakes students make and what they already know, that it is the perfect resource for lexicographers and material writers who want to produce dictionaries and coursebooks that address students' specific needs. The Longman Learners' Corpus was used to write the Usage Notes in the Longman Active Study Dictionary. Evidence from the corpus showed that there was 100% error in the meaning and the use of the word **cloth**: **My cloths and shoes were wet, We have very good cloth stores** etc. Having pinpointed such a problem, our lexicographers were then able to write a corresponding Usage Note (fig. 3).

FIG 3.

USAGE Do not use **cloth** or **cloths** to mean "the things that people wear". Instead use **clothes** *a clothes shop. / The guests all wore casual clothes.*

Figura 7 – About the Longman Learner Corpus
Fonte: LONGMAN, 2014a

ANEXO C – LONGMAN STUDENT GRAMMAR (MODERATE OR SMALL QUANTITY)

C Moderate or small quantity

Some usually denotes a moderate quantity and is used with countable and uncountable nouns:

*Insurance shares produced **some** excitement.* (NEWS)

***Some** performance curves will now be presented.* (ACAD†)

We will go into the contrast between *some* and *any* in 8.8.10–11.

Determiners denoting a small quantity are:

	plural countable	uncountable
a small number/amount less than expected	<i>a few, several</i> <i>few (fewer, fewest)</i>	<i>a little, a bit of</i> <i>little (less, least)</i>

A few and *a little* are used to indicate a small amount:

*With **a little** care he had no difficulty whatever in putting his glass back on the table.* (FICT)

*There were **a few** people sitting at the tables in the back.* (FICT†)

Few and *little* (without *a*) mean 'not many' and 'not much'.

*Very **few** women have hair that's that short.* (CONV†)

*That's why I dislike plans because so much time is spent planning and so **little** time is spent doing anything.* (CONV†)

Figura 8 – Longman Student Grammar (Moderate or small quantity)

Fonte: BIBER; CONRAD; LEECH, 2002, p. 75

ANEXO D – PRACTICAL ENGLISH USAGE (USE OF A)

3 use of a

There is a difference between *little* and *a little*, and between *few* and *a few*. Without *a*, *little* and *few* usually have rather negative meanings. They may suggest 'not as much/many as one would like', 'not as much/many as expected', and similar ideas.

The average MP has little real power.

Few people can speak a foreign language perfectly.

A little and *a few* are more positive: their meaning is generally closer to *some*. They may suggest ideas like 'better than nothing' or 'more than expected'.

Would you like a little soup?

You don't need to go shopping. We've got a few potatoes and some steak.

Compare:

- *Cactuses need little water.* (not much water)
- *Give the roses a little water every day.* (not a lot, but some)
- *His ideas are difficult, and few people understand them.*
- *His ideas are difficult, but a few people understand them.*

Quite a few (informal) means 'a considerable number'.

We've got quite a few friends in the village.

4 formal and informal language

Little and *few* (with no article) are rather formal. In an informal style (e.g. ordinary conversation), we generally prefer *not much/many*, or *only a little/few*.

Come on! We haven't got much time!

Only a few people speak a foreign language perfectly.

However, *very little* and *very few* are possible in an informal style.

He's got very little patience and very few friends.

Figura 9 – Practical English Usage (Use of *a*)

Fonte: SWAN, 2005, p. 309

**ANEXO E – LP (FORMULÁRIO PREENCHIDO PELOS PARTICIPANTES DO
BR-ICLE)**

LEARNER PROFILE

Text code : (do not fill in)

Essay :

Title :

Approximate length required : -500 words +500 words

Conditions : timed untimed

Examination : yes no

Reference tools : yes no

What reference tools ?

Bilingual dictionary :

English monolingual dictionary :

Grammar :

Other(s) :

Surname :

First names :

Age : Male

Female

Nationality :

Native language :

Father's mother tongue :

Mother's mother tongue :

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

Education :

Primary school - medium of instruction :

Secondary school - medium of instruction :

Current studies :

Current year of study :

Institution :

Medium of instruction :

English only

Other language(s) (specify)

Both

Years of English at school :

Years of English at university :

Stay in an English-speaking country :

Where ?

When ?

How long ?

Other foreign languages in decreasing order of proficiency :

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

Date :

Signature : ...

ANEXO F – PARTE INICIAL DA CURVA DE ZIPF PARA A DISTRIBUIÇÃO DE PALAVRAS NO ROMANCE *MOBY DICK*

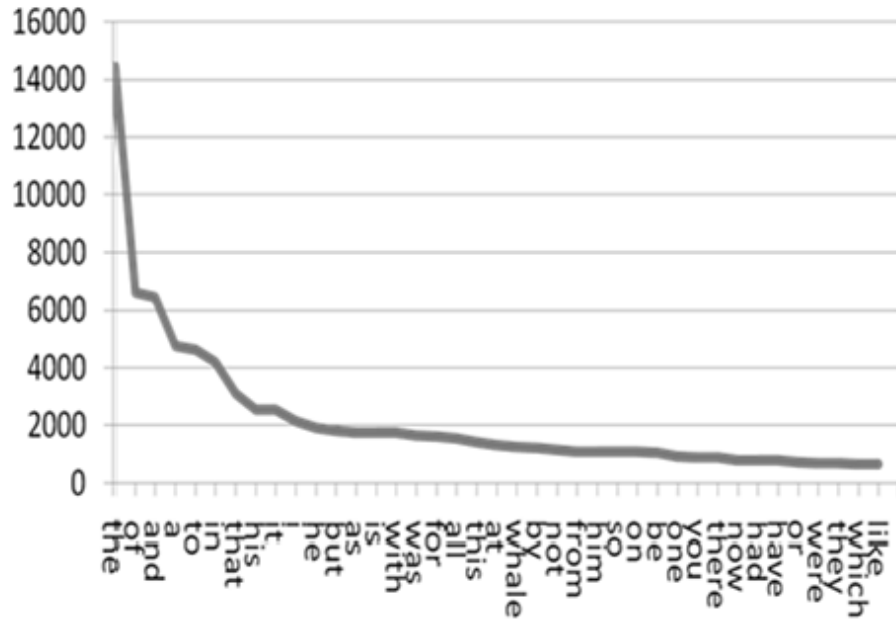


Gráfico 3 – Parte inicial da curva de Zipf para a distribuição de palavras no romance *Moby Dick*
 Fonte: PARDO, 2012.

ANEXO G – LINHAS DE CONCORDÂNCIA DO BR-ICLE CUJO NÓDULO É A *FEW* E OS COLOCADOS À DIREITA POSSUEM SENTIDO TEMPORAL

Quadro 30 – As linhas de concordância do Br-ICLE cujo nóculo é *a few* e os colocados à direita possuem sentido temporal

3 || real world in their paintings, sometimes spending a few days to complete their work. With the invention o
 4 || laces and things that wouldve not been acessible a few decades ago are now avaisible to us through the to
 5 || an sit and study for hours, other must stop after a few hours, or minutes, have a break and continue. I'v
 8 || lings, because we are loosing the power of words. A few moments after the birth, the child is already int
 9 || avorite programs, problems are forgotten even for a few minutes. Moreover, today the programs and TV seri
 10 || f information, which probably they will forget in a few time, and they do not practice, so they do not kn
 13 || ive means which were so enthusiastically welcomed a few decades ago. As a result, it is common to see you
 15 || vão, a nationwide exam created by the government a few years ago with the objective of evaluating the qu
 16 || st in professions that require a technical level. A few years ago the job market favored those who had go
 23 || imagination to live happy or escape the world for a few moments. Keep dreaming and using imagination, it
 24 || for anything with a broken card. Usually it takes a few days to replace a card after a phone call or a pe
 26 || acts through rehabilitation. Both are necessary. A few years ago, the Brazilian TV show Fantástico sho

Fonte: Gerado pelo programa AntConc 3.2.1.

ANEXO H – LINHAS DE CONCORDÂNCIA DO BR-ICLE CUJO NÓDULO É *FEW* E OS COLOCADOS À DIREITA POSSUEM SENTIDO TEMPORAL

Quadro 31 – As linhas de concordância do Br-ICLE cujo nóculo é *a few* e os colocados à direita possuem sentido temporal

1 ||rgent for the mankind do not face consequences in *few* years.

23 ||be one of the most important conquest in the last *few* centuries. fileseparator <BRDEN004> <Feminism: th

42 ||ations this certainly would not be possible in so *few* time. All things considered we should not say tha

46 ||t that was possible to realize by studies on past *few* years is the increase of cancer on people who liv

48 ||an I believe that we are going to over pass it in *few* years.

51 ||aculous life will be in a hundred years from now. *Few* years ago, society dreamt that by the year two th

52 ||eligion. This idea has been altered over the past *few* years and now religion is not societys aim. Rel

66 || improve quality of live. Nowadays, diseases that *few* decades ago were considered fatal can be quickly

68 ||r concern because of poor responses over the past *few* years. In summary, the relationship between AIDS

69 || <Living in a high tech world> Over the past *few* decades, technology and the science have been sev

Fonte: Gerado pelo programa AntConc 3.2.1.