

Uso de *corpora* em aulas de prática de redação de língua inglesa

Adriane Orenha-Ottaiano
UNESP – São José do Rio Preto
adriane@ibilce.unesp.br

Paula Tavares Pinto Paiva
UNESP – São José do Rio Preto
paula@ibilce.unesp.br

Resumo

Neste estudo visamos a discutir aspectos relacionados aos *corpora* de aprendizes, assim como a descrever algumas características linguísticas encontradas em textos escritos por aprendizes da língua inglesa, no que diz respeito ao uso de colocações na produção de texto. Para esta pesquisa, analisamos as colocações empregadas com o verbo *to have* e com os substantivos *prejudice* e *regret*.

Palavras-chave: *linguística de corpus; corpora de aprendizes; colocações.*

Abstract

This study aims at discussing aspects related to learner corpora and linguistic features found in texts written by English learners based on the use of collocations in text production. For this research, we analyzed collocations with the verb “to have” and with the nouns “prejudice” and “regret”.

Keywords: corpus linguistics; learner corpora; collocations

1. Introdução

A observação e a descrição da linguagem em uso em diferentes contextos, como no de ensino e no de aprendizagem de línguas, têm

sido favorecidas pela Linguística de *Corpus*¹, em virtude da abordagem teórico-metodológica de caráter interdisciplinar dessa área. Berber Sardinha (2004) destaca as vantagens que a Linguística de *Corpus* oferece tanto para o professor quanto para o aluno, por permitir o acesso à tecnologia computacional e à linguagem utilizada nos textos contidos em um determinado *corpus*. Tal pressuposto deve-se à proposta de Sinclair (1978, *apud* BERBER SARDINHA, 2004) ao afirmar que a língua deveria ser estudada com base em exemplos reais de uso e não somente por meio de textos criados com propósitos de exemplificação.

Em conformidade com as afirmações destacadas anteriormente, propomos um trabalho, iniciado em 2005, no qual analisamos textos em língua inglesa produzidos por alunos brasileiros do curso de Letras com Habilitação em Tradução e Interpretação de uma universidade privada do interior paulista, com a finalidade de verificar: a) os padrões linguísticos da língua-alvo empregados com maior frequência (*sobreuso*) ou menor frequência (*subuso*) quando comparados à produção de falantes nativos de língua inglesa e b) se há influência da língua materna na utilização das colocações empregadas por aprendizes brasileiros textos de língua inglesa elaborados por eles.

Ressaltamos que a abordagem utilizada para esta pesquisa inicial, de caráter longitudinal (visto que os dados foram coletados durante quatro anos consecutivos) visa à identificação de específicos padrões linguísticos que foram repetidamente utilizados pelos alunos da disciplina de *Produção e Interpretação de Textos em Língua Inglesa*. Esses padrões poderiam nos auxiliar, de modo empírico, a traçar as dificuldades de emprego de vocabulário e de colocações em língua inglesa por parte dos alunos brasileiros. O objetivo não foi, portanto, o de avaliar as produções individualmente, prática que também faz parte da disciplina, mas não desta pesquisa.

¹De acordo com Sanchez (1995, p. 8-9, *apud* BERBER SARDINHA, 2000, p. 338) um *corpus* é “um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise.”

Em uma fase inicial, observamos, no *corpus* de aprendizes investigado com o auxílio do programa *WordSmith Tools* (SCOTT, 2004), um sobreuso do artigo definido *the*, assim como um subuso de advérbios, nos textos produzidos pelos alunos da referida disciplina ministrada pelas professoras/pesquisadoras, conforme detalhamos no item Metodologia. Uma vez constatadas tais características, propusemos, para as turmas seguintes, um trabalho de leitura e de escrita de diferentes gêneros textuais, além de uma verificação das linhas de concordância² com base em *corpora* compostos por textos originalmente escritos em língua inglesa, a fim de aumentar a variação lexical utilizada nas produções dos alunos e, possivelmente, minimizar as características de sobreuso e subuso dos itens observados no estudo.

Dessa maneira, este trabalho inicial teve o objetivo de incentivar os alunos a empregar uma variedade maior de vocabulário e de colocações, de maneira a contribuir para a produção de textos mais fluentes, além de conscientizá-los sobre os aspectos gramaticais, sintáticos e semânticos, de modo a favorecer um aprendizado mais autônomo por meio da observação de textos originalmente escritos em língua inglesa.

O trabalho ora proposto tem por finalidade mostrar algumas análises e resultados colhidos posteriormente e apresentar uma discussão para futuros procedimentos. Para tanto, apresentamos a fundamentação teórica que permeia este trabalho e propomos, em seguida, uma reflexão sobre como textos em língua inglesa escritos por alunos poderiam ter seu conteúdo léxico e sintático mais adequado com base na análise do *corpus* de aprendizes e na utilização de *corpora* compostos por textos originalmente escritos em língua inglesa. Essa análise foi realizada com o auxílio do programa *WordSmith Tools* (SCOTT, 2004).

² Linhas de concordância são todas as linhas de textos destacadas de um *corpus* computadorizado que tenham como referência uma palavra de busca, conforme poderá ser visto nas pesquisas 1 e 2.

2. Revisão da literatura

Ao descrever o uso de *corpora* de aprendizes compostos por textos produzidos por alunos de uma determinada língua materna ou estrangeira, Berber Sardinha (2004) destaca quatro áreas majoritárias que têm sido favorecidas com o uso de *corpora* ou, mais especificamente, de *corpora* de aprendizes: descrição da linguagem nativa; descrição da linguagem do aprendiz; transposição de metodologias de pesquisa acadêmica para a sala de aula e desenvolvimento de materiais de ensino, currículos e abordagens.

O autor enfatiza que os *corpora* desempenham um papel importante na conscientização do aprendiz, uma vez que lhe permitem visualizar e compreender diversas possibilidades existentes no momento de realizar escolhas linguísticas.

A utilização de *corpora* não eletrônicos como auxílio na produção de materiais didáticos e no ensino de línguas não é tão recente (THORNDIKE, 1921; WEST, 1953). No entanto, a pesquisa em *corpora* de aprendizes é uma atividade nova e tem tomado impulso recentemente (TRIBBLE, 1990, 1991; GRANGER, 1993, 1998; NICHOLLS, 2003, ORENHA, PAIVA e CAMARGO, 2007; VIANA e TAGNIN, 2011). *Corpora* de aprendizes são constituídos por textos de falantes não nativos, os quais são utilizados para o estudo da linguagem que produzem. De acordo com Haan (1992), o tamanho ideal de um *corpus* de aprendizes depende do tipo de pesquisa a ser realizada. No entanto, segundo o autor, um *corpus* de vinte mil palavras pode ser suficiente para se realizar um estudo. *Corpora* de aprendizes que variam de vinte a duzentas mil palavras costumam ser mais específicos em relação a *corpora* maiores, no que se refere aos seus tópicos e gêneros discursivos. Granger (1998), em *Learner English on Computer*, defende que duzentas mil palavras seja o número ideal para a compilação de um *corpus* de aprendizes, como foi o caso do ICLE (*International Corpus of Learner English*), inicialmente proposto pela pesquisadora e que, hoje, conta com três milhões e meio de palavras produzidas por seis mil aprendizes de 16 países diferentes (MEURIER, 2011). Conforme a autora, *corpora* de aprendizes maiores seriam mais destinados a pesquisadores que visam, por exemplo, à compilação de dicionários para aprendizes de uma língua estrangeira (LE). Entre os *corpora* de aprendizes compilados

para a elaboração de dicionários podemos citar o *Cambridge Learner Corpus* e o *Longman Learner Corpus* que, no entanto, possuem acesso restrito.

No Brasil, já existem pesquisas sendo realizadas com *corpora* de aprendizes, entre elas, destacamos o Br-ICLE, sub*corpus* do ICLE, organizado pelo Prof. Dr. Tony Berber Sardinha, da PUC/São Paulo, e o *Corpus* Multilíngue de Aprendizes, compilado pela Profa. Dra. Stella E. O. Tagnin, da USP.

Em relação às possibilidades de pesquisa em *corpora* de aprendizes, Leech (1998) destaca algumas questões que podem ser levantadas por meio de observação:

- a) Quais características linguísticas da língua-alvo são empregadas com mais (sobreuso) ou menos (subuso) frequência em comparação com falantes nativos?
- b) Qual é a extensão da influência da língua nativa (transferência) na produção dos aprendizes?
- c) Em que áreas eles tendem a usar estratégias de evitação deixando de explorar a fundo o potencial da língua-alvo?
- d) Em que áreas eles tendem a demonstrar desempenho nativo ou não nativo?
- e) Quais são as áreas nas quais os aprendizes de um dado país parecem necessitar de mais ajuda para desenvolver sua produção na língua-alvo? (LEECH, 1998, p.xiv).

Conforme se verifica na análise deste estudo, de fato, foi possível observarmos características de sobreuso e subuso de algumas estruturas sintáticas e colocações nas produções escritas dos alunos. Além disso, notamos a influência da língua materna nas escolhas léxico-gramaticais, por exemplo, no uso do artigo definido *the* em inglês. Contudo, ainda não podemos afirmar qual a extensão dessa influência, para tanto são necessárias mais pesquisas sobre esse assunto. Apesar de o aumento de pesquisas realizadas com base em *corpora* de aprendizes, Meunier (2011), ao advogar a favor do uso de *corpora* no ensino/aprendizagem de uma segunda língua, elenca algumas razões para a não adoção de metodologias dessa natureza nos ambientes de ensino, por exemplo, a falta de estudos empíricos que evidenciam o real impacto de métodos com o uso de *corpora*. Leech (2011) também aponta para o desconhecimento, por parte dos

professores de LE, dos traços evidenciados com base na observação da frequência de padrões linguísticos em *corpora* eletrônicos. Além disso, Meunier (2011) lembra que Granger (2009) apresenta uma avaliação crítica da real contribuição da análise de *corpora* de aprendizes à aquisição de uma segunda língua e para o ensino/aprendizagem de uma LE”. Granger acrescenta que, de fato, ainda há poucas evidências sobre sua aplicação na área.

Ao considerarmos os estudos citados anteriormente e a escassez de resultados de pesquisas com *corpora*, parte de nosso estudo propõe uma análise das colocações empregadas pelos aprendizes, cujas produções escritas compõem o *corpus* mencionado. Ressaltamos que um dos objetivos dessa investigação é contribuir para uma reflexão acerca de uma possibilidade, entre inúmeras outras existentes, de uso de *corpora* de aprendizes em sala de aula. Destacamos, ainda, que de modo algum entendemos que o uso de concordâncias geradas por meio de *corpora*, de referência ou de aprendizes, seja a única maneira de se ensinar vocabulário.

Conforme relatamos, trataremos neste estudo do emprego (ou não) de colocações por parte de aprendizes de língua inglesa, desse modo, cabe definirmos o termo “colocações”. De acordo com Tagnin (2005, p.37), o termo *collocation* foi introduzido pelo linguista Firth, para “designar casos de coocorrência léxico-sintática, ou seja, palavras que usualmente ‘andam juntas’”.

Por seu turno, Béjoint (1994), referiu-se às colocações como “o fenômeno das afinidades lexicais”. Segundo o autor, as ligações entre as palavras de uma colocação são geralmente lexicais e não semânticas. Com base nessa afirmação, podemos destacar que as ligações entre os elementos que formam as colocações podem não ser semânticas, mas sintagmáticas e convencionais. Tais combinações de palavras são regidas pela convenção daquela comunidade em que estão inseridas e, dessa forma, a combinatória entre seus elementos pode já estar previamente estabelecida e consagrada por essa comunidade.

Conforme poderá ser observado neste estudo, chamamos a atenção ao fato de que, por se tratarem de combinações léxico-gramaticais e convencionais na língua e, em razão de o número de colocações existentes na língua ser muito grande (TAGNIN, 2005), faz-se necessário ensiná-las de modo mais sistemático (BAHNS; EL DOW, 1993; BERBER SARDINHA, 2004; ORENHA, 2004),

ênfatizando aos alunos seu importante papel na redação de textos escritos e em sua produção oral. Por sistemático entende-se, neste estudo, o ensino guiado e formalizado, com a apresentação da definição do fenômeno das colocações e seus exemplos para, em seguida, propor um exercício de sua observação em linhas de concordância geradas por meio de *corpora* compostos por textos originalmente escritos em língua inglesa.

3. Metodologia

Neste artigo, apresentaremos alguns resultados de duas análises, conforme relataremos a seguir. Para tanto, compilamos um *corpus* de aprendizes de 38.594 palavras formado por textos em inglês, de diferentes temas, tais como aspectos culturais de diferentes países, hábitos pessoais e temas mais polêmicos (pena de morte, aborto, etc), produzidos por alunos dos terceiros anos diurno e noturno, da disciplina de *Produção e Interpretação Textual em Língua Inglesa*, turmas de 2005, 2006, 2007 e 2008. Este *corpus* se subdivide em dois *subcorpora*: 1) um *subcorpus* constituído de textos produzidos pelos alunos do turno diurno, somando um total de 16.353 itens; e 2) um *subcorpus* com textos de características idênticas produzidos pelos alunos do noturno, com um total de 22.241 itens. O *corpus* foi subdividido em razão de a mesma disciplina ser ministrada por uma das pesquisadoras no período diurno e pela outra pesquisadora no período noturno. A diferença no número de itens em cada *subcorpus* também se deve ao menor número de alunos nas classes do período diurno em todos os anos em que a pesquisa foi realizada.

Por meio do uso e da exploração do referido *corpus* de aprendizes, com base no arcabouço teórico-metodológico da Linguística de *Corpus* e com o auxílio do programa *WordSmith Tools* (SCOTT, 2004), visamos mostrar como textos escritos em língua inglesa poderiam ter seu conteúdo léxico e sintático mais naturais, ou seja, seguindo os padrões utilizados pelos nativos dessa língua estrangeira. Além disso, buscamos enfatizar o uso de colocações por parte dos alunos em relação a um *corpus* de referência que, para essa pesquisa, foi o *British National Corpus (BNC)*, acessado pelo *site*

<http://www.natcorp.ox.ac.uk>. Tal acesso permite a visualização de cinquenta linhas de concordância retiradas de textos que representam uma extensa variedade de gêneros na língua inglesa, seja falada ou escrita.

Para a pesquisa 1 deste trabalho, geramos uma lista de frequência das palavras contidas no subcorpus de textos produzidos pelos alunos do curso no período noturno. Em seguida, observamos as linhas de concordância a fim de verificar o uso das colocações utilizadas com o verbo *to have*, uma vez que elas haviam sido formalmente apresentadas aos alunos em uma das unidades do livro didático *Passages* (RICHARDS e SANDY, 1998) utilizado em sala de aula. Nessa unidade, os alunos deveriam procurar quais eram as combinações entre o verbo *to have* e os substantivos *influence*, *opinion*, *effect*, *faith*, *belief* e *idea* e, em seguida, deveriam produzir sentenças nas quais as colocações estudadas fossem empregadas. Como a proposta do livro didático se encerrava com o exercício de sentenças, procuramos observar se os alunos utilizavam as colocações estudadas em sua produção textual em língua inglesa e verificar se eles usavam outras colocações com o mesmo verbo. Para tanto, realizamos um levantamento das concordâncias com a palavra de busca *to have* no corpus composto pelas produções escritas em língua inglesa desde de 2005.

Em seguida, selecionamos apenas as colocações utilizadas com o verbo analisado *to have* que constavam nas linhas de concordância. Após a análise, mais uma atividade foi proposta a fim de apresentar outras colocações que não estavam presentes no livro didático. Desse modo, os alunos foram levados ao laboratório de informática durante duas aulas e verificaram as linhas de concordância do *BNC*, corpus composto por textos originalmente escritos em língua inglesa. A atividade proposta era a de selecionarem, com base nas linhas apresentadas, outras colocações com o verbo *to have*. Tal atividade foi realizada para que os alunos percebessem o alto índice de colocações utilizadas por nativos da língua inglesa. Ressaltamos, neste momento, que o propósito não foi o de apresentar textos autênticos completos, mas linhas de concordância com as inúmeras possibilidades de combinatórias tendo como base o verbo *to have*. Dessa maneira, os alunos teriam a chance de observar, mais nitidamente, outros exemplos de produções autênticas que não fossem somente aquelas apresentadas

no livro didático, as quais não sabemos se haviam sido retiradas de textos autênticos. Com as colocações do mesmo verbo em foco, acreditamos que a aquisição ou, ainda, a percepção desse fenômeno linguístico suscitaria o aprendizado por meio de comparações, por exemplo, entre as combinatórias em língua inglesa e as da própria língua materna.

Já a Pesquisa 2 deste estudo focou os substantivos *prejudice* e *regret*, cujo procedimento metodológico foi, de certa forma, semelhante ao realizado na Pesquisa 1. O motivo de havermos selecionado os referidos substantivos ocorreu em razão de os aprendizes terem escrito textos que tratavam de a) preconceitos e b) boas e más recordações da infância – em que os alunos abordaram do tema arrependimento. Dessa forma, por meio da análise das linhas de concordância geradas pela ferramenta *Concord*, do *WordSmith Tools*, fizemos um levantamento dos colocados empregados pelos alunos do período diurno com base nos referidos substantivos, uma vez que tiveram a tarefa de elaborar uma redação sobre *preconceito* e outra sobre *boas e más recordações de sua infância*, tema este em que se acaba abordando a questão dos *arrependimentos* (*regrets*), levando os alunos a falarem também sobre esse assunto.

Em virtude do baixo emprego de colocações e da produção de colocações não convencionais ou pouco frequentes na língua inglesa, ministramos algumas aulas no laboratório de informática, nas quais os alunos puderam acessar o *British National Corpus*, e gerar linhas de concordância dos substantivos pesquisados para que, dessa forma, pudessem observar colocações frequentemente empregadas por falantes nativos da língua inglesa. Essa atividade visou a conscientizar os alunos da existência de combinações lexicais recorrentes, chamar a atenção ao aspecto da coocorrência entre seus elementos e mostrar que a combinação de palavras não se dá de modo aleatório na língua, mas sim, de forma convencional.

4. Resultados

Os resultados das pesquisas foram subdivididos em duas partes: Pesquisa 1 e Pesquisa 2. Na primeira, apresentamos as

investigações com a colocação formada com o verbo *to have* e na segunda parte, apresentamos as colocações junto aos substantivos *prejudice* e *regret*.

4.1. Pesquisa 1

A lista de frequência do subcorpus de textos escritos pelos alunos do período noturno apontou uma frequência de 147 ocorrências do verbo *to have*. Em seguida, verificamos as linhas de concordância que apresentavam colocações com o verbo analisado, das quais listamos algumas a seguir.

<i>N</i>	<i>Concordance</i>
1	<i>We arrived home at 1pm and I went to have a lunch. In the afternoon I</i>
4	<i>te of Wine, is a place where the people have an opportunity to try the</i>
5	<i>good fit, because exercises have transient beneficial effects over the</i>
23	<i>pop band, Backstreet Boys. Now we don't have much time to listen to</i>
32	<i>ople who can't get along with computers have no chances to get a good</i>
36	<i>São José do Rio Preto because I have a great professional opportunity.</i>
48	<i>many subjects, to go out with them and have fun. I could know many</i>
55	<i>my mother. At 6pm I went to college and have normal classes. Tuesday</i>
59	<i>. An important thing is that exercises have a more lasting-long effect</i>
85	<i>ct on cognition. Researchers also don't have a clear idea of how much e</i>
993	<i>like to find a new place to go out and have fun. Our friendship shows</i>
102	<i>terous session of kickball, can also have much more widespread effects</i>
109	<i>an the house in the morning, and then I have a dentist appointment at I</i>
143	<i>with distinct ways. Think differently. Have different beliefs, values,</i>

Ao avaliarmos as colocações encontradas com o verbo *to have* no subcorpus de alunos do terceiro ano noturno, de 2005, 2006, 2007 e 2008, que faziam parte do conteúdo programático da disciplina, verificamos que, embora os alunos as utilizassem com frequência, a maioria das colocações eram parecidas com as utilizadas em língua portuguesa, como em *have time to*, *have classes*, *have an appointment*, *have chances* e *have opportunity* (“ter tempo para”, “ter aulas”, “ter um compromisso”, “ter chances” e “ter oportunidade”). Acreditamos que, embora uma comparação de frequência desta colocação entre o BNC e o corpus de aprendiz não tenha sido realizada, o uso das combinatórias utilizadas pode ter sido motivado por influência da língua materna dos aprendizes. A única colocação de alta frequência

com o verbo *to have* que não é comumente utilizada na língua portuguesa foi *have fun*. Foi possível observar também que, das colocações estudadas no livro didático, os alunos utilizaram *have effect* três vezes e *have (an) idea* uma vez.

Além disso, percebemos alguns desvios de sintaxe na utilização de certas colocações como em *have a lunch*, em vez de *have lunch* e *have beliefs*, utilizado no início de uma sentença e que, nesse caso, poderia ter sido empregado no infinitivo com o *to* ou no gerúndio.

Posteriormente a essas constatações e a aula no laboratório de línguas, os alunos puderam verificar outros tipos de colocações nas linhas de concordância do *BNC* que eram utilizadas com o verbo *to have*, conforme vemos nas linhas de concordância a seguir:

1. *I like to **have a laugh** and that, but there's some things you shouldn't laugh at.*
2. *You've got to **have a laugh** with the kids,'; says the proprietress,'; they're alright.'*
3. *Let's now look at how well you get along with other people **Have a look at the list***
4. *Right, let's **have a look at** you.'*
5. *Mr Rost says electricity users are fully justified in their 'angry determination'; to **have a showdown**.*
6. *And yet, if I did **have a showdown** now, what good would it do?*
7. *It is much better to **have a showdown** on the soccer pitch than on the battlefield,'; he said.*

Conforme os alunos puderam verificar, algumas das concordâncias apresentadas no *BNC* não eram as que eles conheciam, por exemplo, em *have a laugh*, *have a look at* ou *have a showdown*, entre outras. Desse modo, além das colocações estudadas anteriormente, eles puderam avaliar o quão comum o fenômeno da colocação com um mesmo verbo pode ser em textos originalmente escritos em língua inglesa.

Com base nessas análises, verificamos que, embora a colocação apresentada no livro já houvesse sido empregada em exercícios anteriores, seu uso nas produções escritas foi baixo, o que poderia, ou não, ter sido causado pelo tema das redações também. Em contrapartida, a observação das linhas de concordância no *corpus*

composto por textos originais retirados do livro didático, permitiu que os alunos se conscientizassem da variedade de “combinações” e equivalência possíveis, diferentemente daquelas que estavam acostumados a utilizar em seus textos.

4.2. Pesquisa 2

Para levantar as colocações surgidas pelo lema *prejudice*, presentes no subcorpus cujos textos haviam sido produzidos por alunos do turno diurno (16.353 palavras), utilizamos a ferramenta *Concord* e obtivemos as seguintes linhas de concordância:

<i>Concordance</i>	
1	<i>difficult to accept existing differences between the people. Increasingly, prejudice draws our attention by what has been happening. People just haven't</i>
2	<i>extremely difficult to accept existing differences between the people.” In a prejudiced society it's extremely difficult to accept existing differences</i>
3	<i>Topic Sentence concerning Shark Tale Topic Sentence: “In a prejudiced society it's extremely difficult to accept existing differences between</i>
4	<i>values discussed in this movie are the ones about law and the ones about prejudice and the relationship between black and white people. Nowadays, it's</i>
5	<i>and law, behaviors and mixture of people in the USA. The story is all about prejudice against black people and the way they have to deal with it.</i>
6	<i>“A time to kill” This movie talks about values, prejudice, human relations, justice and law, behaviors and mixture of people in</i>
7	<i>in Brazil, and also to study and improve my career. I know that it exists prejudices, racism and injustice there, like in others countries, so you have</i>
8	<i>prejudice. Sometimes, because of religion, races... I think that racial prejudice worse of all. Americans discriminate some groups like blacks,</i>
9	<i>I couldn't immigrate to the USA. I think that foreigners suffer much from prejudice. Sometimes, because of religion, races... I think that racial</i>
10	<i>races, sex, and religious. And especially that they won't know the word “prejudice” in the never case. Then, I agree with Martin, but I think</i>

Quadro 1: Concordâncias da palavra *prejudice** do Corpus de Aprendizes

De acordo com o esse quadro, o número de ocorrências da palavra de busca *prejudice* foi relativamente baixo, já que 1) consistiu de um tema da redação bastante discutido em sala de aula e, por esse

motivo, esperava-se ser empregado com mais profundidade; 2) era um dos temas do texto escrito que deveriam produzir; 3) havia um número de alunos que escreveu sobre o tema relativamente grande (mais de setenta alunos, somando os quatro anos de pesquisa). Quanto ao uso de colocações, as únicas empregadas foram *racial prejudice* e *suffer from prejudice*.

No que tange à palavra *regret*, foram geradas as seguintes linhas de concordância:

1	<i>and memories. I regret have not finished my piano 's course. I learnt this</i>
2	<i>One thing that I regret about my childhood is that I used to be very shy at school. I remember the teachers asked me to answer the homework</i>
3	<i>Like everybody, I had regrets and bad moments and memories.</i>

Quadro 2: Concordâncias da palavra regret do Corpus de Aprendizizes*

Conforme se observa, o termo teve igualmente uma frequência baixa no *corpus*, com base no qual levantamos apenas uma colocação (*have regrets*). A primeira ocorrência contém um erro gramatical quanto ao uso do verbo *regret*, pois o aluno parece não ter assimilado ou pode ter esquecido que, para expressar arrependimento no passado, ele precisa usar o verbo seguinte no gerúndio – este item gramatical foi abordado no segundo ano do curso. Já na segunda ocorrência, o aluno emprega o verbo *regret* seguido da preposição *about*, o que não é gramaticalmente aceito.

Com base nos dados acima apresentados, no número de participantes na pesquisa, e em resultados semelhantes em exercícios realizados em sala de aula que, por motivos de delimitação, não serão apresentados neste artigo, verificamos que, por não terem conhecimento do fator convencionalidade³ na língua, por fazerem apenas uma leitura composicional e não idiomática das estruturas linguísticas da língua, e por não saberem que a língua é também formada por combinações estáveis e recorrentes, acabam empregando colocados inadequados, não produzindo colocações fluentes ou apropriadas ao contexto enfocado. Por essa mesma razão, os

³ O termo convencionalidade está relacionado ao uso que fazemos, ao interagir, de uma série de convenções, expressões e blocos de palavras já preestabelecidos e consagrados em nossa comunidade. (ORENHA, 2004)

aprendizes podem, inclusive, empregar poucas colocações em seus textos, conforme mostra o uso de combinações com o substantivo *regret* (apenas três linhas de concordância). Também foi possível verificar que os aprendizes cujos textos foram analisados neste estudo, tendem a empregar colocações em que elementos – base e colocado – são praticamente isomórficos nas duas línguas tratadas, ou seja, possuem formas semelhantes às aquelas comumente empregadas na língua materna, de acordo com o que notamos em outras análises realizadas em sala de aula.

Em virtude desse resultado, os alunos foram direcionados à pesquisa via *BNC*. Nesse sentido, os aprendizes puderam constatar que junto do substantivo *prejudice* é possível formar uma série de colocações antes não imaginadas, tais como: *to waive prejudice against someone/something*, *encounter strong prejudice against someone/something*, *loud prejudice*, *quiet prejudice*; *flaunted prejudice*, *consciousness prejudice*, *irrational prejudice*, *to balance the prejudice*, conforme mostra o quadro 3.

A mesma constatação ocorreu com o substantivo *regret*: *express regret for*, *at the tinge of regret*, *wistful regret*, *show regret*, entre várias outras colocações, que igualmente poderiam enriquecer os textos escritos dos participantes desta pesquisa (quadro 4).

Após terem sido realizadas buscas pontuais e isoladas das colocações levantadas (exemplo: *encounter strong prejudice against*, *flaunted prejudice*, *at the tinge of regret* etc.) também no *BNC*, verificamos que são combinatórias típicas e naturais da língua inglesa. Ademais, constituem-se de combinações bastante diferentes das tentativas de produzi-las por parte dos aprendizes analisados neste estudo. Percebemos que muitas das colocações apresentadas pelo *BNC*, embora composto por textos de vários gêneros, não poderiam ter sido empregadas pelos referidos aprendizes, a não ser que eles tivessem aprendido ou observado seu uso anteriormente, já que sua combinabilidade poderia não se dar por critérios semânticos, mas, de modo geral, por critérios arbitrários, ditados pela convencionalidade na língua. Diante desse fato, os aprendizes não teriam consciência de sua estabilidade e estariam, assim, limitados linguisticamente. Daí nota-se a importância da atividade realizada, haja vista que os alunos tomaram consciência desse aspecto no processo de aprendizagem de uma LE e,

desse modo, passaram a empregar combinações de palavras mais elaboradas.

A68 871 *The manner of the rejection on the floor of the Chamber was unpleasant, with ranting and **prejudice flaunted and loud**,*
A74 867 *Lindeman **flaunted prejudice** by saying gay marriage "is nothing more than an attempt to legitimize an abnormal behavior"*
B14 411 *Any social debate about AIDS acquires a political dimension with associated **prejudice about race, religion, social class and nationality**;*
BNA 1006 *There is a great deal of **prejudice about age**, with the result that the older you are the more likely you are to remain unemployed*
CA0 2576 *Alejandro was so staggered by Fantasma's progress that he decided to **waive his prejudice against greys**;*
CGE 1023 *Even Phyllis's **racial prejudice** might have had a touch of the humorous about it,*
CGF 678 *Women who want to work in broadcasting, or who are required by their jobs to speak frequently in public, **encounter strong prejudice against their 'shrill'; or 'tinny' voices.***
CLI 528 *This shocked me 'cause I'd left Antigua with no idea what **colour prejudice** was, no idea.*
FAY 1333 *In the second case (necessary **false consciousness/irrational prejudice**) racism is seen as dysfunctional because it distorts the perception of reality*
FYT 660 *This meant a common search for a new quality of international interaction, less dependent on military force and free of **ideological prejudice**.*
G3E 859 *It wasn't fair that your grandmother left her fortune to him instead of sharing it with you just because he's a man and she had an **old-fashioned prejudice** that women shouldn't own money.*
HHV 8155 *The Government are confessing to a unique combination of **political prejudice** and economic weakness*
J6V 366 *It is clear from Thompson v Brown 1 WLR 744 that the s33 discretion is unfettered and can include all relevant matters in **balancing the prejudice**.*
J76 1405 *Bush J held that the child's welfare was a matter to be considered by the court and balanced against the **risk of prejudice** to the accused.*

Quadro 3: Concordâncias da palavra prejudice por meio do BNC

A5Y 637 *Most of the time I find it's **regret on the part of** the victim.*
ABK 656 *We **deeply regret** suggesting that farmers, who use around four-fifths of the available water,*
APD 385 *If there is **regret over** the failure of Resolution 242 as a blueprint for peace, [...] there are grounds for even **greater regret** that the 1949 Geneva Convention has not been vigorously upheld*
AR8 332 *With **deep regret** he decided that he would have to abort his part of the mission.*
BNI *without the sequel of second thoughts, without the **burden of regret.***
CKC 974 *I **deeply regret** both your loss of the members of your crew and*
FSP 131 *He could never **entirely regret** it, because it reminded him of working with Willie,*
G0N 1388 *It was one of **wistful regret** rather than fretful anxiety*
G17 11 *He controlled the urge to laugh **at the tinge of regret** which had coloured Artai's tone.*
H10 338 *I would like to believe that he is right though I notice **with regret** that the academic professions*
HA5 107 *A **pang of regret** lanced through her as she recognised how little she still knew of Suzie's innermost feelings.*
HD0 944 *As reported last month four boys have been suspended for unacceptable behaviour and have **expressed no regret** for their actions.*
HHX 14285 *Perhaps they will live to regret that.*
HPX 1609 *Later, as though **echoing their regret** that schools are not businesses, Handy and Aitken (1986:45) sum up the question about differences*
JXY 1659 *Suddenly she was surprised that she could talk about Ian White **without regret**; surprised to find that it didn't hurt any more.*
JYE 1013, *she felt a **pang of regret** that their relationship could not be friendly.*
K4B 284 *On increasing income tax, he said: 'One would see a change in that **with the utmost regret.***
K54 1012 *He **shows no regret** at not maintaining the movie career*

Quadro 4: Concordâncias da palavra regret por meio do BNC

5. Implicações para o ensino

Com este estudo, percebemos que houve uma contribuição para o ensino de colocações: a apresentação em diversos contextos de uso, assim como o emprego de várias colocações com as mesmas palavras de busca encontradas em um *corpus* composto por textos originalmente escritos em inglês, ou seja, por nativos da LE estudada.

Com base nos resultados gerados pelo *BNC*, em comparação àqueles levantados do *Corpus* de Aprendizes, pudemos verificar, assim como os estudos de Sinclair (1991), Tagnin (2005), Orenha (2004), Orenha, Paiva e Camargo (2007), entre outros, que as colocações têm um papel fundamental no comando de uma língua. Além disso, que o domínio de tais combinações recorrentes e convencionais, na língua, facilita a comunicação já que, uma vez adquiridas e armazenadas, o discurso, tanto escrito quanto oral, dá-se com maior rapidez e fluência e a mensagem é transmitida com mais fluidez e naturalidade (SINCLAIR, 1991). Ao analisarmos a produção escrita dos alunos envolvidos, tanto do curso diurno quanto do noturno (mais de 150 participantes), notamos esse aspecto descrito, no sentido de que seus textos poderiam soar com mais fluidez, se houvesse realmente um maior domínio das combinações de palavras.

Acreditamos que a conscientização do aprendiz em relação à língua em uso, mais especificamente quanto ao uso das colocações, pode promover a internalização da estrutura enfatizada e seu futuro emprego em textos escritos.

Outra implicação para o processo de ensino e de aprendizagem de língua, seja estrangeira ou materna, é a análise e a discussão dos textos dos alunos, individualmente ou em grupo, com o intuito de desenvolver a autoavaliação e o senso crítico em relação ao emprego de novas estruturas. Ressaltamos que as atividades realizadas em sala de aula aqui descritas são apenas uma, entre as inúmeras atividades ou, até mesmo, estratégias, que podem ser empregadas para o ensino de LE, mais especificamente de produção textual em uma LE. O uso de *corpora on-line* ou *corpora* de aprendizes apresenta-se como mais um recurso, por meio do qual o professor de LE pode explorar, a fim de auxiliar os alunos no processo de aquisição de vocabulário e de desenvolvimento da escrita, entre outros assuntos que também devem ser tratados ou desenvolvidos em aulas de produção textual em LE. Dentre os quais: a leitura de textos de diversos gêneros; a discussão de estratégias para a produção de textos, segundo o gênero focado; os comentários acerca dos problemas verificados nos textos produzidos etc.

6. Conclusões e encaminhamentos

Neste trabalho buscamos apresentar alguns pressupostos da Linguística de *Corpus* voltados para o ensino de LE, com base em parte de um estudo de textos produzidos por alunos da disciplina de Produção Textual em Língua Inglesa. Além disso, procuramos destacar a importância do uso e da exploração de *corpora* de aprendizes e de referência em aulas de produção textual em língua inglesa.

De acordo com os dados apresentados, observamos que ambas as pesquisas realizadas, entre outras atividades realizadas em sala de aula e não discutidas no âmbito deste artigo, demonstram que os aprendizes investigados neste estudo parecem combinar alguns itens lexicais semelhantes aos padrões colocacionais da língua materna e, dessa maneira, parecem não se arriscarem a escolhas mais complexas, preferindo ficar em uma zona de conforto.

Notamos, também, que a maioria deles, conforme indicam as produções escritas que formam nosso *corpus* de aprendizes, podem conhecer, teoricamente, que as colocações são combinações coesas, estáveis e recorrentes na língua. No entanto, ainda não se mostram capazes de empregá-las com naturalidade em seus textos.

Nesse sentido, cabe a nós, professores de língua inglesa como língua estrangeira, auxiliá-los nesse processo e ensinar-lhes tais aspectos, para que consigam aumentar o repertório de colocações e possam, dessa maneira, produzir textos mais naturais.

Referências

BAHNS, Jens; ELDAW, Moira. Should we teach EFL students collocations? In: *System*, v. 21, n. 1, p. 101-114, 1993.

BÉJOINT, Henri. Dictionaries and the dictionary. In: *Tradition and innovation in modern English dictionaries*, p. 6-41, 1994.

BERBER SARDINHA, Antonio. *Linguística de Corpus*. Barueri, SP: Editora Manole, 2004.

British National Corpus. Disponível em <http://www.hcu.ox.ac.uk/BNC>. Acesso em 2005.

HAAN, Pieter de. The Optimum Corpus Sample Size? In: LEITNER, G. (Org.) *New Directions in English Language Corpora*. Berlin: Mouton de Gruyter, p. 3-19, 1992,

GRANGER Sylviane. The International Corpus of Learner English. In: AARTS, J.; DE HAAN, P.; OOSTDIJK, N. (Orgs.) *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam: Rodopi, 57-69, 1993.

_____. *Learner English on Computer*. London and New York: Longman, 1998.

_____. The contribution of learner corpora do second language acquisition and foreign language teaching: a critical evaluation. In: ALIJMER, K. (Org.) *Corpora and language teaching*. Studies in Corpus Linguistics 33. Amsterdam: John Benjamins, 2009.

LEECH, Geoffrey. 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.

_____. Frequency, corpora and language learning. In: MEUNIER, Fanny; DE COCK, Silvie.; GILQUIN, Gaëtanelle; PAQUOT, Magali.(Ed). *A Taste for Corpora*. In honour of Sylviane Granger. Amsterdam/Philadelphia: Benjamins, 2011.

MEUNIER, Fanny. Corpus linguistics and second/foreign language learning: exploring multiple paths. In: *Revista Brasileira de Linguística Aplicada*. v. 11, n. 2 2011.p. 459-477.

NICHOLLS, Diane. The Cambridge Learner Corpus – error coding and analysis for lexicography and ELT. In: ARCHER, Dawn, et al. (Orgs.). *Proceedings of the Corpus Linguistics 2003 Conference (CL*

2003). Technical Papers 16. Lancaster University: University Centre for Computer Corpus Research on Language, 2003.

ORENHA, Adriane. *A compilação de um glossário bilíngüe de colocações, na área de jornalismo de Negócios, baseado em corpus comparável*. São Paulo, 2004. 241 f. Dissertação (Mestrado em Estudos Lingüísticos e Literários) - FFLCH/USP.

ORENHA, Adriane; PAIVA, Paula. T. P.; CAMARGO, Diva. C. O uso de corpora de aprendizes no ensino de produção escrita em língua inglesa. *MOARA*, v. 26, p. 281-293, 2007.

RICHARDS, Jack C.; SANDY, Chuck (Orgs.). *Passages - an upper-level multi-skills course*. Cambridge: Cambridge University Press, 1998.

SCOTT, Mike. *WordSmith Tools: Version 4*. Oxford: Oxford University Press, 2004.

SINCLAIR, John. MCH; RENOUF, Antoinette. A lexical syllabus for language learning. In: CARTER, Ronald; McCARTHY, Michael. (Orgs.). *Vocabulary and language teaching*. Londres: Longman, 1988, p.140-60.

SINCLAIR, John. *Corpus, concordance and collocations*. Oxford: Oxford University Press, 1991.

TAGNIN, Stella. E. O. *O jeito que a gente diz: expressões convencionais e idiomáticas*. São Paulo: Editora Disal, 2005.

THORNDIKE, Edward L. *Teacher's wordbook*. Nova York, Columbia Teachers College, 1921.

TRIBBLE, Chris. *Concordancing and an EAP writing programme*. *CAELL Journal*, v.1, n.2, p.10-15, 1990.

_____. Some uses of electronic text in English for academic purposes. In: MILTON, John C.; TONG, Keith S. T. (Orgs.). *Text Analysis in*

Computer Assisted Language Learning. Hong Kong: Hong Kong University of Science and Technology, p. 4-14, 1991.

VIANA, Vander; TAGNIN, Stella. E. O. (Orgs.). *Corpora no ensino de línguas estrangeiras*. 1. ed. São Paulo: HUB Editorial, 2011 v. 1. 375 p.

WEST, Mike. *A general list of English words*. Londres: Longman, 1953.

Recebido em: 03/02/2012

Aceito em: 29/11/2012

Title: The use of corpora in English writing classes

