

APRENDIZADO MOVIDO POR DADOS PARA INICIANTES: O USO DE CORPUS DE APRENDIZES PARA O ESTUDO DA ORDEM DO ADJETIVO EM LÍNGUA INGLESA

Lílian Figueiró TEIXEIRA

(Universidade do Vale do Rio dos Sinos – UNISINOS)

lilianjoy@gmail.com

Rove Luiza de Oliveira CHISHMAN

(Universidade do Vale do Rio dos Sinos – UNISINOS)

rove@unisinossinos.br

RESUMO: No ensino de língua estrangeira, há uma crescente tendência de valorização do estudo da linguagem autêntica através de metodologias que conferem ao aprendiz um papel central. Este trabalho objetiva utilizar um desses métodos, o Aprendizado Movido por Dados, para a análise da ordem dos adjetivos em inglês e para a elaboração e a aplicação de atividades pedagógicas voltadas para aprendizes iniciantes. Optamos por uma proposta envolvendo um corpus de aprendizes, constituído por produções dos alunos, cujas ocorrências foram comparadas com as produzidas por falantes nativos tanto na fase de análise quanto na elaboração e na aplicação das atividades pedagógicas.

PALAVRAS-CHAVE: Lingüística de Corpus; ensino de língua estrangeira; corpus de aprendizes; ordem do adjetivo em língua inglesa.

ABSTRACT: In the foreign language teaching field, there is a tendency in studying authentic language through methodologies in which the learner plays a central role. The purpose of this paper is to use the Data Driven Learning method to analyse the adjective order in English and to develop and carry out pedagogical tasks for beginners. We decided to work with learner corpus, which is formed by students' productions, whose occurrences were compared with those produced by native speakers during the analysis, development and application of the pedagogical tasks.

KEYWORDS: Corpus Linguistics; foreign language teaching; learner corpus, adjective order in English.

0. Introdução

A utilização de corpus para o estudo da língua e para a elaboração de materiais como dicionários e gramáticas já não é mais novidade, mas a sua aplicação direta em sala de aula ainda está dando os primeiros passos.

Imaginemos uma aula de língua estrangeira em que os alunos possam buscar as respostas para as suas dúvidas em exemplos de usos reais da língua. Esse tipo de postura do aprendiz substitui o que se faz normalmente, como aceitar incontestavelmente a resposta do professor ou buscar uma resposta extremamente comportada em uma gramática ou dicionário. Sob essa nova perspectiva, o aluno chega às suas próprias conclusões a partir da observação de dados autênticos.

Neste trabalho, apresentaremos uma sugestão sobre como realizar o estudo de corpus para a análise da ordem do adjetivo em língua inglesa e para a elaboração de tarefas em uma aula de língua estrangeira.

Partindo das produções textuais de aprendizes de inglês de uma escola pública do Rio Grande do Sul, percebe-se certa dificuldade quanto à ordem dos adjetivos, o que justifica a escolha desse fenômeno para o presente estudo. Essas mesmas produções foram organizadas em um corpus e trazidas para a sala de aula com o objetivo de servirem de material de estudo para os alunos que as produziram.

Abaixo, apresentaremos os princípios teóricos deste trabalho, assim como as suas principais etapas e conclusões.

1. Aprendizado movido por dados

A proposta de ensino de Tim Johns conhecida como *Data Driven Learning (DDL)* pode ser traduzida por Aprendizado Movido por Dados, e, segundo Berber Sardinha (2004), é a abordagem mais consistente que utiliza corpus no ensino de língua. O estudo da linguagem autêntica através de concordâncias, uma lista com todas as ocorrências de uma palavra de busca, é um dos traços mais específicos do *DDL*.

Além de representar uma inovação no ensino de língua estrangeira, ou seja, mais uma forma de se utilizar os recursos computacionais na sala de aula, há várias vantagens ao estudar a língua através de concordâncias. Entre os benefícios do seu uso, Stevens (1995) afirma que a concordância é uma forma econômica, em termos de tempo, de manipular textos, já que, após ter um programa e uma base de dados,

pode-se simplesmente incluir novos textos, implementando assim essa base rapidamente.

Há outros métodos, como o *Lexical Approach* (Abordagem Lexical) e o *Lexical Syllabus* (Currículo Lexical), que utilizam os conceitos e a metodologia da lingüística de corpus, no entanto para Berber Sardinha (op.cit) o *DDL* traz mais vantagens para o ensino. Na Abordagem Lexical, o foco é no léxico, ensinado através de textos escritos ou falados. O léxico é apresentado por meio de porções de palavras, de colocação e nunca de palavras isoladas, e as atividades têm como foco o ensino da forma e não de tarefas.

O principal problema com essa abordagem, conforme Berber Sardinha (op cit), é a falta de clareza por parte dos materiais de referência, quando e como os corpora foram utilizados para a preparação dos recursos e quais os tipos de ferramentas e critérios utilizados para a organização do léxico.

O centro no léxico também é uma característica do Currículo Lexical, cuja proposta de ensino partiu da descrição de corpus. Com base em uma lista com as palavras mais freqüentes da língua, os aprendizes são expostos inicialmente a certo número de itens lexicais mais freqüentes, número esse que é ampliado ao longo o curso.

O léxico é apresentado a partir de textos retirados do próprio corpus que serviu como base de dados para a seleção das formas a serem ensinadas. O aspecto positivo desse método, para Berber Sardinha (op.cit.), é a centralidade no aluno.

Embora os seus autores afirmem que ele seja baseado em tarefas, isso não ocorre na prática, e o foco, de fato, é na forma (Berber Sardinha, op.cit). Outra crítica que o Currículo Lexical recebe está relacionada aos exemplos, pois não parecem autênticos, e dão a impressão de terem sido fabricados ou simplificados.

A maneira como o corpus é utilizado parece ser o maior problema do Currículo Lexical e da Abordagem Lexical. No entanto, uma aula de língua estrangeira guiada pela proposta do *DDL* tem o corpus como um instrumento muito importante, já que as tarefas são elaboradas com base nos dados do corpus.

Segundo Granger e Tribble (1998), o *DDL* tem sido utilizado com corpora de falantes nativos do inglês, com o objetivo de refletir sobre a forma como as pessoas realmente falam ou escrevem. É no corpus que os alunos esclarecem as suas dúvidas quanto às especificidades da língua, e não mais simplesmente fazendo questionamentos ao professor.

Entre as principais justificativas para que os alunos criem regras e generalizações sobre a língua a partir do estudo de corpus, Aston

(2000) enfatiza o fato de que os materiais didáticos tradicionais ignoram a frequência de uso presente em dados de falantes nativos. Isso faz com que o aluno perca tempo estudando estruturas na aula de língua estrangeira que raramente ocorrem em uma situação concreta de comunicação.

O DDL é um método indutivo, partindo dos dados observados para a elaboração de regras ou padrões da língua. Dessa forma, pode-se dizer que o aluno exerce um papel central e mais independente, já que ele é responsável pela pesquisa no corpus e cria suas próprias conclusões acerca da língua.

De certo modo, o aprendiz não faz esse trabalho totalmente sozinho nem aprende a ser um pesquisador de uma hora para outra. Cabe ao professor orientar e diminuir o controle das atividades à medida que o aprendiz torna-se mais autônomo em relação ao seu aprendizado.

Bernardini (2004) refere-se ao *DDL* também como "o aprendiz como pesquisador". A autora destaca alguns efeitos da utilização de rotinas de aprendizagem indutivas, tais como: o professor torna-se o coordenador da pesquisa ou o facilitador; o aprendiz aprende a partir de exercícios que envolvem observação e interpretação de padrões de uso.

Para Hadley (2005), que introduziu o método em aulas de língua inglesa para alunos japoneses, ao se trabalhar com o *DDL*, deve-se oferecer uma série de atividades de conscientização desde o início. Essas atividades devem envolver as técnicas de identificação, classificação e generalização.

Encontramos essas técnicas em Berber Sardinha (2004), que define a identificação como o primeiro contato dos alunos com a concordância. É nesse momento que eles identificam padrões ou aspectos recorrentes. Logo após, classificam esses aspectos de acordo com a sua interpretação. Para concluir o trabalho, criam generalizações sobre a ocorrência da classificação.

A seqüência das atividades nem sempre ocorre dessa forma; o aluno pode iniciar o trabalho a partir de generalizações e tentar confirmá-las ou refutá-las com a identificação e a classificação, podendo gerar uma nova generalização.

Inspiramo-nos no DDL para a elaboração de atividades didáticas, porém, partimos de um corpus formado por textos escritos por aprendizes, denominado na literatura da área como corpus de aprendizes.

2. Corpus de aprendizes

Por corpus de aprendizes, compreende-se uma “coleção de textos sistematicamente computadorizados produzidos por aprendizes” (Nesselhauf, 2004:127). Esses dados, de acordo com Granger, Hung e Petch-Tycon (2002), podem ser analisados com ferramentas lingüísticas que realizam tarefas simples como pesquisas e contagens, até com as mais complexas, que fornecem análises sofisticadas dos dados.

A partir dessa idéia, é possível imaginar que um professor em poder de produções textuais de seus alunos poderia construir um corpus simplesmente digitando todas e guardando-as em um arquivo. Porém, a construção de um corpus de aprendizes deve seguir alguns critérios, que estão intimamente ligados com o objetivo de análise desse professor.

Geralmente, ao se compilar um corpus, o professor busca identificar as principais dúvidas que os seus aprendizes demonstram ter, e, a partir da identificação desses problemas nos textos, propõe algumas atividades de retomada do conteúdo.

Essa é uma possibilidade, mas existem outras, como, por exemplo, o projeto *International Corpus of Learner English (ICLE)*, que busca reunir o mesmo tipo de produções textuais de aprendizes de níveis intermediário e avançado de inglês como língua estrangeira. Os aprendizes são de variadas nacionalidades, e a parte do projeto com as produções de brasileiros aprendizes de inglês está sendo organizada pelo professor Tony Berber Sardinha (PUC-SP¹).

Nesselhauf (2004) afirma que os textos de um corpus de aprendizes devem ser o mais natural possível, mas que, no ambiente de língua estrangeira, não há como encontrar textos que ocorram naturalmente. Dessa forma, busca-se coletar textos que foram escritos sob circunstâncias de muito pouco controle.

Esses materiais podem ser resultados de atividades pedagógicas ou elaborados com o objetivo de pertencerem a um corpus. Ao seguirmos as orientações de Nesselhauf (op cit), evitamos redações guiadas por figuras ou por traduções, exatamente para obter material produzido sob circunstâncias de pouco controle.

Quanto ao tamanho, o corpus de aprendiz não precisa ser muito grande. O mais importante, segundo Ferreira (2003), é que o corpus seja autêntico e tenha uma formatação que dê sustentação para a pesquisa.

¹ Maiores informações em: <<http://www2.lael.pucsp.br/corpora/bricle/index.htm>>

Granger (1998) apresenta dois tipos de critérios a serem considerados durante a compilação de um corpus: os relacionados à língua e os que dizem respeito às características do aprendiz. Como critérios da língua, temos o meio, que pode ser escrito ou falado, o gênero discursivo, o tópico da atividade e o cenário, ou seja, as condições para que os alunos produzissem o material.

Entre os critérios referentes ao aprendiz, é importante destacar a idade, o gênero (masculino ou feminino), a sua língua materna, o nível de conhecimento da língua por parte dos aprendizes e o contexto de aprendizagem, como língua estrangeira ou materna.

Em nosso experimento, os aprendizes são adolescentes que cursam o ensino médio de uma escola pública no estado do Rio Grande do Sul; portanto, podem ser considerados iniciantes, já que muitos não demonstram ter conhecimentos anteriores de língua inglesa, mesmo tendo tido aulas de inglês desde a quinta série do ensino fundamental.

Optou-se por produções escritas, visto que as habilidades de escrita e leitura têm sido prioridade no ensino médio de língua estrangeira em detrimento da habilidade oral. Todos os textos fizeram parte de uma tarefa maior, que era trocar correspondências com alunos de outras turmas com o objetivo de fazer novos amigos.

Após o recebimento de uma resposta, eles deveriam descrever esse novo amigo expressando a sua opinião sobre ele ou ela. Essa tarefa resultou em dois tipos de textos: cartas e textos descritivos. Optamos por cartas e não e-mail, pois não havia na escola na época um laboratório de informática em que os alunos pudessem ter acesso à Internet. Abaixo, apresentamos dois quadros com um resumo dos critérios do corpus de aprendizes compilado para esse estudo.

Critérios da língua	
Meio:	escrito
Gênero:	cartas e descrições
Tópico:	fazer novos amigos
Cenário da atividade:	os alunos iniciaram a produção na sala de aula e alguns terminaram em casa; como recurso podiam consultar o caderno e o livro; as redações foram individuais e fizeram parte da avaliação.

Quadro 1: critérios para seleção dos textos incluídos no corpus de aprendizes

Critérios dos aprendizes	
Idade:	adolescentes (15 – 18)
Gênero:	masculino e feminino
Língua materna:	português
Nível:	ensino médio / iniciantes
Contexto de aprendizagem:	inglês como língua estrangeira.

Quadro 2: critérios para seleção dos textos incluídos no corpus de aprendizes

Todos os alunos das turmas selecionadas participaram da experiência. Eles produziram os textos e participaram da aplicação das atividades. No entanto, para a inclusão dos textos em um corpus para pesquisa, era necessário que os pais dos alunos menores de idade assinassem um termo de consentimento livre e esclarecido.

Essa exigência fez com que o número de redações incluídas no corpus fosse reduzido, posto que muitos se esqueciam de trazer o documento assinado ou perdiam. Dessa forma, 106 alunos participaram das atividades propostas, totalizando a inclusão de 101 textos no corpus de aprendizes.

3. Objetivos e metodologia do trabalho

Ao ler os textos produzidos pelos aprendizes que participaram deste estudo, deparamo-nos com ocorrências como *eyes brown*, *his actor favorite* e *a girl very special*. Em um primeiro momento, pensa-se que a dificuldade com a ordem do adjetivo enfrentada por alunos iniciantes deve-se à interferência da língua portuguesa. Dessa maneira, analisamos a posição do adjetivo nas línguas portuguesa e inglesa por meio de consultas a algumas gramáticas, visando a identificar as especificidades das línguas no que tange ao fenômeno em questão.

Ao considerar como fonte de consulta o corpus de aprendizes, foi possível concretizar um segundo objetivo: analisar o uso dos adjetivos em redações de aprendizes iniciantes de língua inglesa utilizando os recursos da Lingüística de Corpus.

Além de analisar como os aprendizes utilizam os adjetivos em suas produções escritas, objetivamos elaborar e aplicar os recursos da Lingüística de Corpus em uma situação concreta de ensino com alunos iniciantes.

Parte da motivação em aplicar atividades que envolvem corpus com alunos iniciantes deve-se a Hadley (2005), pois a maioria dos trabalhos tanto com *DDL* quanto com corpus de aprendizes é voltada

para alunos de nível intermediário ou avançado. O mesmo autor demonstra que é possível trabalhar com alunos iniciantes, desde que esses sejam preparados para o trabalho com as concordâncias e bem orientados durante o seu primeiro contato com essa metodologia.

Nas próximas seções, as etapas da realização deste trabalho serão brevemente apresentadas.

4. Gramáticas analisadas

Ao partir da hipótese de que a dificuldade com a ordem do adjetivo em inglês dos aprendizes brasileiros se dá por interferência da sua língua materna, decidimos analisar algumas gramáticas, a fim de comparar de que forma as descrições das duas línguas assemelham-se ou diferem. As consultas foram feitas apenas pelos pesquisadores, e os alunos não tiveram acesso a regras gramaticais antes de realizar a atividade.

Em língua portuguesa, analisamos duas gramáticas baseadas em corpora, a "Gramática de Usos do Português" de Moura Neves (2000), e "Uma Gramática de Valências para o Português", de Borba (1996). Em inglês, usamos a gramática "*Collins Cobuild English Grammar*" (Sinclair, 1990).

Para ampliarmos nosso estudo, escolhemos outras gramáticas de língua inglesa que são referências em cursos de graduação, sendo elas a "Longman English Grammar" (Alexander, 2002), "The Grammar Book" (Celce-Murcia, Larsen-Freeman, 1999) e "A Communicative Grammar of English" (Leech, Svartvik, 1977). A última foi escolhida por ter sido a primeira gramática considerada comunicativa.

Nas gramáticas de Borba (1996), de Neves (2000) e de Sinclair (1990), ou seja, nas gramáticas baseadas em estudo de corpus, os adjetivos são divididos em classificadores e qualificadores. Optamos por manter esta divisão, pois Sinclair (op cit) consegue propor uma ordem dos adjetivos que é comprovada através do estudo do seu uso em corpus.

Outra razão para essa escolha é que os grupos dos classificadores e dos qualificadores são mais abrangentes, incluindo vários tipos de adjetivos. Nas outras gramáticas mais tradicionais, os autores buscam especificar esses tipos de adjetivos, posicionamento que leva a uma discordância entre os autores.

Enquanto Celce-Murcia e Larsen-Freeman (1999) sugerem que a ordem seja determinante, **opinião, tamanho, formato, condição, idade, cor, origem e substantivo**, para Alexander (2002), a ordem

mais adequada é **modificadores (tais como *big* e *little*), qualidade/opinião, tamanho/idade/formato/temperatura/sabor, etc., pátrios, formados pelo particípio passado e substantivo.** (grifo nosso)

Ao mantermos a divisão dos adjetivos em classificadores e qualificadores, também facilitamos a nossa tarefa de comparação com as gramáticas do português, pois as gramáticas escolhidas, por serem baseadas em corpus, também classificam os adjetivos dessa forma.

A seguir, apresentamos um quadro com um resumo geral sobre o uso dos adjetivos nas línguas inglesa e portuguesa a partir do estudo das gramáticas baseadas em corpora.

	Português	Inglês
1 – Os adjetivos podem ocorrer na posição atributiva e predicativa.	SIM	SIM
2 – A maioria dos adjetivos ocorre na posição atributiva.	SIM	SIM
3 - Alguns adjetivos mudam de significado quando em posição diferente.	SIM	SIM
4 – A posição atributiva mais comum é a pré-nominal.	NÃO	SIM
5 – A posição atributiva mais comum é a pós-nominal.	SIM	NÃO
6 – A posição anteposta é mais subjetiva.	SIM	NÃO
7 – O substantivo pode ter a função de adjetivo.	SIM	SIM
8 – Geralmente não usamos mais do que dois adjetivos com o mesmo substantivo.	SIM	SIM
9 – Ordem: qualificador + cor + classificador + substantivo	NÃO	SIM
10 – Ordem: qualificador + substantivo + classificador	SIM	NÃO

Quadro 3: Comparação dos adjetivos em português e inglês.

Ao comparar as duas línguas, vemos que há mais semelhanças do que diferenças entre o português e o inglês. A principal diferença é que, normalmente, em português, usa-se o adjetivo após o substantivo, e em inglês, o adjetivo é usado antes do substantivo. Porém, essa questão não é tão simples quanto parece ser.

Como as duas opções em português (anteposta e posposta) são possíveis, cada posição pode atribuir um sentido diferente ao adjetivo, como em *velho amigo* e *amigo velho*. Não sendo esse recurso possível,

em inglês, o aprendiz pode ter dúvidas, sendo obrigado a escolher outra palavra mais adequada, ou a inferir o sentido da palavra de acordo com o contexto.

Em ambos os idiomas, o substantivo pode ser usado como modificador de um outro substantivo. Entretanto, esse recurso parece ser mais freqüente na língua inglesa. Em português, algumas vezes, a ocorrência de substantivos modificadores não é percebida, já que, no uso, há certa dificuldade em diferenciarmos se determinada palavra é usada como substantivo ou adjetivo.

Os dois últimos itens do quadro 3 sugerem uma ordem específica quando temos dois adjetivos modificando o mesmo substantivo. Em inglês, eles ocorrem antes do substantivo, enquanto em português, quando temos três formas de adjetivo em uma frase, a palavra do meio será o substantivo.

Mais adiante, a partir do estudo de corpus, veremos que nem sempre essa regra é seguida pelos usuários falantes nativos, cujas ocorrências foram registradas nos corpora estudados neste trabalho.

5. Coleta do corpus

Para a construção do corpus, foram coletadas redações de alunos de ensino médio de uma escola pública do Rio Grande do Sul. Os participantes desse projeto eram alunos de três turmas de primeiro ano e de uma turma de segundo ano do ensino médio do turno da tarde.

Todos os textos foram digitados cuidadosamente um a um e armazenados em um arquivo em formato *txt*. Para evitar constrangimentos, todos os nomes dos alunos foram substituídos por XXX. É muito importante que, ao serem digitados, os textos mantenham a sua forma original, caso se queira aproveitar o mesmo material para outros tipos de estudo futuramente.

Também é aconselhável manter os textos na íntegra, pois, se ocorrer alguma dúvida a partir das linhas de concordância, temos como recorrer ao seu texto de origem. Para uma futura conferência, todos os originais foram guardados. Utilizou-se o formato *txt*, pois, para a obtenção de uma lista de palavras e de concordância, utilizamos o concordanciador do *site* do Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem (LAEL²) da PUC-SP, que exige esse formato para que o corpus seja analisado.

² Disponível em: <<http://lael.pucsp.br>>

No quadro abaixo, temos os dados gerais do corpus de aprendizes coletado obtidos através do software do LAEL.

Número de cartas:	49
Número de descrições:	52
Total de textos:	101
Número de formas (<i>types</i>):	732
Número de palavras (<i>tokens</i>):	8629

Quadro 4: Dados do corpus de aprendizes

6. Análise dos dados

Como base de referência - a lista de palavras gerada pelo programa do LAEL-, observamos que os adjetivos são freqüentes no corpus e selecionamos quatro para nossa proposta. Analisamos os três adjetivos mais freqüentes: *old* (91 ocorrências), *brown* (72 ocorrências) e *favorite* (62 ocorrências). Também analisamos a palavra *grade* (16 ocorrências), pois é o único substantivo modificador encontrado nas redações. O quadro abaixo traz os dados referentes à utilização dessas palavras pelos aprendizes a partir das suas freqüências de uso identificadas na lista de palavras:

Adjetivos	605	7% do total de <i>tokens</i>
<i>old</i>	15%	Porcentagem sobre o total de adjetivos
<i>brown</i>	12%	
<i>favorite</i>	10%	
<i>grade</i>	2,6%	

Quadro 5: Freqüência de uso dos adjetivos do corpus de aprendizes a partir da lista das 100 palavras mais freqüentes do corpus.

Após a escolha desses quatro itens, geramos as linhas de concordância de cada um, também com o auxílio do programa do LAEL. As etapas dessa análise assemelham-se aos procedimentos que os alunos devem adotar ao pesquisarem sobre a língua em sala de aula, conforme Nesselhauf (2004).

Nesse momento, o trabalho foi dividido basicamente em duas etapas: obtenção das linhas de concordância de um corpus de aprendizes e de falantes nativos e identificação das diferenças entre as concordâncias.

Como não conseguimos ter acesso a um corpus de falantes nativos, utilizamos as concordâncias disponibilizadas nos sites do *Collins Cobuild*³ (*Bank of English*) e do *British National Corpus*⁴ (*BNC*).

O *Collins Cobuild* é constituído por 420 milhões de palavras de inglês escrito e falado. Em sua maior parte, contém textos do inglês britânico, mas há também outras variações da língua. Já o *BNC* possui aproximadamente 100 milhões de palavras do inglês britânico escrito e falado. Salientamos que, na versão disponível on-line, não temos acesso ao corpus integral, e o *Cobuild* fornece apenas 40 linhas de concordância para cada busca enquanto que o *BNC*, 50 linhas.

No quadro 6, há um esquema com as principais conclusões de nossa análise ao comparar as ocorrências do corpus de aprendizes com as do site do *Cobuild* e do *BNC*.

Adjetivo	Corpus de Aprendizes	Concordâncias de Falantes Nativos ⁵		
			<i>Collins</i>	<i>BNC</i>
Old	<ul style="list-style-type: none"> • Todos os usos são relacionados à idade (pergunta/resposta): <i>how old are you? I am sixteen years old.</i> • Usado após substantivo: <i>years old</i> – 79%⁶. • O verbo <i>to have</i> é usado para informar a idade: <i>I have fourteen years old</i> – 40%. 	<ul style="list-style-type: none"> • <i>Old</i> com o sentido de velho – antes do substantivo: <i>many more old men died.</i> 	75%	72%
		<ul style="list-style-type: none"> • Adjetivo composto indicando a idade – antes do substantivo: <i>41-year-old teacher.</i> 	22%	2%
		<ul style="list-style-type: none"> • Substantivo modificador, após o adjetivo <i>old</i>: <i>an old submarine movie.</i> 	15%	8%
		<ul style="list-style-type: none"> • Usado após substantivo: <i>years old.</i> 	X	8%
Grade	<ul style="list-style-type: none"> • Usado como modificador de <i>student</i>: <i>I am a 1st grade student</i> – 73%. • Significa <i>série</i> em todas as ocorrências: <i>She is a 1st grade student.</i> • Usado como substantivo: <i>What grade are you in?</i> – 27%. 	<ul style="list-style-type: none"> • Outros significados, tais como <i>classe, grau e nota</i>: <i>(minimum grade C) in Mathematics is required.</i> • Quando informa a série de alguém, usa-se adjetivo composto: <i>second-grade student.</i> 	55%	98%
			15%	X

³ Disponível em: <http://www.collins.co.uk/corpus/CorpusSearch.aspx>

⁴ Disponível em: <http://www.natcorp.ox.ac.uk/>

⁵ Como não tivemos acesso a todos os dados dos corpora *BNC* e *Collins*, a percentagem foi calculada a partir do resultado de amostra de cada, sendo 50 linhas de concordância para o *BNC* e 40 para o *Collins*.

⁶ A percentagem das ocorrências desse quadro é calculada a partir do número total de ocorrências presentes no corpus.

Brown	<ul style="list-style-type: none"> • Descreve sempre a cor dos olhos ou do cabelo: <i>I have long curly brown hair and brown eyes.</i> • Mesma ordem do português: <i>eyes brown, hair brown</i> – 7%. • Falta de consenso quanto à ordem mais adequada dos adjetivos quando usa-se mais de dois adjetivos, pois é possível encontrar ocorrências como: <i>brown long hair, long straight brown hair, long curly hair brown.</i> 	<ul style="list-style-type: none"> • Substantivo – nome próprio: <i>Former California Governor Jerry Brown.</i> • Raramente usa-se mais de dois adjetivos: <i>long brown wavy hair.</i> • Quando há dois adjetivos, o de cor ocorre mais próximo ao substantivo: <i>a large brown teddy in the corner.</i> • Ordem relacionada com o contexto, com a intenção do falante, quando usa-se mais de dois adjetivos: <i>long brown wavy hair, strong curly black hair.</i> 	57%	42%
			10%	X
			17%	10%
Favorite	<ul style="list-style-type: none"> • A ordem é determinante, <i>favorite</i> e o substantivo: <i>her favorite actor</i> – 88%. • O adjetivo é utilizado no plural: <i>my/her favorites bands</i> – 8%. • Mesma ordem do português: <i>My band/actress/food favorite, the favorite movie his</i> – 9%. 	<ul style="list-style-type: none"> • A ordem é determinante, <i>favorite</i>, podendo ter um substantivo modificador e o substantivo, como em: <i>our favorite wildlife images, his own favorite subway story.</i> 	60%	45%

Quadro 6: Comparação entre as ocorrências do corpus de aprendizes e as concordâncias de falantes nativos.

A seguir, explicitaremos algumas conclusões às quais chegamos a partir dos dados observados no quadro 6. Como os produtores dos textos são alunos iniciantes, não é de se estranhar que eles tenham um menor conhecimento da língua, tanto dos diversos sentidos de uma mesma palavra quanto dos diversos usos que se pode fazer.

No corpus de aprendizes, não foi possível encontrar adjetivos compostos, algo freqüente nas linhas de concordância de falantes nativos quando se referem à idade de alguém. Vale ressaltar que, nas concordâncias on-line, o substantivo modificador é bastante usado, e os alunos utilizaram apenas a palavra *grade* nessa função em seus textos.

Há poucas ocorrências, conforme os dados do quadro 6, mas alunos preservam em suas produções a mesma ordem dos adjetivos presente na língua portuguesa. Exemplos disso são os casos de *eyes brown, My band favorite* e *The favorite movie his*.

No caso específico do adjetivo *favorite*, destacamos que, na utilização do determinante *my*, a ordem determinante, adjetivo e substantivo aparece com maior frequência, mas, na utilização de outros determinantes como *his* ou *her*, há interferência da língua portuguesa. Isso pode ter acontecido porque os alunos já aprenderam primeiro a falar sobre si e só depois sobre os outros, ou seja, o uso de terceira pessoa ainda requer maior prática.

Nos fragmentos de falantes nativos raramente encontramos mais de dois adjetivos modificando um mesmo substantivo. Apenas 10% de ocorrências no corpus do *Collins* e nenhuma no do *BNC*. Quando isso ocorre, não parece haver uma regra fixa; a ordem dos adjetivos está relacionada com a intenção do falante. Já no corpus de aprendizes, percebemos que, ao descreverem o cabelo, eles geralmente usam três adjetivos e não demonstram um critério claro quanto à sua ordem.

Todas essas questões que surgiram a partir da análise são importantes para a elaboração das atividades, já que os aprendizes deverão chegar a conclusões semelhantes ao analisarem as linhas de concordância.

7. Elaboração de atividades didáticas

A criação das atividades foi inspirada grandemente no trabalho de Bertóli-Dutra (2002). A autora sugere algumas questões práticas que devem ser seguidas no primeiro contato dos aprendizes com as concordâncias. Primeiramente, evitamos utilizar a palavra concordância, utilizando fragmentos de frases no seu lugar.

Esse procedimento parece importante, pois os alunos já conhecem a palavra concordância dos seus estudos de gramática, e podem confundir o seu significado. A expressão "fragmentos de frases" também parece mais coerente, visto que nas linhas de concordância nem sempre temos as frases completas, mas partes de frases.

Outra sugestão de Bertóli-Dutra (op cit) adotada em nosso trabalho é que os aprendizes sejam orientados através de perguntas bem específicas durante os exercícios. Não seria muito produtivo se apresentássemos linhas de concordância e pedíssemos que eles relatassem o que puderam concluir.

Durante a comparação das concordâncias do corpus de aprendizes com as do site do *Collins* e do *BNC*, perguntas como estas foram feitas:

Quais as palavras encontradas à esquerda do núcleo?; Quais são os usos semelhantes aos usos identificados nos fragmentos de frases dos textos de vocês?; De acordo com o exercício

anterior, podemos criar uma regra sobre a ordem dos adjetivos em inglês?; Qual(is) seria(m)?; Verifique os fragmentos dos textos de vocês. Quais exemplos seguem a regra que você criou e quais são diferentes?; Reescreva as frases na forma correta.

Após a comparação das linhas de concordância guiadas por perguntas específicas, passamos a uma etapa final sugerida por Nesselhauf (2004): a aplicação de exercícios adicionais com o objetivo de praticar o fenômeno lingüístico. Entre os exemplos de exercícios, temos: organizar frases na ordem correta, completar espaços em branco e elaborar frases que seguissem as regras identificadas nas concordâncias.

Com o objetivo de não assustar os aprendizes com uma quantidade grande de dados, antes de lhes entregar as concordâncias fizemos uma simplificação dessas. Selecionamos algumas linhas mais significativas, pois, no corpus de aprendizes, havia muitas ocorrências semelhantes e, nas linhas de concordância de falantes nativos, podia-se perceber também a repetição de alguns padrões.

Escolhemos alguns exemplares de cada padrão identificado, o que diminuiu o tamanho das concordâncias. Esse procedimento foi sugerido por Hadley (2005), já que, ao trabalhar com iniciantes, a grande quantidade de dados pode fazer com que eles percam muito tempo na atividade tentando compreender cada frase ou palavra. Também reorganizamos algumas concordâncias, agrupando as linhas conforme sua semelhança, a partir da sugestão de Jacobi (2001).

8. Aplicação e avaliação das atividades

As atividades foram aplicadas durante o período de um mês, totalizando oito períodos de 45 a 50 minutos (dois períodos por semana). As turmas que participaram do experimento foram as mesmas que contribuíram com suas redações para a constituição do corpus de aprendizes.

Para introduzir essa nova metodologia de estudo, no primeiro dia, as linhas de concordância foram apresentadas no retro-projetor aos alunos e as atividades foram feitas com o grande grupo.

Como não havia um laboratório de informática na escola na época, as atividades e as concordâncias tiveram de ser impressas e distribuídas aos alunos. Eles trabalharam em duplas e, no final de cada atividade, a

correção foi feita com o grande grupo, a fim de esclarecer quaisquer dúvidas.

Ao saberem que as concordâncias tinham sido extraídas dos textos escritos por eles próprios, os aprendizes motivaram-se bastante para a realização das tarefas. Sua reação inicial ao observarem a primeira concordância foi tentar identificar os erros, rir e depois imaginar quem poderia ter sido o possível autor de cada erro. Salientamos que essa atitude foi espontânea; o professor não solicitou a identificação dos erros.

Quando trabalharam nas duplas, alguns se incomodaram com a grande quantidade de palavras novas. Mesmo tendo sido orientados para o fato de que a leitura deveria iniciar do meio para as suas palavras à direita e à esquerda, eles sentiam necessidade de compreender toda a frase ou fragmento de frase. Durante o estudo, eles podiam consultar o caderno e o livro, mas não tiveram acesso a dicionários.

Após as duas primeiras semanas, eles demonstraram certo cansaço em relação ao tipo de atividade, pois precisavam refletir e discutir com o colega sobre a língua. No final do período de um mês, foram aplicados exercícios de reforço das construções abordadas. Os aprendizes também responderam a um questionário aberto, em que manifestaram suas opiniões quanto aos pontos positivos e negativos do trabalho. Destacamos a seguir alguns dos comentários mais freqüentes:

- 1- Os trabalhos ajudaram a gente aprender mais a língua.
- 2- Os trabalhos são fáceis.
- 3- Podemos interagir com os colegas.
- 4- É legal de fazer.
- 5- Vimos coisas que não tínhamos aprendido antes.
- 6- É muito trabalhoso, muito puxado, deveríamos trabalhar mais com música.
- 7- Isso enche o saco.
- 8- Matéria difícil.
- 9- Muita conversa e nos atrapalhamos, pois não sabíamos traduzir as palavras.

Os alunos aprenderam mais sobre a língua, e alguns gostaram da interação com os colegas. Porém, provavelmente, por nunca terem realizado atividades dessa natureza, constataram que tiveram certa dificuldade e que o exercício exige esforço. Embora motivados no início do trabalho, o fato de analisarem suas próprias produções e a

permanência da mesma metodologia nas aulas seguintes tornaram o trabalho um tanto cansativo.

Essa foi uma experiência única para os alunos que realizaram essas atividades. Além de terem estudado a língua através de uma metodologia nova para eles, que os motivou bastante, os alunos desenvolveram outras competências, como o trabalho de pesquisa e o compartilhamento do conhecimento com os colegas.

Apontamos, como um desafio para uma futura aplicação, tornar as atividades mais interessantes, transformando as mesmas tarefas em brincadeiras ou em atividades, com as quais os alunos possam se movimentar pela sala ou mesmo trabalhar com músicas.

Comentários finais

Este trabalho mostra que é possível realizar atividades de análise da língua utilizando-se os recursos da Lingüística de Corpus, mesmo com alunos ainda em fase inicial do seu aprendizado em língua estrangeira.

O fato de as tarefas e as linhas de concordâncias terem sido adaptadas pelo professor evitou que os aprendizes se deparassem com concordâncias enormes que pudessem confundi-los. Outra atitude que os deixou mais seguros foi como o professor orientou as aulas, sempre buscando deixar claro o objetivo de cada tarefa.

Manter um registro das produções dos alunos pode servir como um valioso instrumento para o professor. Através da sua análise, é possível constatar em que pontos os aprendizes têm dificuldade e de que forma seu aprendizado progrediu.

A construção de um corpus de aprendizes também nos mostrou um novo modo de retornar os textos dos alunos de forma que eles reflitam sobre sua produção e sobre as estruturas da língua. Para o professor, é no estudo de corpus que se confirmam ou rejeitam as regras encontradas nas gramáticas.

De forma geral, a experiência foi positiva para os participantes, pois eles motivaram-se e tornaram-se mais independentes quanto ao aprendizado.

Para estimular essa postura de aluno pesquisador, mais trabalhos desse tipo devem ser realizados durante o ano. À medida que os alunos se familiarizarem mais com esses recursos, em um futuro experimento, pode-se tentar diminuir o controle das atividades, chegando a ponto de os alunos manejarem diretamente as concordâncias.

Vale salientar que a reflexão sobre a língua a partir das técnicas da Lingüística de Corpus foi apenas um momento da aula. Trabalhar somente dessa forma pode tornar as aulas cansativas e desinteressantes para os aprendizes. Em uma boa aula, vários métodos são utilizados de forma que diversas habilidades dos alunos sejam exploradas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALEXANDER, L. G. *Longman English Grammar*. New York: Longman, 2002.
- ASTON, G. Corpora and Language Teaching. In: BURNARD L., MCENERY T. (eds.) *Rethinking Language Pedagogy from a Corpus Perspective*. Frankfurt: Peter Lang, 2000.
- BERBER SARDINHA, T. *Lingüística de Corpus*. São Paulo: Manole, 2004.
- BERNARDINI, S. Corpora in the classroom. In: SINCLAIR, J. M. *How to Use Corpora in Language Teaching*. Studies in Corpus Linguistics. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2004.
- BERTÓLI-DUTRA, P. *Explorando a lingüística de corpus e letras de música na produção de atividades pedagógicas*. Dissertação de Mestrado em Lingüística Aplicada e Estudos da Linguagem - LAEL, Pontifícia Universidade Católica de São Paulo, 2002.
- BORBA, F. S. *Uma gramática de valências para o português*. São Paulo: Ática, 1996.
- CELCE-MURCIA, M.; LARSEN-FREEMAN, D. *The grammar book*. 2.ed. Boston: Heinle & Heinle, 1999.
- FERREIRA, W. R. *Deslexicalização no inglês de alunos brasileiros: um estudo baseado em corpora de aprendizes*. Dissertação de Mestrado em Lingüística Aplicada e Estudos da Linguagem - LAEL, Pontifícia Universidade Católica de São Paulo, 2003.
- GRANGER, S. The computer learner corpus: a versatile new source of data for SLA research. In: GRANGER, S. (org.) *Learner English on computer*. New York: Longman, 1998.
- GRANGER, S., HUNG, J., PETCH-TYCON, S. (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. John Benjamins Publishing, 2002.
- GRANGER, S., TRIBBLE, C. Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In:

- GRANGER, S. (org.) *Learner English on computer*. New York: Longman, 1998.
- HADLEY, G. *Sensing the winds of change: an introduction to Data-Driven learning*. Disponível em: <<http://web.archive.org/web/20030404025654/web.bham.ac.uk/johnstf/winds.htm>> Acesso em: 04 fev. 2005.
- JACOBI, C. C. B. *Lingüística de corpus e ensino de espanhol a brasileiros*: Descrição de padrões e preparação de atividades didáticas. Dissertação (Mestrado em Lingüística Aplicada e Estudos da Linguagem), LAEL, Pontifícia Universidade Católica de São Paulo, 2001.
- LEECH, G., STARTVIK, J. *A Communicative Grammar of English*. Londres: Longman, 1977.
- NEVES, M. H. M. *Gramática de usos do português*. São Paulo: UNESP, 2000.
- NESSELHAUF, N. Learner corpora and their potential for language teaching. In: SINCLAIR, J. M. *How to Use Corpora in Language Teaching*. Studies in Corpus Linguistics. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2004.
- SINCLAIR, J. (ed.) *Collins Cobuild English grammar*. Londres: Collins, 1990.
- STEVENS, V. Concordancing with Language Learners: Why? When? What? *CAELL Journal*, 6:2:2-10, 1995.

Recebido em setembro de 2007
Aprovado em maio de 2008