

A elaboração de uma atividade didática de língua inglesa direcionada por dados de aprendizes

DOI: <http://dx.doi.org/10.21165/el.v49i1.2568>

William Danilo Garcia¹

Paula Tavares Pinto²

Resumo

Este artigo objetiva apresentar uma tarefa desenvolvida no intuito de auxiliar o processo de aprendizagem de inglês como língua estrangeira por estudantes universitários de um curso de Letras. Para isso, um *corpus* de aprendiz foi compilado a partir da escrita autêntica desses alunos com a intenção de coletar informações sobre como eles usam algumas preposições da língua inglesa. Em seguida, após análise dessas informações com o auxílio da ferramenta computacional AntConc, uma tarefa com base nos preceitos conceituais da Aprendizagem Direcionada por Dados foi elaborada no sentido de melhorar o uso dessas preposições. Como suporte teórico deste estudo, recorreu-se a Granger (1998, 2002, 2013), Johns (1991, 1994), Berber Sardinha (2011) e Boulton (2010). Com relação aos resultados deste estudo, foi possível desenvolver tarefas das quais os graduandos se colocaram na posição de professor para investigar, por meio do uso de um *corpus*, se os usos das preposições trazidas em algumas linhas de concordância estavam adequados ou não. Desse modo, este estudo sugere que a Aprendizagem Direcionada por Dados pode contribuir para o processo ensino-aprendizagem de certas preposições do inglês, pois desperta nos alunos um perfil mais autônomo com relação à investigação linguística.

Palavras-chave: *corpus* de aprendiz; aprendizagem direcionada por dados; tarefa; preposições.

1 Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), São José do Rio Preto, São Paulo, Brasil; wil.danilo@hotmail.com; <https://orcid.org/0000-0001-6898-1931>

2 Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), São José do Rio Preto, São Paulo, Brasil; paula.pinto@unesp.br; <https://orcid.org/0000-0001-9783-2724>

Designing a teaching activity for the English language driven by learners' data

Abstract

The purpose of this paper is to present a task designed to assist undergraduate students from a language teaching course in their process of learning English as a Foreign Language. To do so, their free compositions were compiled in a learner corpus in order to analyze how these students use some prepositions in English. Later, having gathered the information from corpus analysis with the aid of AntConc software, a task under the conceptual basis of Data-Driven Learning was developed to improve the use of such prepositions. As theoretical support, we referred to Granger (1998, 2002, 2013), Johns (1991, 1994), Berber Sardinha (2011) and Boulton (2010). Concerning the results of this study, it was possible to develop some tasks in which students played the role of a teacher to analyze, through corpus database, whether the prepositions shown by the concordance lines were appropriately used. Therefore, these findings suggest that Data-Driven Learning can contribute to the learning-teaching process of such prepositions in English because it provides the students with a more autonomous profile in relation to linguistics investigation.

Keywords: Learner *Corpus*; Data-Driven Learning; Task; Prepositions.

1. Introdução

Granger (2002) afirma que dados de aprendizes levantados em um contexto artificial – em que o aluno apenas utiliza a língua a partir de técnicas de elicitación³ – não configuram um *corpus* de aprendiz autêntico. A autora ainda destaca a questão de o contexto de ensino de inglês como língua estrangeira apresentar certa artificialidade, o que pode fazer com que o *corpus* de aprendiz possa ser visto como menos autêntico do que o *corpus* nativo. No entanto, a autora postula que o uso das *free compositions* auxilia essa problemática da autenticidade, uma vez que a liberdade de escolha sobre o que escrever favorece o uso “natural” da língua pelo aprendiz.

Desta forma, este artigo apresenta uma proposta de compilação de um *corpus* de aprendiz de graduandos do primeiro ano de um curso de Letras – realizada a partir de uma escrita mais livre (as mencionadas *free compositions*) –, atribuindo-lhe, portanto, mais autenticidade, se analisado sob a reflexão de Granger (2002). O intuito é que esse *corpus* de aprendiz autêntico possibilite, por meio da Aprendizagem Direcionada por

3 De acordo com o *site* do British Council, o termo designa uma gama de técnicas que possibilitam que o professor obtenha informações dos alunos ao invés de lhes fornecê-las. Disponível em: <https://www.teachingenglish.org.uk/article/eliciting>. Acesso em: 19 maio 2019.

Dados (no inglês *Data-Driven Learning*), a elaboração de uma atividade didática que auxilie o processo de aprendizagem dos usos das preposições.

No que se refere aos procedimentos metodológicos adotados, foi inicialmente elaborada e aplicada uma aula em que os alunos redigiram uma *free composition*. Em seguida, os dados desses aprendizes foram analisados com o auxílio do concordanciador *AntConc*, objetivando identificar desvios relacionados às preposições. Feito isso, elaborou-se, em um terceiro momento, uma atividade direcionada por dados a ser aplicada aos alunos no intuito de fazê-los analisar seus próprios desvios dos usos das preposições de modo mais autônomo e utilizando uma abordagem que desperta um perfil pesquisador nos graduandos.

Sendo assim, o objetivo era que, a partir dos dados linguísticos dos aprendizes, fossem levantados os desvios com relação às preposições *with, in, on, at, for* e *to* encontradas na escrita desses alunos para que atividades didáticas direcionadas por esses dados fossem desenvolvidas e, posteriormente, aplicadas a esses aprendizes na intenção de auxiliá-los no processo de aprendizagem desses itens gramaticais.

A escolha pela análise e desenvolvimento de atividades interventivas no que se refere ao uso das preposições se mostra relevante pelo fato de elas serem um tópico de dificuldade aos alunos. Conforme Swan (2005, p. 425, tradução nossa⁴), “é difícil aprender a usar as preposições corretamente em uma língua estrangeira. Muitas preposições do inglês têm várias funções diferentes [...] e elas podem corresponder a diversas preposições diferentes em outra língua.”.

Por fim, para melhor compreensão dos embasamentos envolvidos na compilação de um *corpus* de aprendiz e no que se baseia a Aprendizagem Direcionada por Dados, faz-se necessária certa revisão bibliográfica, que se encontra em seguida.

2. A Linguística de *Corpus* e o ensino

Inicialmente, julga-se pertinente a consideração acerca da relação da Linguística de *Corpus* com o ensino de línguas. Como destaca Berber Sardinha (2000, p. 325), a Linguística de *Corpus* teve, no início do século XX, forte relação com a área de ensino de línguas em pesquisas como a de Thorndike. O autor reforça ainda que, embora a LC tenha mudado seu foco para a descrição da linguagem nos anos seguintes, voltou a dar ênfase a pesquisas relacionadas ao ensino, especialmente a partir da década de 90, momento em que a pesquisadora Granger lança um olhar mais minucioso na área conhecida como *corpus* de aprendiz.

4 No original: “It is difficult to learn prepositions correctly in a foreign language. Most English prepositions have several different functions [...] and these may correspond to several different prepositions in another language.”.

Com o objetivo de entender melhor sobre o *corpus* de aprendiz, já que é um dos fatores precursores da reutilização da LC no ensino e também o *corpus* que se foi compilado para este artigo, a citação de Granger (2013, p. 1, tradução nossa⁵) se faz relevante: “*corpora* de aprendizes são coleções eletrônicas de dados linguísticos produzidos por aprendizes de uma segunda língua, sejam eles aprendizes de segunda língua ou de língua estrangeira”. Como se pode notar, o que difere o *corpus* de aprendiz é o fato de ele ser compilado a partir da produção de alguém que não seja nativa daquele idioma.

Granger (1998, 2002), além de definir o *corpus* de aprendiz e assimilá-lo a uma nova forma de se gerar informações relacionadas ao processo de aquisição de uma língua, afirma que as contribuições do *corpus* para o ensino de língua inglesa se dão a partir da compilação de *corpora* de nativos (que são revertidos em materiais que trazem textos autênticos). Porém, reconhece ainda que a utilização de *corpus* de aprendiz também oferece vantagens para o ensino, pois evidencia as dificuldades linguísticas encontradas pelos alunos.

Ainda conforme Granger (1998), sua ressalva com relação ao fato de um *corpus* de aprendiz raramente atingir uma proporção gigantesca se mostra válida mesmo em tempos atuais, afinal há diversas limitações com relação ao levantamento de um *corpus* de aprendiz de larga escala.

Entendido do que se trata o *corpus* de aprendiz e considerado que ele foi um dos responsáveis por reaproximar a LC da área de ensino de línguas, faz-se necessário maior explanação referente à utilização do *corpus* (seja ele de aprendiz ou não) em sala de aula.

Em sua publicação de 2011, Berber Sardinha traça um panorama sobre o uso da Linguística de *Corpus* no contexto de ensino, indicando três diferentes formas de se elaborar atividades didáticas pautadas na LC. A primeira delas se origina, conforme o autor, nas décadas de 80 e 90 com o que ficou conhecido como *classroom concordancing*. Fica bastante evidente que esse período do uso dos concordanciadores em sala de aula remete ao consagrado Johns e à Aprendizagem Direcionada por Dados (conhecida no inglês como DDL).

Como descreve Johns (1991), essa abordagem evidencia o uso do concordanciador como uma forma de informante em um processo passivo, no qual a resposta é provida a partir de um questionamento do aprendiz. Ainda no sentido de uma definição para a DDL, Johns (1991, p. 2, tradução nossa⁶) alega que “a Aprendizagem Direcionada por Dados fornece ao aluno o acesso a fatos da performance linguística.”

5 No original: “Learner corpora are electronic collections of language data produced by L2 learners, that is, second or foreign-language learners.”

6 No original: “[...] and that is a data-driven approach which gives the learner access to the facts of linguistic ‘performance’.”

Johns (1991) acrescenta também o fato de o concordanciador não ser uma novidade, porém o que se mostrou inovador na época foi a descentralização do acesso à ferramenta, feito apenas por pesquisadores até então, tornando-o disponível aos alunos e os colocando em uma posição em que um perfil de pesquisador se mostre necessário.

Em conformidade com Johns (1994), a DDL busca afastar a necessidade de um intermediador e aproximar o aprendiz ao dado linguístico por um acesso direto. Como resultado dessa aproximação, Johns (1991) apresenta que a Aprendizagem Direcionada por Dados se postula como um processo de aprendizagem de línguas em que o aluno faz especulações sobre a língua, que o ajuda a ver os padrões linguísticos e até mesmo a fazer generalizações sobre o funcionamento da língua.

No referente ao papel do professor sob as considerações da DDL, destaca-se também certa mudança de papéis, por assim dizer. Segundo Johns (1991, p. 3, tradução nossa⁷), o professor “tem que se tornar um diretor e um coordenador de pesquisas iniciadas pelos alunos.”. Johns (1994) descreve ainda que o professor tem que deixar a posição de detentor do conhecimento para assumir essa posição de alguém que organiza a pesquisa.

Com os princípios fundamentais da DDL expostos, mostra-se também relevante evidenciar algumas das reflexões mais atuais acerca de como utilizar essa abordagem de forma eficaz na sala de aula. Boulton (2010) apresenta uma opção de operacionalização de atividades DDL em que o contato do aprendiz com o dado linguístico seja feito por meio de materiais preparados pelo professor ou pesquisador. Embora alguns teóricos acreditem que a não utilização do computadores para acessar os concordanciadores possa ferir, de certo modo, a essência da Aprendizagem Direcionada por Dados, Boulton (2010) aponta que esses materiais preparados pelos professores oferecem também benefícios e se mostram possíveis de serem utilizados por aprendizes com baixo nível linguístico, sem mencionar que podem ser preparados por professores que não tenham vasto conhecimento da Linguística de Corpus. Nos dizeres do próprio autor, “eliminar o computador não representa menosprezar a base conceitual da DDL, mas sim tornar a tarefa do aprendiz mais fácil.” (BOULTON, 2010, p. 559, tradução nossa⁸).

Por fim, visto que o contato com o dado linguístico pode ocorrer por meio de materiais preparados por professores e ainda sim representar uma postura de pesquisador nos alunos, como defende Boulton (2010), torna-se necessário apontar algumas formas com as quais as atividades que exploram a concordância podem ser elaboradas. Há, conforme

7 No original: “The second main effect of DDL is on the role of the teacher, who has to learn to become a director and a coordinator of student-initiated research.”.

8 No original: “We have argued that eliminating the computer from the equation, far from fatally undermining the conceptual basis of DDL, can in fact make the learner’s task considerably easier.”.

Berber Sardinha (2011) as que são integrais (com todas as ocorrências, as selecionadas (com apenas algumas ocorrências), as editadas (em que as linhas são modificadas), as preenchidas (com exibição da palavra de busca) e as lacunadas (em que a palavra de busca não é disponibilizada).

Em conclusão, viu-se, nesta seção, que o reaparecimento do uso da LC no ensino se deu por estudos relacionados aos *corpora* de aprendizes e também pela abordagem que permite que o aluno acesse dados linguísticos em busca de respostas a perguntas das quais ele queira maior entendimento. Quanto ao professor e/ou pesquisador, cabem algumas escolhas, tais como, tornar esse acesso mais direto ou não ao se escolher usar materiais preparados previamente ao invés de levar o aluno ao computador e acessar dados linguísticos por um concordanciador. Será evidenciado em seções seguintes que esta pesquisa opta por uma mescla de atividades com informações previamente preparadas (*hands off*) e aquelas em que o aluno acesse diretamente o *corpus* de referência para obter respostas (*hands on*).

3. Metodologia

Esta seção objetiva descrever a metodologia envolvida em dois procedimentos significativos para o desenvolvimento da tarefa de análise dos dados, que almeja colaborar no processo de aprendizagem das preposições dos alunos do primeiro ano do curso de Letras.

O primeiro momento está relacionado à compilação dos textos dos alunos, feita de forma a se levantar um *corpus* de aprendiz a ser analisado com auxílio do concordanciador AntConc. Com relação ao segundo passo metodológico, esse se refere a como os dados dos aprendizes foram analisados em comparação aos dados do *corpus* de referência, o COCA. Análise essa que também norteou a criação da atividade didática direcionada por dados com ênfase nas preposições.

Antes de descrever, porém, o estágio inicial, referente à operacionalização necessária para o levantamento do *corpus* de aprendiz, julga-se válida uma descrição dos alunos e do contexto em que eles estiveram inseridos. Sendo assim, é válido dizer que se tratou de uma turma de primeiro ano do curso de Letras e que cursava a disciplina de Língua Inglesa. Estima-se que o total de matrículas na disciplina ficou em torno de 15 alunos, embora apenas nove tenham participado assiduamente das aulas. O nível dos alunos se mostrou heterogêneo, pois eles variaram entre uma faixa de nivelamento mais básica (A1 e A2) àqueles níveis mais independentes (B1 e B2).

Com relação ao *corpus* de aprendiz desses alunos, ele se pautou em uma aula com duração aproximada de 1h30 e cuja tarefa dos graduandos foi a de produzir uma

*fanfiction*⁹ como forma de se alcançar as *free compositions* sugeridas por Granger (2002). No que tange à alimentação desses textos escritos pelos alunos do primeiro ano de Letras na ferramenta computacional AntConc, o procedimento metodológico não trouxe nenhum aspecto diferente da operacionalização usual do *software*. Esses textos foram compilados em um único arquivo em formato de “texto sem formatação” (txt), que foi inserido no concordanciador por meio da aba “file”, e da seleção da opção *open file(s)*, ou por meio do atalho Ctrl+F. Esse procedimento de inserção do *corpus* de aprendiz no AntConc pode ser visualizado na figura a seguir.

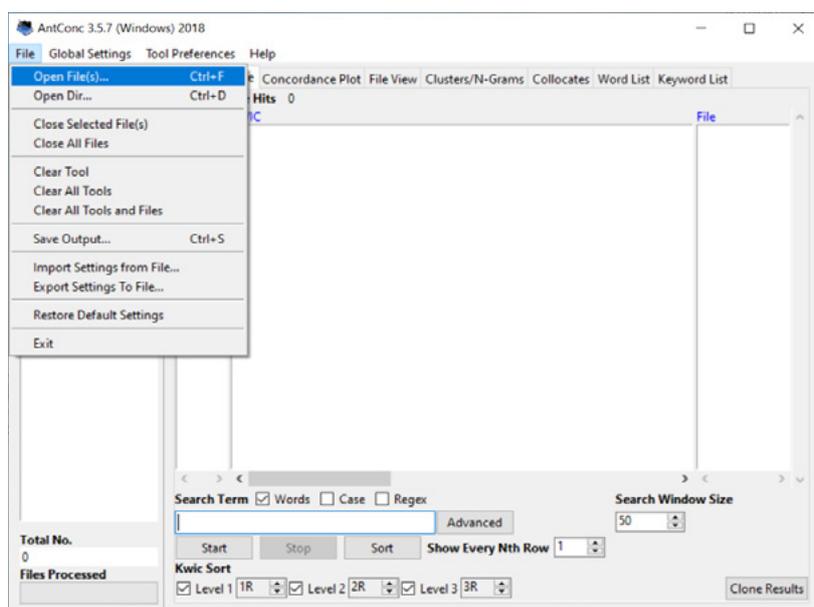


Figura 1. Inserção de textos no AntConc
Fonte: AntConc

Feito isso, o programa abre uma nova janela para buscar o arquivo que se deseja abrir. Um detalhe importante é que esse arquivo deve estar salvo no computador. Assim que se escolhe o documento txt, ele é introduzido no AntConc e aparecerá no canto superior esquerdo da tela. Em seguida, deve-se escolher a opção que melhor represente o objetivo da pesquisa, que no caso foi o de gerar a lista de palavras. Para isso, escolheu-se a opção *Word List* e, então, clicou-se no botão *start*. Veja a figura a seguir.

9 Como descreve Neves (2014, p. 100), “as *fanfictions* ou simplesmente *fanfics* são histórias alternativas em prosa [...] escritas por fãs de determinada série ou *fandom*”.

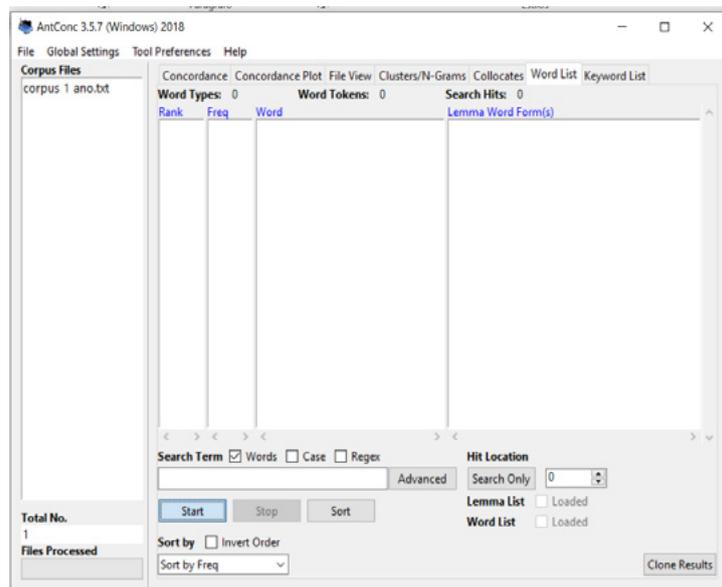


Figura 2. Como gerar listas de palavras
Fonte: AntConc

Em seguida, a partir da relação das palavras evidenciadas na *Word List*, buscou-se pelas preposições de interesse neste estudo (*with, in, on, at, for* e *to*) de forma a gerar as linhas de concordância de cada uma delas. Para isso, o procedimento é inserir o item que se deseja buscar no campo *Search Term* e clicar no botão *Search Only*. Veja o procedimento de busca a seguir.

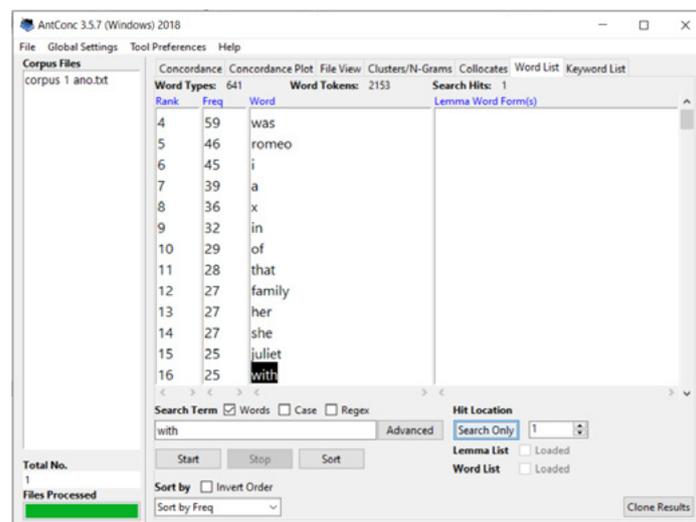


Figura 3. Busca de um termo na *Word List*
Fonte: AntConc

Como se observa na figura, o termo buscado recebe esse destaque na cor preta. Para gerar, então, as linhas de concordância desse termo, basta clicar na palavra destacada.

Há diferentes possibilidades de se gerar as linhas de concordância, porém nesta pesquisa a escolha foi pela busca por uma palavra à esquerda e duas à direita na opção *Kwic Sort*. Isso significa que as linhas foram organizadas de forma alfabética com base na primeira palavra que ocorresse à esquerda da preposição buscada (destaque em vermelho); e nas duas palavras posicionadas à direita, que também receberam destaque nas linhas de concordância (realçadas em verde e em roxo).

A seguir, tem-se uma ilustração da linha de concordância da preposição *with* com as palavras que ocorrem tanto à esquerda quanto à direita destacadas.

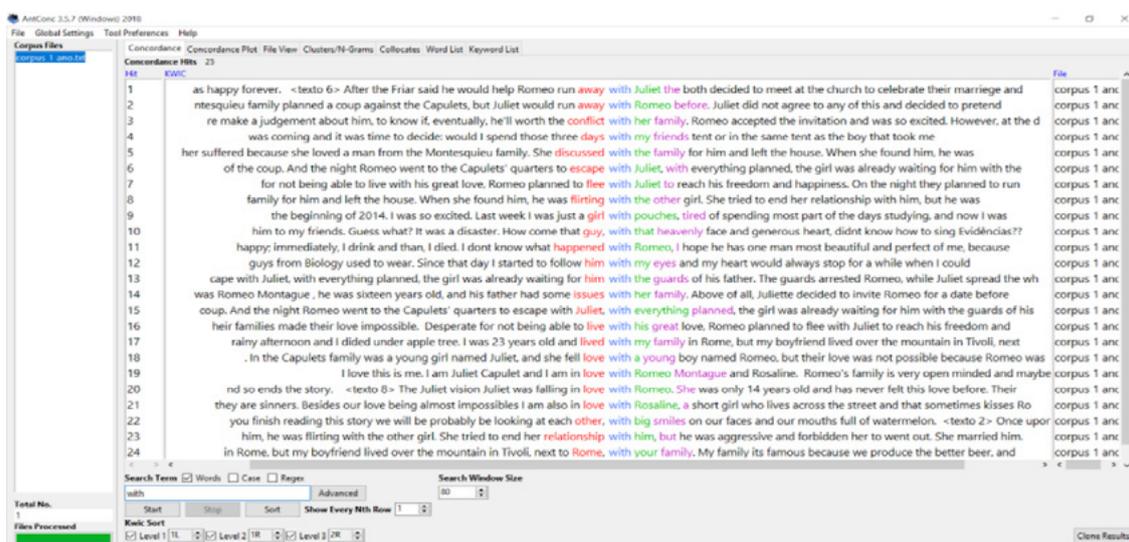


Figura 4. Captura de tela da linha de concordância da preposição *with*
Fonte: AntConc

Os destaques, embora nem sempre suficientes para a realização das análises das preposições, facilitam a visualização das palavras que ocorrem junto à preposição e norteiam uma primeira observação no sentido de analisar se a preposição está usada de forma adequada ou não. Obviamente, em inúmeras situações um co-texto que além dessas palavras destacadas foi utilizado, de modo que foi preciso recorrer a uma análise mais minuciosa em quase toda a linha.

Os mesmos procedimentos foram repetidos com as outras preposições, levantando as linhas de concordância de cada uma delas, no intuito de analisar como os alunos empregaram cada um desses itens gramaticais em seus textos. Essa análise, portanto, configurou o segundo estágio do qual esta seção se propõe a descrever.

Faz-se preciso esclarecer que cada uso das preposições *with*, *in*, *on*, *at*, *for* e *to* foi, a partir da observação da linha de concordância, analisado pelos pesquisadores para que se concluísse os que foram feitos de forma adequada e os que não. Para isso, recorreu-se à utilização de dicionários *on-line* (Cambridge, Macmillan e Oxford), além do *Corpus of Contemporary American English*, o COCA.

Vale ressaltar que essas análises, apesar de parecerem complexas em termos de demandar muito tempo, são práticas, pois o AntConc apresenta todas as linhas de concordância com uma preposição em específico em uma mesma tela. Desse modo, os pesquisadores puderam observar, de modo prático, cada caso e consultar nos dicionários *on-line* e no COCA aqueles que lhes causaram dúvidas.

Assim, quando houve a necessidade, os pesquisadores consultaram o *corpus* de estudo. Portanto, considera-se relevante a disponibilização do acesso ao COCA, que se faz por meio do *site* corpus.byu.edu/coca. Uma vez acessada essa página, o usuário se depara com a seguinte tela:



Figura 5. Página inicial do COCA
Fonte: corpus.byu.edu/coca

Como se observa na figura, aparece a mensagem “not logged in”, que significa que o usuário não está logado. Caso seja o primeiro acesso, será necessário se registrar. Para isso, é necessário que se faça um cadastro para a utilização do *corpus*.

Evidentemente o usuário não sabe qual é o comando que o *site* utiliza para indicar que se busca por todas as preposições posicionadas logo após o verbo *go*, como foi o exemplo dado anteriormente. Por esse motivo, o *site* traz uma aba intitulada POS que serve como uma espécie de legenda para esses comandos. Pode-se visualizar essa aba na figura anterior, em que o cursor do *mouse* selecionou a opção prep.ALL, que representa a abreviação do termo preposição. Uma vez clicada nessa opção, gera-se, no campo de busca, o comando `_i*` que indica que se busca por qualquer preposição que se combina com o verbo *go*. Isso só é possível, pois o COCA traz a etiquetagem por POS, que como descreve Frankenberg-Garcia (2016) corresponde a *part-of-speech*. Essa etiquetagem, basicamente, divide o léxico por suas classes de palavras.

O resultado dessa busca (`go _i*`) pode ser visto na figura 7.



Figura 7. Resultado da busca `go _i*`

Fonte: corpus.byu.edu/coca

Outro procedimento importante e que facilitou a consulta dos pesquisadores em alguns momentos foi o fato de o COCA também possibilitar a busca lematizada. Assim, se o pesquisador tiver a intenção, por exemplo, de saber quais preposições são mais frequentemente posicionadas após o verbo *go* em qualquer uma de suas possíveis conjugações, basta colocar esse verbo entre colchetes e em seguida inserir o comando POS da preposição (`_i*`) separado por um espaço. Veja o procedimento a seguir.

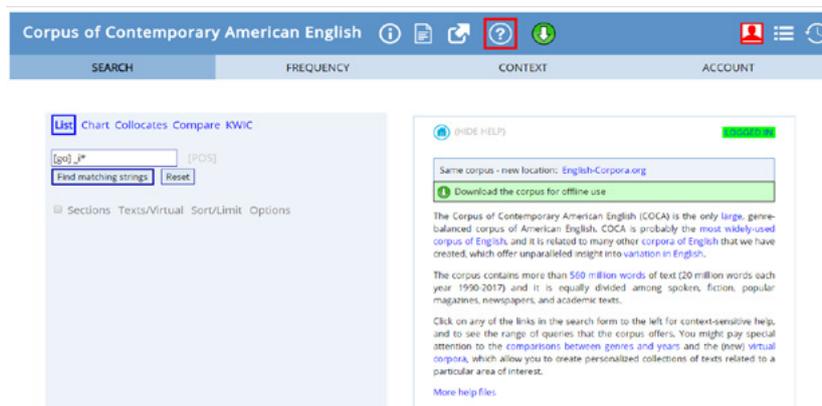


Figura 8. Busca lematizada
Fonte: corpus.byu.edu/coca

Já no caso contrário, considerando, por exemplo, uma busca em que se queira analisar as preposições que acontecem antes de um verbo ou de outra palavra, basta apenas inverter a ordem da busca, de modo que o comando `_i*` seja inserido primeiro e, após uma separação de um espaço, a palavra ou o verbo seja incluída em seguida.

No caso da busca que vise descobrir quais palavras, independentemente da classe gramatical, se colocam a uma outra, basta utilizar o asterisco. Assim, se o objetivo for saber, por exemplo, quais são as palavras que mais frequentemente se colocam após o verbo *go*, a busca deve ficar: `go *`. Pode-se saber também quais as palavras que mais frequentemente antecedem uma outra. Desse modo, caso haja a necessidade de saber quais palavras se colocam mais comumente antes do verbo *go*, por exemplo, o comando para a busca deve ser o asterisco seguido pelo verbo, como em: `* go`.

Em conclusão, nota-se que as possibilidades de busca no COCA são bastante diversificadas, no entanto, tentou-se aqui oferecer elucidações relacionadas às formas de busca mais frequentemente utilizadas pelos pesquisadores. É válido destacar ainda que esse manuseio do COCA também foi feito pelos alunos do primeiro ano do curso de Letras, porém após os pesquisadores terem concluído as análises de cada preposição e elaborado a atividade com ênfase nos casos em que se observou o uso inapropriado de algum desses itens.

4. Resultados

Esta seção traz os dados obtidos com a compilação e análise do *corpus* de aprendiz, a elaboração, a partir desses dados, da atividade didática com ênfase nas preposições *with, in, on, at, for* e *to* e a aplicação dessa tarefa em sala de aula.

No concernente ao *corpus* de aprendiz, compilou-se, a partir de uma produção de língua autêntica em que os alunos escreveram uma *free composition*, um *corpus* com 649 *word types* e 2099 *word tokens*. Desse total, houve 160 ocorrências das preposições observadas, sendo 13 dessas ocorrências consideradas usos inadequados quando analisados em comparação ao COCA, *Grammar for English Language Teachers* (PARROTT, 2000), *Student Grammar of Spoken and Written English* (BIBER et al., 2002) e *Practical English Usage* (SWAN, 2005), ou aos dicionários *on-line* Cambridge, Oxford e Macmillan.

Em uma análise pormenorizada, observa-se as seguintes informações:

Tabela 1. Preposições analisadas

Preposição	Total de ocorrências	Uso inadequado
With	25	1
In	32	3
On	10	1
At	7	1
For	16	7
To	70	0

Fonte: Elaboração própria

Percebe-se, ao observar a tabela, que apesar de haver uma alta frequência no uso da preposição *to*, essa não apresentou nenhum uso indevido de acordo com as gramáticas e os dicionários *on-line* já referenciados ou o COCA. Entretanto, a preposição *for*, que apresentou uma frequência relativamente mais baixa no total de ocorrências, foi a que mais obteve usos inadequados. Por meio das análises individuais de cada linha de concordância, conclui-se que a preposição *for* foi usada onde se deveria usar a preposição *to*, como se vê na linha de concordância a seguir.

"Above of all, Juliette decided to invite Romeo **for** a date before make a judgement about him"

Nesse caso, tanto uma pesquisa no COCA quanto no dicionário *on-line* Cambridge confirma que a regência verbal de *invite* é a preposição *to*. Isso aconteceu em cinco dos sete usos inapropriados da preposição *for* e evidencia, portanto, que apesar de a preposição *to* não ter sido usada inadequadamente em nenhuma de suas ocorrências, houve situações consideradas inadequadas pela falta de seu uso.

Sobre o segundo passo – a elaboração da atividade direcionada por esses dados e com ênfase no auxílio do processo de aprendizagem das preposições – o resultado foi a criação de uma tarefa. Essa atividade oferece aos graduandos do primeiro ano de Letras uma situação hipotética em que eles têm que se posicionar não mais como alunos da graduação, mas como professores de língua inglesa. Professores esses que precisam fazer correções e dar *feedback* aos seus alunos com base em linhas de concordâncias retiradas de uma produção textual desses supostos alunos. No entanto, essa simulação também prescreve que esse professor analisa a língua por meio de dados linguísticos reais, ou seja, que recorre ao uso real da língua. Assim, para que esses dados sejam acessados, esse professor estipulado na tarefa deve acessar algum *corpus* (no caso deste artigo o COCA) para ser capaz de dar o *feedback* preciso aos seus alunos.

A atividade completa está disponibilizada na seção anexo, porém a proposta da tarefa desenvolvida pode ser visualizada a seguir:

TASK (1st year)

You are an English teacher who takes **real language data information** into account when preparing your classes and when giving your students feedback. Below, you will find a list of sentences produced by some of your students. Using the **Corpus of Contemporary American English (COCA)** as a reference to your language analysis, prepare your feedback to your students in terms of which **prepositions** were used appropriately and which ones were not. Justify your explanations with real and authentic data and/or evidence taken from the corpus.

Uma observação interessante na dinâmica dessa tarefa é que, a partir dessa simulação proposta pela atividade, os graduandos do curso de Letras analisam, na verdade, linhas de concordância produzidas nas *free compositions* que eles mesmos redigiram. Assim, os princípios da Aprendizagem Direcionada por Dados são vistos como condutores dessa tarefa, uma vez que os alunos precisam atuar com um perfil pesquisador e autônomo para acessarem o COCA e obterem informações quanto ao funcionamento da língua.

Outra ressalva é a de a atividade direcionada por dados elaborada conter 15 linhas de concordância, ou seja, duas linhas a mais do total de erros delimitados na análise do *corpus* de aprendiz compilado na pesquisa. Isso se explica pelo fato de se considerar relevante que os graduandos possam concluir, a partir da pesquisa no COCA, que nem todas as linhas de concordância apresentam inadequações. Desse modo, a busca no *corpus* de referência pode confirmar que em duas situações as preposições foram utilizadas pelos alunos em conformidade com o uso real que se faz da língua. Isso traz a contribuição também de os professores em formação perceberem que o professor de línguas pode (e deve) dar *feedback* naquilo que o aluno produz assertivamente, e não apenas nos casos de erros ou usos indevidos.

Há de se considerar também que a atividade proposta se caracteriza por um material preparado previamente pelos pesquisadores, configurando um procedimento *hands off*, ou seja, em que o acesso às linhas de concordâncias (consideradas integrais) não se faz de forma direta aprendiz-computador. Obviamente que isso ocorre pelo fato de a pesquisa já ter um critério de análise definido: as preposições. No entanto, o momento seguinte, em que os alunos precisam consultar o COCA para obterem as informações necessárias para darem *feedback* aos supostos alunos, esse contato com as linhas de concordância ocorre de forma direta (*hands on*), ou seja, sem a utilização de materiais previamente elaborados, são os alunos e os dados do *corpus*.

Com relação à terceira ênfase desta seção, que envolve a aplicação dessa proposta aos alunos, fazem-se necessárias algumas colocações, tais como, o fato de ela envolver duas aulas, ao invés de uma. Além disso, destaca-se que para que os alunos conseguissem realizar a tarefa com êxito, o professor estabeleceu a proposta e, logo em seguida, proporcionou-lhes as explicações de como manusear o COCA (vide seção 3). Dadas essas elucidações, os alunos se reuniram em duplas para acessarem o COCA e analisarem cada linha de concordância de forma a identificarem se o uso por ela evidenciado está adequado ou não e, por fim, elaborar o *feedback* que dariam ao aluno. Evidencia-se, então, um requisito importante com relação à estrutura física para a realização dessa tarefa, que é a necessidade de uma sala de aula com computadores ou um laboratório de informática.

Como conclusão da aula, o professor/pesquisador pediu para que cada dupla compartilhasse as descobertas feitas a partir da busca no COCA, assim como o *feedback* que eles dariam a seus supostos alunos. Esse foi um momento enriquecedor, pois os graduandos puderam contrastar as informações levantadas por cada uma das duplas, tais como, os métodos e os comandos utilizados nas consultas, e os resultados obtidos.

Além disso, essa etapa final da aula também contemplou uma discussão feita de forma aberta com a sala toda a respeito da atividade. Com isso, os alunos relataram que esse tipo de busca em um *corpus* era algo novo e que causou certo estranhamento de

início, porém, após algumas buscas, eles já conseguiram se sentir mais habituados à ferramenta. Eles afirmaram ainda que acharam esse tipo de pesquisa uma maneira eficaz de se tirar dúvidas com relação ao funcionamento da língua e que conseguem ver a aplicabilidade da busca em um *corpus* para aspectos linguísticos que vão além das preposições.

5. Conclusão

Como visto durante o artigo, nota-se que as ferramentas da Linguística de *Corpus* (tal como o concordanciador AntConc) podem ser usadas em prol da análise da produção linguística dos alunos de uma língua estrangeira. Além disso, há de se destacar que as contribuições desta área da Linguística podem ir além. Isso se evidencia especialmente se considerada a ênfase aqui proposta de delimitar usos inapropriados de cinco preposições (*with, in, on, at, for* e *to*) para que uma atividade didática amparada pelos princípios da Aprendizagem Direcionada por Dados fosse desenvolvida e aplicada aos graduandos de Letras. Isso tudo visando melhorar o processo de aprendizagem desses referidos itens gramaticais ao expô-los a dados de um *corpus*, o que proporciona o desenvolvimento de um perfil de investigação linguística mais autônomo. Ressalta-se, ainda, que essa exposição aos dados de funcionamento da língua não precisa acontecer de forma a causar estranhamento nos alunos. Na verdade, ela se mostra mais benéfica se feita de modo mais ameno, pois faz com que o aluno se sinta mais confortável para realizar as suas buscas, o que explica, portanto, o porquê da intercalação dos procedimentos *hands off* e *hands on*.

Por fim, destaca-se que não só a produção de atividades didáticas a partir de dados linguísticos reais dos alunos é possível. Tangível também se mostra a utilização dessa mesma ferramenta computacional (o AntConc) para traçar, futuramente, o quanto essa atividade com foco nas preposições contribuiu para o desenvolvimento do processo de aprendizagem dessa classe gramatical (as preposições). Sendo assim, um possível desdobramento deste estudo poderia visar a compilação, após a aplicação da tarefa proposta neste artigo, de um novo *corpus* de aprendiz com esses mesmos graduandos do primeiro ano do curso de Letras e em situação semelhante (envolvendo uma produção de uma *fanfiction*). Assim, o objetivo do levantamento desse segundo *corpus* de aprendiz seria o de confirmar com uma análise quali-quantitativa se a atividade direcionada por dados aqui desenvolvida beneficia, de fato, o uso das preposições da língua inglesa pelos alunos.

REFERÊNCIAS

- ANTHONY, L. *AntConc*. Versão 3.5.7. Tokyo: Waseda University, 2018. Computer Software. Disponível em: www.laurenceanthony.net/software/antconc/. Acesso em: 10 ago. 2018.
- BERBER SARDINHA, T. Como usar a Linguística de Corpus no ensino de língua estrangeira. In: TAGNIN, S. E. O.; VIANA, V. (org.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2011. p. 301-356.
- BERBER SARDINHA, T. Linguística de Corpus: histórico e problemática. *Delta*, São Paulo, v. 16, n. 2, p. 323-367, 2000.
- BIBER, D. *et al.* *Student grammar of spoken and written English*. Harlow: Pearson Education, 2002.
- BOULTON, A. Data-driven learning: taking the computer out of the equation. *Language Learning*, p. 534-572, 2010.
- FRANKENBERG-GARCIA, A. Corpora in ELT. In: HALL, G. (ed.). *The routledge handbook of English language teaching*. Abingdon: Routledge, 2016. p. 383-398.
- GRANGER, S. Learner Corpora. In: CHAPELLE, C. A. (ed.). *The encyclopedia of Applied Linguistics*. Oxford: Blackwell Publishing, 2013.
- 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*. Amsterdã: John Benjamins, 2002. p. 03-33.
- GRANGER, S. The computer learner corpus: a versatile new source of data for SLA research. In: GRANGER, S. (ed.). *Learner English on computer*. New York: Longman, 1998. p. 03-18.
- 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. Should you be persuaded: two samples of data-driven learning materials. *English language research journal*, Birmingham, v. 4, p. 1-16, 1991.



NEVES, A. J. *Cibercultura e literatura identidade e autoria em produções culturais participatórias e na literatura de fã (fanfiction)*. Jundiaí: Paco Editorial, 2014.

PARROTT, M. *Grammar for English language teachers*. 2nd. Cambridge: Cambridge University Press, 2000.

SWAN, M. *Practical English usage*. 3rd. Oxford: Oxford University Press, 2005.

ANEXO – A tarefa proposta

TASK

You are an English teacher who takes **real language data information** into account when preparing your classes and when giving your students feedback. Below, you will find a list of sentences produced by some of your students. Using the **Corpus of Contemporary American English (COCA)** as a reference to your language analysis, prepare your feedback to your students in terms of which **prepositions** were used appropriately and which ones were not. Justify your explanations with real and authentic data and/or evidence taken from the corpus.

1) so they agreed to punish Romeo for such a tragedy. He was sent **away for a jail**, where he died. Rosaline had the baby and Romeo's Family supported

Feedback:

2) the hate between their families made their love impossible. **Desperate for not being** able to live with his great love

Feedback:

3) The girl was Juliet, she was a very young girl and **loved to walk alone** in the city to see new places.

Feedback:

4) she loved a man from the Montesquieu family. She discussed with the **family for him and** left the house. When she found him, he was flirting.

Feedback:

5) somebody getting closer and finally could see who was the **person. For their surprise** it was Rosaline, Romeo's ex girlfriend who was crying.

Feedback:

6) Above of all, Juliette decided to invite Romeo for a date before make a judgement about him
Feedback:
7) The Friar ran immediately of he church and told what he had saw for the both families, so they agreed to punish Romeo for such a tragedy.
Feedback:
8) he crossed the street and kept following her until he reached her and talked to her. The girl was Juliet, she was a very young girl and loved to
Feedback:
9) and started asking Romeo what was it about. After Rosaline tell the story for them- that she was pregnant of Romeo and they planned to stole
Feedback:
10) Romeu and Juliano In my light bed, I close my eyes and I thinks in one person, only a person, Romeo. My name is Juliano and this is the
Feedback:
11) And for each violent end in the stage, are violent delights in the backstage, theses are the dreams of the sleep of death.
Feedback:
12) One is Count Paris, the other, Romeu. And for each violent end in the stage, are violent delights in the backstage,
Feedback:

13) We started to have lunch together and talk to each other on parties. It all seemed perfect until the day I had the chance

Feedback:

14) Romeo accepted the invitation and was so excited. However at the day of the date, Juliette had to do an extra homework

Feedback:

15) immediately, I drink and than, I died. I don't know what happened with Romeo, I hope he has one man most beautiful and perfect of me

Feedback:

Concordance line	Research Method	Appropriateness	Feedback
1)			
2)			
3)			
4)			
5)			
6)			
7)			
8)			
9)			
10)			
11)			
12)			
13)			
14)			
15)			