

Corpus computadorizado e concordância no ensino de língua

Ednei Nunes de Oliveira*

Resumo: Para o presente artigo, fundamentado na linguística de *corpus* e no ensino de línguas mediado pelas tecnologias de informação e comunicação (TIC), teve-se o objetivo de apresentar os conceitos de *concordância* e *concordanciador* e algumas formas de utilizar a concordância no processo de ensino e aprendizagem de língua. Como atividades com concordâncias e concordanciadores vêm sendo desenvolvidas tanto por docentes e alunos em práticas educativas como por linguistas em suas pesquisas, demonstramos, ao final do trabalho, alguns exemplos de atividades e/ou projetos que podem ser desenvolvidos com estudantes de língua. Entretanto, não há a pretensão de passar modelos prontos para serem seguidos, mas apenas indicar como os trabalhos iniciais podem ser desenvolvidos.

Palavras-chave: Ensino de Língua; Linguística de *Corpus*; Tecnologias de Informação e Comunicação.

Abstract. This paper is based on Corpus Linguistics and in the teaching of languages mediated by Information and Communication Technologies (ICT). The aim was to present concepts of agreement and concordancers and forms of using agreement in the language teaching/learning process. Activities with agreements and concordancers have been developed by teachers and students as well as by linguists. Examples of activities/projects that can be developed with language students are presented. There is no intention of presenting models to be followed, but to indicate how the initial work can be developed.

Keywords. Language Teaching; Corpus Linguistics; Information and Communication Technologies.

Introdução

A criação de uma concordância completa de palavras da Bíblia Latina, completada em 1230, é considerada o primeiro trabalho com concordâncias. Ela foi produzida por um grupo de 500 monges dominicanos liderados por Hugo de San Charo.

Para a linguística de *corpus*, uma concordância é a exibição de uma ou diversas ocorrências de determinada palavra, partes de palavras ou mesmo uma expressão, geralmente apresentada por linhas, com a palavra-chave no meio e o respectivo contexto de cada lado. Estas exibições são chamadas de *Kwic* (*keyword in context*).

Cada palavra-chave tem o contexto onde ocorre dos seus lados esquerdo e direito. As concordâncias geradas por computador são semelhantes aos

* Professor doutor. Universidade Federal da Grande Dourados – UFGD. E-mail: edneioliveira@gmail.com

índices bíblicos que facilitam a compreensão de certos temas, em virtude de facilitar a localização e o acesso a outras representações de uma informação. Pode-se gerar uma concordância de uma palavra-chave dada ou de todas as palavras que têm um determinado prefixo, sufixo, ou que contenham algumas letras (comuns em determinados alfabetos, como o ñ no Espanhol), sinais particulares ou expressões.

Os programas que geram uma concordância são chamados de *concordanciadores*. Segundo Berber Sardinha (2004), trata-se de um *software* com ferramentas básicas para o linguista que trabalha com *corpus*. Como os *corpora*, na sua maioria, são muito grandes, seria uma atividade infrutífera sua exploração sem a ajuda de um computador. Os concordanciadores varrem a base de dados de *corpus* à procura de palavras que sejam do interesse do usuário. Geralmente, grande parte dos programas também oferece a possibilidade da busca de expressões ou combinações de palavras e, dependendo do grau de sofisticação, também pode fornecer listas de palavras com informações sobre a frequência de cada uma delas no *corpus*.

Para Oliveira (2005), a utilização de concordâncias no processo de ensino e aprendizagem justifica-se porque aprendizes lembram melhor do conhecimento que formularam que de formulações que lhe foram impostas.

1. Trabalhando com concordância em sala de aula

Um dos pioneiros que utilizaram a concordância na pesquisa e no ensino de língua foi Tim Johns, que, desde os anos 1970, tem feito uso de concordâncias na University of Birmingham e tem escrito programas para a geração de exercícios para o ensino de língua. Ele é um dos autores do concordanciador *Sinclair ZX81*, um dos primeiros programas populares a rodar em microcomputadores, e também foi o autor do primeiro concordanciador comercial, o *MicroConcord*, produzido pela Oxford University Press.

Especificamente em relação ao ensino de gramática da língua inglesa, Johns elaborou a *aprendizagem movida a dados* (*data-driven learning – DDL*), que, segundo seu autor (JOHNS, 1991), é uma forma efetiva de aprendizagem e de pesquisa linguística que oferece um recurso inigualável para a estimulação de estratégias de aprendizagem indutivas. Por meio da DDL, Johns (1988) procurou utilizar concordâncias no ensino de línguas porque elas propiciavam

aos estudantes o contato com exemplos autênticos de uso da língua no processo de aprendizagem. Essa forma de trabalho propicia a construção da competência de estudantes, dando-lhes acesso a fatos de desempenho linguístico. Com a DDL, o aspecto mais importante do processo de aprendizagem é a descoberta de conceitos gramaticais pelo estudante, feita a partir de evidências de uso autêntico de uma língua.

Portanto, as concordâncias passaram a ser adotadas como uma técnica para a exploração de *corpus* com estudantes, principalmente porque elas proporcionam a oportunidade para que os aprendizes realizem atividades de descoberta. Conforme Johns (1994), as concordâncias têm um impacto positivo sobre a autoaprendizagem de uma segunda língua, tanto para os estudantes como professores. Para o autor, os professores não devem ser vistos como a maior fonte de conhecimento, por isso eles podem recorrer a *corpus* para obterem diversas respostas.

professor tinha isso claro), mas porque ele podia, através dessa prática, entender melhor o que é Dourados, MS. A coleta de dados foi realizada através de levantamento bibliográfico, observações individuais e coletivos realizados entre eles e através de parcerias e na prática adquirida nos tecnologias como suporte, foi possível verificar, através da observação, da análise de documentos e uma concordância através de um software. Através da wordlist, o professor poderá orientar as escolas de conectar escolas às redes telemáticas através de modernos equipamentos e alfabetização através da simulação de diferentes interações através da construção de outros exemplos de u

Figura 1. Exemplo de concordância

Essa oportunidade para usar exemplos pertinentes, autênticos e interessantes, em vez de exemplos provenientes da intuição de autores de livros didáticos ou de gramáticas tradicionais, é uma das principais vantagens da utilização de concordâncias no ensino, pois, como afirma Rocha (2001), as críticas à gramática tradicional são reações aos problemas gerados pelo ensino com base em uma especificação predeterminada de sentenças gramaticalmente corretas. Nesse sentido, para Stubbs, as

[...] concordances can identify syntactic features of texts which can then be semantically interpreted. Independent of intuitive judgment (within the limitations noted), a concordancer identifies every example in the data, and helps to ensure that analysts do not merely pick evidence to

fit their preconceptions. This also helps to present quantitative evidence in ways which can be checked by readers. (STUBBS, 1994, p. 218)¹

Utilizando concordâncias em sala de aula, os estudantes podem ser mais ativos na aprendizagem de seu próprio vocabulário e, dependendo de seu nível, podem até questionar algumas das regras, baseando-se em suas observações dos padrões da língua. Em geral, os aprendizes podem ser convidados a descobrir vocábulos e significados novos, observar diferenças estilísticas, relacionar palavras sintaticamente e, entre outras atividades, analisar definições de dicionário.

O professor pode usar um concordanciador para achar exemplos de uso para demonstrar aos alunos características de vocabulário, colocações típicas, um ponto da gramática, ou a estrutura de um texto.

No começo das atividades, os estudantes têm muita curiosidade de experimentar um concordanciador. É possível que o professor tenha muito trabalho para administrar esse tipo de tarefa, pois, segundo Ball (2001), os estudantes acham divertido trabalhar com *corpora* computadorizados quando a tarefa é orientada para buscar diferenças culturais, mas eles mostram menos habilidade em explorações de dados gramaticais ou sintáticos. É nesse tipo de atividade que necessitam da mediação de seus professores.

Os usuários podem empregar diferentes categorizações daquelas já empregadas pelos linguistas de *corpus*, pois nem sempre certas categorias que eles desejam trabalhar podem estar bem representadas numericamente e nos termos de sua variação no *corpus* (ASTON, 2004). Isso significa que, caso quisesse estudar alguns aspectos de determinados gêneros textuais, muitas vezes um usuário teria melhor resultado se compilasse seu próprio *corpus*. No entanto, para isso o professor-pesquisador deve ter relativo conhecimento de teorias linguísticas e, especificamente, de *corpus*.

Aston (1997) afirma que um projeto que envolve a exploração de material para a compilação de *corpus* pode permitir ao usuário encontrar tipos específicos de artigos que podem ser usados para uma variedade de metas co-

¹ “[...] concordâncias podem identificar características sintáticas de textos que podem ser, então, semanticamente interpretadas. Independentemente do julgamento intuitivo (dentro das limitações notadas), um concordanciador identifica todos os exemplos nos dados, e ajuda a assegurar que os analistas não escolham evidências somente para ajustar as suas pré-concepções. Isso também ajuda a apresentar evidência quantitativa de tal forma que pode ser conferido por leitores.” - tradução nossa.

municativas, tais como descobrir sobre um país e seus problemas, investigar argumentos e atitudes relativas à poluição, proliferação de material nuclear, informações sobre outras nações etc. Para ele, é mais fácil envolver os alunos no processo de seleção de textos, na decisão sobre os critérios de busca e a produção de estratégias para a identificação de artigos pertinentes aos seus propósitos e interesses. Esse envolvimento pode aumentar a motivação dos alunos para participar do projeto e de discussões que poderão ocorrer durante o processo.

2. Formas de obter *corpora*

Se o professor não deseja trabalhar com os *corpora* que estão disponíveis na internet (alguns gratuitos), pode obtê-los por meio da digitalização de textos e das ações de cópia e colagem de textos de CD-ROM, de páginas de rede local ou da internet. Os textos impressos (livros, revistas, jornais etc.) também podem ser digitalizados, contudo a digitalização pode produzir alguns problemas em virtude da leitura óptica e do reconhecimento de caracteres, tais como a substituição de um caractere não reconhecido por outro. Embora esse problema possa causar algum incômodo aos usuários, não chega a ser tão grave a ponto de comprometer os benefícios que a digitalização de textos oferece (OLIVEIRA, 2005).

Há duas formas práticas de obter textos de língua portuguesa para construção de *corpus* utilizando a internet.

A primeira é fazendo busca de textos eletrônicos em sítios especializados no serviço de disponibilização de obras literárias para estudo.

A outra é navegando em páginas de interesse específico, realizando-se busca em revistas eletrônicas especializadas e em sítios com matérias jornalísticas sobre temas esportivos diversificados e sobre acontecimentos regionais, nacionais e mundiais da atualidade.

Outro produto que está à disposição de interessados em trabalhar com concordâncias é o *corpus on-line*. Uma das características de *corpus* e de concordâncias *on-line* é o seu tamanho: eles são de fato muito grandes e podem ser usados por estudantes para criar seus próprios textos para tê-los como referência.

Caso o professor queira consultar um *corpus on-line* em português, é possível buscar a concordância de palavras na página do Processamento Computacional do Português (<<http://www.linguateca.pt/>>). Nesse sentido, há diversas formas de construir ou ter acesso a um *corpus*, todavia o mais importante (considerando a concordância em sala de aula) não é o tamanho ou as características do *corpus*, mas a metodologia que norteará o trabalho do professor e dos alunos.

3. Algumas atividades e pesquisas com *corpus*

A seguir, citaremos alguns exemplos de atividades e/ou projetos que podem ser desenvolvidos com estudantes de língua. Sabemos que algumas dessas atividades também podem ser realizadas sem a utilização de *corpus*. Não há, no entanto, o interesse de discutirmos esse ponto nem passar alguns modelos para serem seguidos, mas apenas indicar como, geralmente, os trabalhos iniciais são desenvolvidos. Algumas dessas sugestões são feitas a partir das reflexões de Ball (2001).

3.1 Explorando *corpus* com textos dos alunos

Uma das atividades mais interessantes que podem ser realizadas é utilizar as próprias produções textuais de estudantes para a compilação de *corpus*, com o objetivo de chamar-lhes a atenção sobre as diferentes formas de produzir textos coesos e coerentes. Após a compilação do *corpus*, o professor pode orientá-los a fazerem uma lista de palavras (*wordlist*) e uma concordância por meio de um *software*.

O professor pode orientar seus alunos a observarem, por meio de uma *wordlist*, a frequência das palavras que utilizam. Na sequência, ele pode dividir a turma em diversos grupos e solicitar-lhes que criem concordâncias das palavras mais recorrentes no *corpus*. O docente pode orientá-los a explorar os diferentes contextos em que essas palavras ocorrem e solicitar-lhes que tentem substituí-las por algum sinônimo, sem, no entanto, alterarem o significado do enunciado entendido em cada contexto, chamando a atenção para o fato de que estão trabalhando em um *corpus* de textos produzidos por eles mesmos. Caso os alunos não tenham condição de encontrar sinônimos

para fazer a substituição de algumas das ocorrências, o professor pode lhes facultar a utilização de dicionários, enciclopédias ou, em último caso, sugerir-lhes algumas alterações.

Ainda, como exercícios específicos de gramática normativa, o professor pode solicitar que os alunos observem a ocorrência, por exemplo, de conjunções coordenadas aditivas, entre outras, e verificar se há um equilíbrio de utilização ou se são utilizadas em mais de uma forma no texto.

Outro ponto interessante para os alunos observarem em seu *corpus* é como eles lidam com a ideia de causalidade. Após fazerem uma concordância da palavra *porque*, por exemplo, o professor pode sugerir-lhes que tentem pôr em seu lugar *por causa de*, *devido a*, *a razão por que*, *como*, *em virtude de*, entre outros. Nesse tipo específico de atividade, o mais importante é levar os alunos a refletirem sobre as diferentes formas de expressar, em seus textos, causa, adição, contraste etc.

Levando-se em consideração a quantidade de textos inseridos no *corpus*, o professor pode solicitar aos alunos que observem, na *wordlist*, a ocorrência, por exemplo, de conectivos que indiquem o início de uma conclusão. É possível que todos os alunos tenham feito suas conclusões; no entanto, há a probabilidade de que alguns deles tenham esquecido de apontá-las por meio de um conectivo adequado. É provável que na composição ocorram palavras ou expressões tais como *em resumo*, *resumindo*, *em uma palavra*, *ao todo*, *em conclusão*, *assim*, *por conseguinte*, *como resultado*, *então*. Entretanto, o professor deve considerar que os indicadores de conclusão podem variar muito, por diversos fatores. Essa variação pode gerar a necessidade de conferir quais os usos mais frequentes em um *corpus* maior que, talvez, tenha o predomínio de textos com gêneros discursivos diferentes do *corpus* produzido com textos dos alunos. Isso possibilitaria, então, a exploração de outros *corpora* para que os alunos estabeleçam diferenças e semelhanças na variação linguística entre textos predominantes em determinados *corpora*.

3.2 Explorando diferentes *corpora*

Caso o professor tenha o interesse de trabalhar com outros tipos de *corpora*, há outras formas bastante atrativas para os alunos. Uma delas é com os diferentes estilos de linguagem. É possível, por exemplo, que o professor compile, ele mesmo, ou oriente seus alunos a compilarem dois *corpora*: um

utilizando textos de autores do século XVII ou XVIII e outro a partir de vários assuntos de revistas semanais, publicadas em datas próximas à da realização da pesquisa. Um dos objetivos desse tipo de atividade pode ser, por exemplo, apresentar aos estudantes a ideia de que um *corpus* pode refletir a sociedade de seu tempo, como é o caso dos *corpora* histórico, sincrônico e diacrônico.

Os alunos poderão observar a mudança de estilo linguístico e literário com o passar do tempo. Nesse caso, eles devem ser conscientizados de que o resultado de uma busca de palavra é altamente dependente do que está no *corpus* em primeiro lugar, pois determinadas palavras sofrem mudanças ou deixam de ser utilizadas com o tempo.

Outra atividade interessante é a comparação de palavras que ocorrem entre dois *corpora*, para verificar se elas conservam os mesmos significados. A comparação entre diferentes *corpora* pode levar os estudantes a várias conclusões não previstas pelo professor; contudo, sua postura deve ser de estar aberto às opiniões colocadas pelos alunos e não limitá-los em suas descobertas.

Além da verificação do significado de diferentes palavras, outro objetivo a ser atingido envolvendo dois *corpora* é a verificação de diferentes culturas regionais. A compilação de *corpus* com textos que revelam e descrevem a cultura dos povos de regiões distintas e distantes do mesmo país pode, por exemplo, levar os alunos a perceberem o fato de que as escritas de uma cultura refletem hábitos das pessoas daquela cultura. O professor pode solicitar, por exemplo, que os alunos criem concordâncias, a partir dos dois *corpora*, de palavras que provavelmente não serão encontradas ou não terão a mesma frequência em ambos. Os alunos poderão ser questionados sobre as diferentes conclusões a que eles chegaram sobre os interesses das diferentes regiões em relação a alimentação, vestuário, danças típicas, trabalho, religião etc.

3.3 Um exemplo de atividade utilizando o WordSmith

O professor pode colocar seus alunos diante do desafio de observarem diferenças de significado e mudanças estilísticas em uso sugerindo, por exemplo, que os alunos criem concordâncias da palavra *absorver* a partir de dois *corpora*. Primeiramente, haveria a necessidade da compilação desses *corpora*. Essa compilação poderia ser feita a partir de textos copiados da internet, um com matérias jornalísticas e/ou textos atuais e o outro com obras dos clássicos da literatura brasileira.

Uma ferramenta que pode ser utilizada para a busca dos textos atuais é o *Google*, bastante utilizada por linguistas e usuários comuns da internet. Cabe lembrar que, em virtude da contínua e intensa atividade de inclusão e exclusão de textos na internet, se a coleta de dados linguísticos for repetida poucas horas depois, poder-se-ia apresentar resultados diferentes.

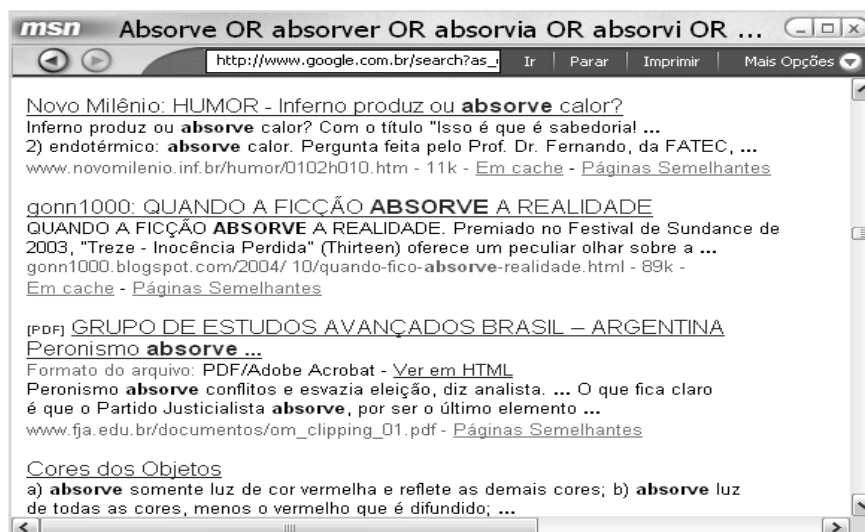


Figura 2. Exemplo de busca realizada pela ferramenta do Google

Os passos iniciais para a compilação desse *corpus*, utilizando o *Google*, podem ser os seguintes:

- abrir o *browser* no sítio <http://www.google.com.br>;
- clicar na opção “pesquisa avançada”;
- digitar o verbo *absorver* em várias flexões;
- selecionar a opção de apresentação da quantidade de resultados, solicitar “100 resultados” e clicar no botão “pesquisa *Google*”;
- selecionar e copiar uma quantidade relativa de exemplos de uso, aleatoriamente, e posteriormente colar no editor de textos;
- salvar o arquivo como tipo “somente texto”.

O *corpus* dos clássicos da literatura brasileira pode ser compilado a partir da cópia de obras completas disponibilizadas, por exemplo, na página <<http://www.casadobruzo.com.br>>. Os passos para a compilação desse *corpus* podem ser os seguintes:

- abrir o *browser* no site <<http://www.casadobruzo.com.br>>;
- clicar no nome de cada obra a ser copiada e salvá-las no computador;
- copiar todas as obras para um único arquivo no editor de textos;
- salvar o arquivo como tipo “somente texto”.

Após a compilação dos *corpora*, os alunos poderão buscar o significado de *absorver* em dicionários e/ou enciclopédias, conforme abaixo.

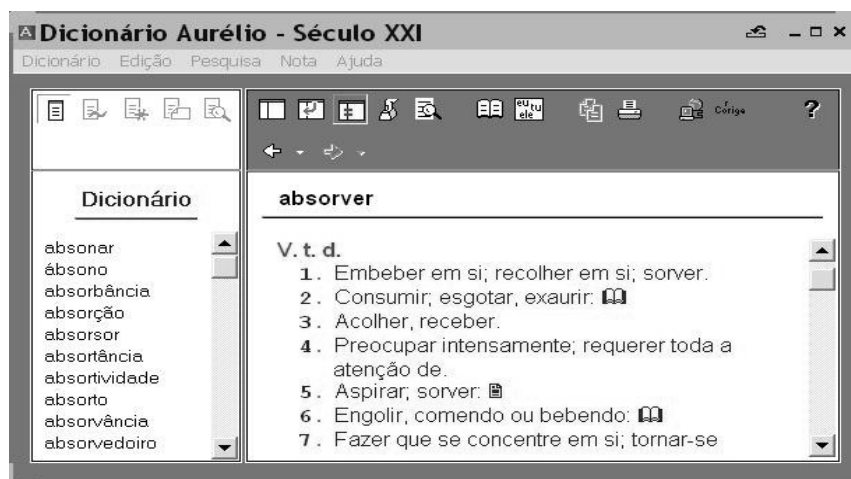


Figura 3. Exemplo de dicionário eletrônico

Ferreira traz as seguintes informações a respeito da palavra *absorver*:

[Do lat. absorbere.] V. t. d. 1. Embeber em si; recolher em si; sorver. 2. Consumir; esgotar, exaurir. 3. Acolher, receber. 4. Preocupar intensamente; requerer toda a atenção de. 5. Aspirar; sorver. 6. Engolir, comendo ou bebendo. 7. Fazer que se concentre em si; tornar-se objetivo exclusivo de. 8. Arrebatado, enlevar. 9. Recolher, apreender, assimilar: 2. 10. Aprender, assimilar: 2. 11. Passar por cima; superar: 2. 12. Fís.-Quím. Realizar a absorção de (uma substância). V. t. d. e. i. 13. Concentrar, aplicar, ocupar inteiramente: 2. V. p. 14. Concentrar-se, aplicar-se

Corpus computadorizado e concordância no ensino de língua

detidamente em. 15. Ficar absorto; concentrar-se nos seus pensamentos.

16. Arrebatarse, enlevar-se, entusiasmar-se. (FERREIRA, 1999).

Levantado os significados, os alunos poderão gerar concordâncias da palavra utilizando o *WordSmith* para explorar os dois *corpora*.

Após abrir a janela inicial do *WordSmith*, o aluno deverá ativar a ferramenta *Concord*, conforme abaixo.

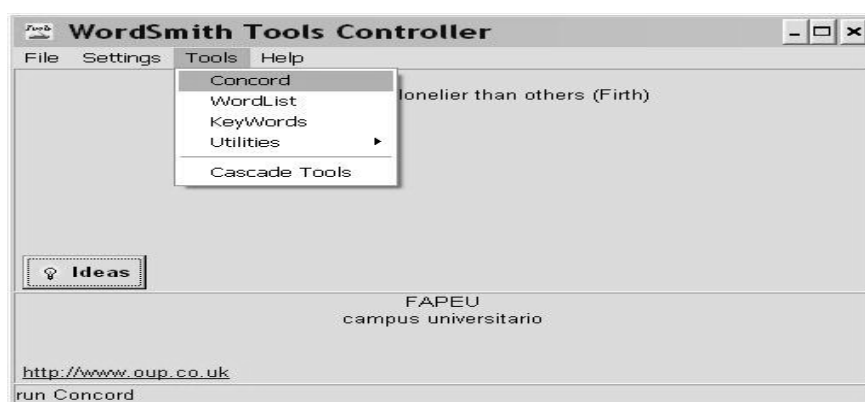


Figura 4 – Tela inicial do WordSmith

A janela a ser aberta é a que segue.

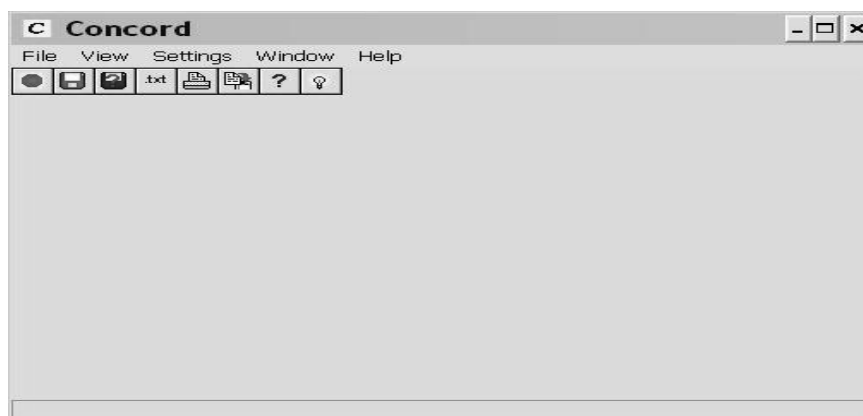


Figura 5 – Tela inicial do Concord

Ao clicar em “file” e, posteriormente, em “start”, será aberta a ferramenta “getting started”, que possibilita a definição dos parâmetros para a geração da concordância a partir de um determinado *corpus*, conforme abaixo.



Figura 6 – Ferramenta “getting started” do Concord

A próxima ação deverá ser clicar em “choose texts now” para escolher o *corpus* que será analisado. A janela para a escolha do *corpus* é a que segue.

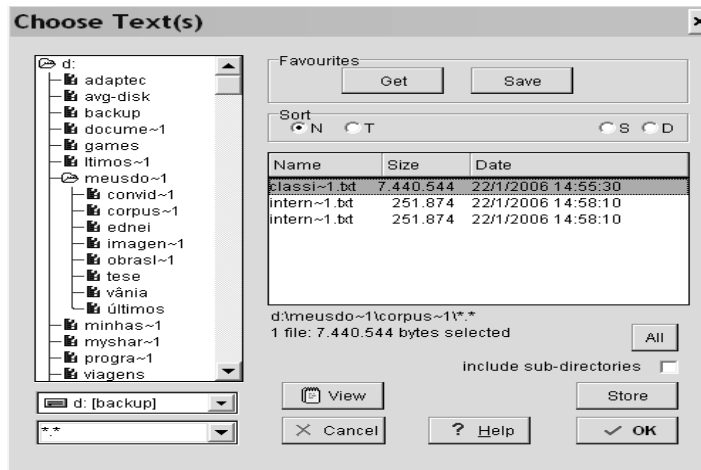


Figura 7. Ferramenta “choose texts now” do Concord

Nesta caixa de ferramenta, o aluno pode selecionar um ou mais textos da pasta onde estão arquivados os *corpora* que pretende analisar. Feita a devida seleção do *corpus*, basta clicar no botão “ok” e a ferramenta retornará para a janela “getting started”, que ainda está em uso.

Retornando a essa última janela, o novo passo será escolher a palavra-chave que comporá a geração da concordância. Isso deverá ser feito clicando-se no botão “specify search-word”, e a janela para a sua inserção será “concordance settings”, conforme abaixo.

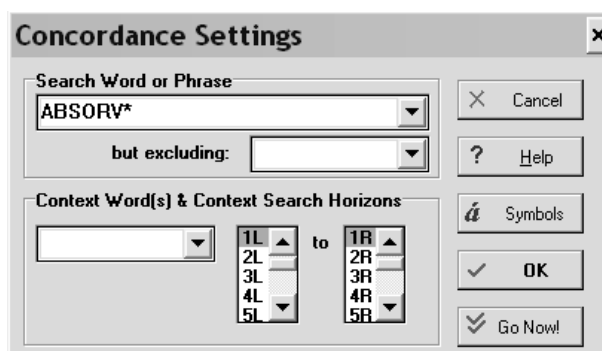


Figura 8 – Ferramenta “concordance settings” do Concord

É interessante que a palavra a ser digitada (na caixa “search word or phrase”) o seja parcialmente, colocando-se um asterisco na parte final, para que o concordanciador gere uma concordância com todas as possibilidades de variação dessa palavra. Por exemplo, na busca das ocorrências da palavra *absorver*, foi digitado “ABSORV*”. Após o preenchimento do campo destinado à palavra, basta clicar no botão “ok” que o concordanciador retornará para janela “getting started”, restando a alternativa de clicar no botão “start concordance”. Logo após essa ação, a concordância será gerada e o concordanciador apresentará a janela abaixo.

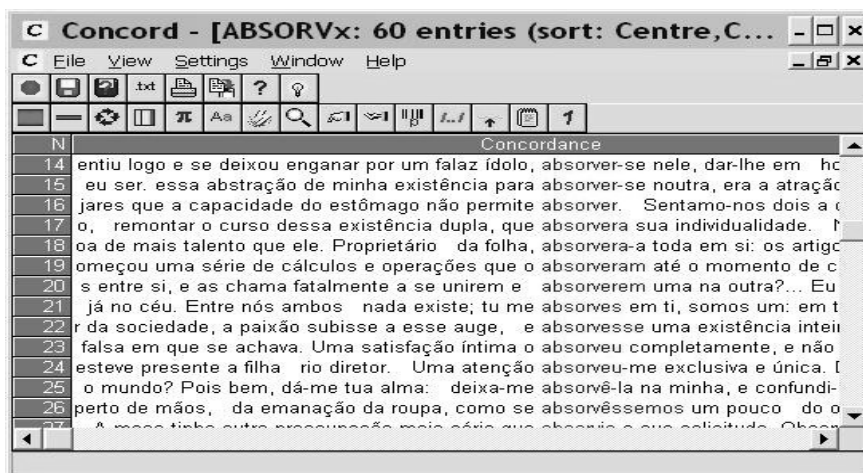


Figura 9. Resultado de concordância gerada pelo Concord

A análise (meta)linguística propriamente dita utilizando concordanciadores só passaria a ocorrer a partir da análise e da reflexão a respeito dos dados gerados até esse ponto pelas ferramentas do *WordSmith*.

Os alunos poderão manipular, observar e comparar os dados dos dois *corpora* entre si e verificar com quais significados a palavra é mais utilizada em cada um deles; se apenas literal ou se também com sentidos derivados, levando-se em conta o contexto em que a palavra ocorre no *corpus* e de sua comparação com as definições do dicionário.

Considerando os significados apresentados por Ferreira (1999) para a palavra estudada, teríamos os seguintes dados levantados nas concordâncias geradas a partir dos dois *corpora* compilados para esse exemplo de atividade:

Quantidade de ocorrências da palavra pesquisada ²

(continua)

² A numeração é correspondente à apresentada por Ferreira (1999) no *Dicionário Aurélio: Século XXI*.

Significado dado pelo dicionário	Quantidade de ocorrências no <i>corpus</i> literário	Quantidade de ocorrências no <i>corpus</i> atual
1	0	01
2	04	06
3	0	10
4	13	0
5	0	0
6	04	01
7	0	0
8	02	0
9	01	05
10	02	01
11	02	02
12	0	21
13	02	0
14	11	0
15	11	0
16	08	0

(conclusão)

Estes dados poderiam levar alunos a perceberem que os significados 4, 8, 13, 14, 15 e 16 da palavra *absorver*, utilizados há muitas décadas por escritores da clássica literatura brasileira, não chegaram a ser encontrados no *corpus* com textos atuais. Da mesma forma, inversamente, observariam que os significados 1, 3, e 12 não ocorreram no primeiro *corpus*. Os alunos também poderiam notar que os significados mais utilizados no passado se referem aos dos números 14 a 16 e, na atualidade, aos de números 3 e 12.

Munido desses dados, o professor pode orientar seus alunos a uma série de reflexões a respeito da escolha lexical que o produtor de textos pode fazer de determinadas palavras, dependendo do sentido que pretende dar às ideias desenvolvidas, e que geralmente essas escolhas refletem um estilo de determinadas épocas e escolas literárias.

Considerações finais

Apresentamos algumas atividades que podem ser desenvolvidas pelos professores utilizando *corpus* no ensino. Muitas outras podem ser feitas; entretanto, chamamos a atenção para o fato de que o professor não pode reduzir o trabalho com *corpus* a atividades para confirmação de regras gramaticais. O trabalho que envolve diferentes usos que o falante faz da língua pode ser mais produtivo e interessante aos alunos, e é possível que eles demonstrem maior interesse por esse tipo de atividade do que pelos exemplos dados neste trabalho.

Entretanto, a utilização de um *corpus* deve ir além da confirmação da intuição linguística do falante nativo ou do aprendiz de uma nova língua, bem como ir além da simples observação estatística (por meio de diferentes ferramentas de concordanciadores) das ocorrências de determinado dado, em diferentes posições em uma concordância. O “aprendiz de pesquisador” deverá refletir sobre possíveis razões de ocorrências em determinados contextos para, depois disso, formular hipóteses. Devem-se buscar novos significados e formas de uso de palavras, já que a competência linguística e comunicativa do falante pode ser construída por meio não apenas do uso mas também do processo de investigação metalinguística.

Mesmo assim, é de esperar que nem todos os aprendizes tenham habilidade para lidar com todo tipo de ocorrência dos termos que estejam pesquisando. É nesse ponto que a mediação do professor é mais importante para a aprendizagem. No entanto, ele deve tomar o devido cuidado para não conduzir o aluno a regras pré-formuladas pelos gramáticos ou por ele mesmo, mas sim a respostas surgidas a partir dos dados no *corpus*. Um experiente professor de língua deve sempre considerar o nível linguístico de seus alunos antes de sugerir-lhes determinadas atividades. Dessa forma, ele poderá fazer uma utilização mais adequada de *corpus* e de concordanciadores como suporte ao estudo de língua.

Referências

ASTON, G. Enriching the Learning Environment: corpora in ELT. In: WICHMANN, A. et al. (Orgs.). **Teaching and language corpora**. Nova York: Longman, 1997, p. 51-63.

_____. **Text categories and corpus users**: a response to David Lee. Disponível em: <<http://lit.msu.edu/vol5num3/aston/default.html>>. Acesso em: 5 mar. 2004.

BALL, C. N. **Tutorial:** concordances and corpora. Disponível em: <<http://www.georgetown.edu/cball/corpora/tutorial.html>>. Acesso em: 12 dez. 2001.

BERBER SARDINHA, T. **Linguística de corpus**. Barueri: Manole, 2004.

FERREIRA, A. B. H. **Dicionário Aurélio eletrônico:** século XXI. Versão 3. Rio de Janeiro: Nova Fronteira, 1999, 1 CD-ROM.

JOHNS, T. From printout to handout: grammar and vocabulary teaching in the context of Data-driven learning. In: ODLIN, T. (Org.). **Perspectives on pedagogical grammar**. Cambridge: Cambridge University Press, 1994, p. 293-313.

_____. Should you be persuaded: two examples of data-driven learning materials. **ELR Journal**, Birmingham, v. 4, n.1, p. 1-16, 1991,

_____. Whence and whither classroom concordancing? In: BONGAERTS, T. et al. (Eds.). **Computer applications in language learning**. Dordrech: Foris, 1988, p. 9-27.

OLIVEIRA, E. N. de. **A linguística de corpus no ensino-aprendizagem de línguas:** um estudo de caso das práticas de professores da educação básica de Dourados (MS). 235 f. Tese (Doutorado em Linguística). Universidade Federal de Santa Catarina, Florianópolis, 2005.

ROCHA, M. A. E. O uso de *corpora* computadorizados no ensino de língua portuguesa: metodologia e avaliação. In: CABRAL, L. G. et al. (Orgs.). **Linguística e ensino:** novas tecnologias. Blumenau: Nova Letra, 2001, p.137-55.

STUBBS, M. Grammar, text, and ideology: computer-assisted methods in the linguistics of representation. **Applied Linguistics**, v. 15, n. 2, 1994, p. 201-23.

Recebido para publicação em 02 julho 2009

Aceito para publicação em 13 maio 2010