

Colocações especializadas na área médica extraídas a partir do *corpus House M.D.*

Jean Michel Pimentel Rocha e Adriane Orenha-Ottaiano

Submetido em 30 de abril de 2012.

Aceito para publicação em 16 de junho de 2012.

Publicado em 30 de junho de 2012.

Cadernos do IL, Porto Alegre, n.º 44, junho de 2012. p. 295-318

POLÍTICA DE DIREITO AUTORAL

Autores que publicam nesta revista concordam com os seguintes termos:

- (a) Os autores mantêm os direitos autorais e concedem à revista o direito de primeira publicação, com o trabalho simultaneamente licenciado sob a [Creative Commons Attribution License](#), permitindo o compartilhamento do trabalho com reconhecimento da autoria do trabalho e publicação inicial nesta revista.
- (b) Os autores têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada nesta revista (ex.: publicar em repositório institucional ou como capítulo de livro), com reconhecimento de autoria e publicação inicial nesta revista.
- (c) Os autores têm permissão e são estimulados a publicar e distribuir seu trabalho online (ex.: em repositórios institucionais ou na sua página pessoal) a qualquer ponto antes ou durante o processo editorial, já que isso pode gerar alterações produtivas, bem como aumentar o impacto e a citação do trabalho publicado.
- (d) Os autores estão conscientes de que a revista não se responsabiliza pela solicitação ou pelo pagamento de direitos autorais referentes às imagens incorporadas ao artigo. A obtenção de autorização para a publicação de imagens, de autoria do próprio autor do artigo ou de terceiros, é de responsabilidade do autor. Por esta razão, para todos os artigos que contenham imagens, o autor deve ter uma autorização do uso da imagem, sem qualquer ônus financeiro para os Cadernos do IL.

POLÍTICA DE ACESSO LIVRE

Esta revista oferece acesso livre imediato ao seu conteúdo, seguindo o princípio de que disponibilizar gratuitamente o conhecimento científico ao público proporciona sua democratização.

<http://seer.ufrgs.br/cadernosdoil/index>

Sábado, 30 de junho de 2012

23:59:59

COLOCAÇÕES ESPECIALIZADAS NA ÁREA MÉDICA EXTRAÍDAS A PARTIR DO *CORPUS HOUSE M.D.*

Jean Michel Pimentel Rocha*

Adriane Orenha-Ottaiano**

RESUMO: À luz da *Linguística de Corpus* e da *Fraseologia*, discutiremos aspectos teórico-metodológicos que nos possibilitaram o levantamento de colocações e de colocações especializadas na área médica, a partir do corpus formado pelas transcrições dos episódios da série *House M.D.* Apresentamos também as colocações especializadas formadas com base nos nódulos 'blood', 'brain', 'tumor' e 'heart', a fim de exemplificar como as colocações serão organizadas em um glossário bilíngue de colocações especializadas na área médica, que se encontra em fase de compilação.

PALAVRAS-CHAVE: *Linguística de Corpus*; *Fraseologia*; *colocações especializadas*; *glossário bilíngue*.

1. INTRODUÇÃO

Neste artigo, trataremos dos aspectos teórico-metodológicos necessários para a extração de colocações e de colocações especializadas, a partir de um *corpus* paralelo bilíngue formado pelas transcrições das sete temporadas da série *House M.D.* Acreditamos que, ao desenvolvermos um estudo focado nas colocações e baseado em *corpora*, passamos a ver a língua sob um novo aspecto, pois, ao contrário da abordagem gramatical tradicional na maioria das vezes focada em uma abordagem intuitiva para explicar o funcionamento das estruturas da língua, atentamos para a companhia mantida pelas palavras, ou seja, pelas combinações lexicais e pelas frequências e recorrências destas na língua em uso. Nesse sentido, ao enfatizarmos e adotarmos uma perspectiva que privilegie o falar natural no estudo das combinações lexicais, em especial, das colocações, esperamos contribuir para uma formação mais sólida e crítica do estudante de Letras e futuro profissional de ensino de línguas, para que este profissional possa também transmitir esse conhecimento e competência fraseológica para seus próprios alunos.

Antes de detalharmos os passos metodológicos para o levantamento das colocações, faz-se necessário discorrer sobre o arcabouço teórico que promoveu o embasamento adequado para o desenvolvimento deste trabalho.

* Aluno de Graduação do Curso de Licenciatura em Letras (Iniciação Científica), do Instituto de Biociências, Letras e Ciências Exatas, da Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP): jeanpimentel_sp@yahoo.com.br

** Docente do Instituto de Biociências, Letras e Ciências Exatas, da Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP) - Departamento de Letras Modernas (área de Língua Inglesa) e do PPG em Estudos Linguísticos: adriane@ibicle.unesp.br

2. FUNDAMENTAÇÃO TEÓRICA

Em primeiro lugar, trataremos da Linguística de *Corpus* e de sua relevância para os estudos linguísticos baseados em *corpora*. A seguir, discutiremos a Fraseologia, área da Linguística voltada para as combinações lexicais recorrentes. Por fim, focaremos nas colocações e colocações especializadas, nosso principal objetivo neste artigo.

2.1. Linguística de Corpus

De acordo com a definição de Berber Sardinha (2004):

A Linguística de Corpus ocupa-se da coleta e da exploração de corpora, ou conjuntos de dados linguísticos textuais coletados criteriosamente, com propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador. (BERBER SARDINHA, 2004, p. 03).

Tal definição é concisa e ressalta a importância do uso de *corpora* eletrônicos em pesquisas linguísticas, o que permitiu a expansão dos estudos na área, por meio de análises quantitativas e qualitativas. Quanto maior a amostragem, maiores as chances de descrever os aspectos de frequência da língua dentro de um contexto. No caso desta pesquisa, se não tivéssemos o auxílio do computador, o levantamento do nosso *corpus* e, principalmente, o levantamento de dados e sua análise seriam tarefas muito mais árduas, sobretudo por questões de tempo, tratando-se de um *corpus* extenso (1.161.213 palavras que compõem o *corpus* em inglês e 2.780.089 no *corpus* em português).

Para Berber Sardinha (2004), os estudos da descrição da linguagem não são recentes. Já no século XX, muitas pesquisas com *corpora* estavam em desenvolvimento. Ele explica que a diferença dos estudos atuais para os antigos está na forma de coleta dos dados, pois antes ela era feita manualmente e a credibilidade das pesquisas era questionada. O autor afirma que em 1950, as teorias racionalistas da linguagem, principalmente a teoria gerativa de Noam Chomsky, estavam em destaque, e os trabalhos com *corpora*, baseados no empirismo, perderam força. Os processamentos de *corpora* gigantescos eram os mais criticados, pois o ser humano não era capaz de analisá-los e as chances de erro eram muitas. Assim, “faltava justamente um instrumento que permitisse a análise de grandes quantidades de dados de modo confiável.” (BERBER SARDINHA, 2004, p. 04). Nesse contexto, o advento do computador foi determinante para o fortalecimento e para a expansão da Linguística de *Corpus*.

Neste trabalho, em razão de ser pautado por uma abordagem empirista, observaremos os padrões e as frequências das colocações levantadas, para depois interpretá-las, ou seja, partimos dos dados quantitativos para um estudo qualitativo. De acordo com Berber Sardinha (2004), a abordagem empirista, representada por Halliday (1961), vê a linguagem como um sistema probabilístico, ou seja, descreve a probabilidade dos sistemas linguísticos, dado os contextos em que os falantes a empregam; e se contrapõe à visão racionalista da linguagem, defendida pelo gerativismo chomskyano, fundamentado na introspecção e em estudos estruturais e funcionais da língua. Dessa maneira, segundo Berber Sardinha (2004, p. 30):

A Linguística de Corpus trabalha dentro de um quadro conceitual formado por uma abordagem empirista e uma visão da linguagem como sistema probabilístico; [...] empírico significa primazia dos dados provenientes da observação da linguagem, em geral reunidos sob a forma de corpus. (BERBER SARDINHA, 2004, p. 30)

Conforme explica o autor, o sistema probabilístico pressupõe que o fato de muitos traços linguísticos serem possíveis teoricamente, eles podem não ocorrer com a mesma frequência e não são aleatórios. Ou seja, a linguagem é padronizada e a padronização se evidencia pela recorrência e regularidade. A esse respeito, Orenha-Ottaiano (2004) explica que a área da Linguística que trata das combinações lexicais recorrentes é a Fraseologia, cujos aspectos serão tratados a seguir.

1.2. Fraseologia

A vida em sociedade é marcada por uma série de regras compartilhadas pelos seus indivíduos, as convenções sociais. Elas são de cunho moral, político, educacional etc., e permitem uma boa convivência entre os membros dessa comunidade. A língua também é uma convenção social e, como afirma Saussure (1973), é parte social da linguagem, não podendo existir a não ser por meio de um contrato estabelecido pelos membros da comunidade. Ao observarmos a língua, nos deparamos com expressões, blocos de palavras e construções que aparecem com maior ou menor frequência, são arbitrárias e se consagram pelo uso (ORENHA-OTTAIANO, 2004). Se arbitrários, um mesmo objeto é nomeado diferentemente por diferentes línguas, por exemplo, o que em língua portuguesa comumente chamamos de cachorro, em língua inglesa é nomeado *dog*, e em língua espanhola é denominado *perro*. Apesar dos nomes diferentes, o objeto continua o mesmo, ou seja, não há um elo intrínseco entre o signo (significante e significado) e o objeto em si. Assim, cada sociedade usa a língua de maneira diversa para representar a realidade.

Orenha-Ottaiano (2004) menciona que o termo convencionalidade foi empregado por Fillmore (1979, apud Tagnin, 1998) para designar “o conjunto dos elementos linguísticos, cuja coocorrência não é explicada sintática ou semanticamente, mas sim pelo uso”. As regras sintáticas, por exemplo, não explicam a posição dos adjetivos em português e em inglês. Orenha-Ottaiano (2004) explica que, em português, o adjetivo pode aparecer anteposto ou posposto ao substantivo, enquanto que em inglês o adjetivo aparece anteposto. A autora chama atenção para a troca de posições do adjetivo na língua portuguesa (o que pode acarretar ou não mudança no significado), por exemplo, a colocação *um grande negócio*, significa um negócio próspero, se dissermos *um negócio grande*, nos referimos a um negócio de grandes proporções, amplo. Em língua inglesa isso não acontece, e, dessa maneira, dois adjetivos são usados para expressar diferentes significados. Usando o exemplo anterior temos: *a big business*, para referirmos a um importante, próspero negócio e, *a large business*, em referência a um negócio amplo, de grandes extensões (ORENHA-OTTAIANO, 2004). Semanticamente, no caso das colocações, não há uma relação entre seus elementos e seu significado (ORENHA-OTTAIANO, 2004).

Dentro da convencionalidade, insere-se a Fraseologia, área de pesquisa da Linguística voltada para os estudos das combinações recorrentes na língua. Ortiz Álvarez (2000, p. 70) afirma que os fraseologismos “são unidades lexicais múltiplas que

apresentam vários graus de transparência semântica, que variam desde uma maior transparência à total opacidade”. Alguns fraseologismos são mais transparentes, ou seja, seu significado pode ser decodificado literalmente, como ocorre com os binômios e as colocações. Entretanto, outros fraseologismos são caracterizados pela não literalidade do significado, tais como expressões idiomáticas e verbos frasais (ORENHA-OTTAIANO, 2004). Orenha-Ottaiano (2004) resalta uma importante característica dos fraseologismos: sua relativa estabilidade. Afirma que a maioria de seus elementos não permite substituição por uma expressão ou palavra sinônima. Podemos dizer *it's none of your business!* (não é da sua conta!), mas não podemos dizer *it's none of your deal*, embora *deal* e *business* possam ser considerados sinônimos. A autora refere-se a Saussure (1973) e explica que os fraseologismos são transmitidos pela tradição, não são improvisados e, portanto, caracterizam a *langue* e não a *parole*. Dessa forma, o falante nativo, ao longo de sua vida, constrói um repertório de combinações mais ou menos fixas e as armazena em seu léxico mental, resgatando-as de maneira automática quando discursa.

As unidades fraseológicas, além de recorrentes, são mais ou menos estabilizadas/cristalizadas e abrangem as colocações, as coligações, as expressões idiomáticas, os binômios, os provérbios, os marcadores conversacionais, entre outros (ORENHA-OTTAIANO, 2004). Para esta pesquisa, focaremos as colocações, em razão de assumirem um papel importante, posto tratar-se de nosso objeto de estudo.

1.3. Colocações

Foi Firth (1951), autor da frase “uma palavra é conhecida pela companhia que mantém”¹, em seu artigo *Modes of meaning*, quem criou o termo *collocation*: “A partir de então, em meus argumentos, [...] proponho antecipar um termo técnico, chamado ‘colocação’, e aplicar o teste da ‘colocabilidade’” (FIRTH, 1951, apud KRISHNAMURTHY, p. 33).

Segundo Nesselhauf (2005), o termo *collocation* é muitas vezes usado em sentidos diferentes e bastante vagos na linguística e no ensino de línguas. O único denominador comum é que o termo se refere a “um tipo de relação sintagmática entre as palavras” (NESSELHAUF, 2005, p. 11). Para a autora, entre os diversos usos do termo, podem ser identificadas duas principais visões. Uma dessas visões, conhecida como *frequency-based approach* (abordagem baseada em frequência), define colocação como a coocorrência de palavras a certa distância, e distingue as coocorrências que são mais frequentes daquelas que não são. Segundo a outra visão, conhecida como *phraseological approach* (abordagem fraseológica), as colocações são vistas como um tipo de combinação de palavras, mais comumente aquelas que são fixas em certo grau, mas não completamente (NESSELHAUF, 2005). Ainda segundo Nesselhauf (2005), a abordagem baseada em frequência remonta aos estudos de Firth (1951), e tem sido desenvolvida, particularmente, por Halliday (1966) e Sinclair (1991). Esta abordagem é frequentemente adotada por pesquisadores envolvidos na análise computacional das relações sintagmáticas. A autora explica que a abordagem fraseológica tem sido fortemente influenciada pela fraseologia russa, cujos trabalhos voltam-se para as áreas

¹ Todas as traduções são de nossa autoria.

de lexicografia e pedagogia. Destaca que entre seus principais representantes estão Cowie (1998) e Hausmann (1984).

Sinclair (1991) define colocações como “a ocorrência de duas ou mais palavras dentro de um curto espaço no texto”. O espaço entre as palavras é definido pela distância de cerca de quatro palavras à direita ou à esquerda da palavra investigada, nomeada por Sinclair (1991) de *node*. Sinclair (1991), em oposição à gramaticalidade na produção de enunciados, propõe o princípio idiomático (*idiom principle*). Tal princípio preconiza que o usuário da língua tem a seu dispor um grande número de expressões “semi-pré-construídas”, o que permite a produção de textos naturais. Por este princípio, o falante não cria novas combinações a cada momento em que discursa, ele usa aquelas que já estão consolidadas na língua. A gramaticalidade não explica, por exemplo, por que o falante prefere dizer *pay a compliment* (fazer um elogio), e não *make a compliment*. Isso reforça a ideia de convencionalidade nas línguas (ORENHA-OTTAIANO, 2004).

No tocante à classificação das colocações, cabe mencionar Hausmann (1985), cujas contribuições para os estudos fraseológicos são bastante significativas. O autor fez uma distinção entre combinações fixas (formada pelas expressões idiomáticas) e não fixas, nestas últimas inserem-se as colocações. Hausmann propõe uma terminologia para referir-se aos elementos que compõem uma colocação: a **base** e o **colocado**. Para o teórico, existe uma hierarquia entre esses elementos. A base determina e o colocado é determinado. Por exemplo, na colocação *heart biopsy*, *heart* é colocado e *biopsy* é a base. Resumidamente, a base (*basis*) é:

- um elemento independente;
- semanticamente autônoma;
- traduzível, independentemente de seu uso na colocação;
- determina padrões lexicais que podem combinar com ela. (HEID, 1991).

Quanto ao colocado (*kolokator*), ele:

- funciona como um conceito modificador;
- é semanticamente interpretável somente dentro da colocação;
- sua tradução depende do uso na colocação;
- é escolhido por uma dada base para formar uma colocação. (HEID, 1991)

Orenha-Ottaiano (2004) resumiu algumas características que qualificam uma combinação como colocação, discutidas por Tagnin (1999), são elas:

- **Recorrência** – há necessidade de que a combinação seja recorrente (ter frequência superior a 1);
- **Não-idiomaticidade**, ou seja, seu significado deve ser composicional;
- **Coesão** – é necessário que haja uma ligação muito forte entre seus elementos, muito mais forte do que se esperaria de uma combinação qualquer;
- **Restrição contextual** – deve haver uma probabilidade de que ocorra dentro de um contexto específico;
- **Coocorrência arbitrária entre seus elementos** – não há razão semântica que explique tal coocorrência.

A partir das características acima, podemos diferenciar, por exemplo, as colocações das expressões idiomáticas. Estas são relativamente opacas, e possuem significado metafórico, como observamos em: *to kick the bucket* (bater as botas), cujo significado é “morrer”. Já as colocações não são tão fixas quanto às expressões idiomáticas e tampouco podem ser consideradas combinações livres. Algumas colocações, conforme discute Hill (2000) são consideradas fortes (*strong*) como em *moved to tear*, (“cair em lágrimas”), *ulterior motives* (“segundas intenções”); outras podem ser caracterizadas como médias, como em *heavy smoker* (“fumante viciado”), *hold a conversation* (“manter uma conversa”); outras, ainda, pode ser vistas como fracas (*weak*): *have lunch* (“almoçar”), *nice day* (“bom dia”), *red car* (“carro vermelho”). Há aquelas consideradas únicas (*unique*): *foot the bill* (“pagar a conta”). Neste último caso, o nativo não produziria, por exemplo, *foot the invoice* (“pagar a fatura”).

Apresentamos, a seguir, algumas colocações especializadas encontradas no *corpus*, exemplificadas com base na tipologia apresentada por Orenha-Ottaiano (2004, 2009), a partir da proposta de Hausmann (1985):

Verbais

- Verbo **base** + Substantivo **colocado** = *draw blood* (recolher sangue)
- Substantivo **base** + Verbo **colocado** = *the heart stopped* (o coração parou)
- Verbo **base** + Determinante + Substantivo **colocado** = *Do an MRI* (fazer uma ressonância)
- Substantivo **base** + Preposição + Verbo **colocado** = *a biopsy to confirm* (uma biópsia para confirmar)
- Verbo **colocado** + Pronome + Substantivo **base** = *take my meds* (tomar meus remédios)

Nominais

- Substantivo **base** + Substantivo **colocado** = *heart rate* (frequência cardíaca)
- Substantivo **base** + Preposição + Substantivo **colocado** = *MRI with contrast* (ressonância com contraste)

Adjetiva

- Adjetivo **colocado** + Substantivo **base** = *benign tumor* (tumor benigno)

Adverbial

- Verbo **base** + Advérbio **colocado** = *breathe heavily* (respirar pesadamente)

Conforme pode ser notado nos exemplos, todas as bases acima citadas são termos. De acordo com Orenha-Ottaiano (2009), diferentemente das colocações da língua geral, que têm como núcleo uma unidade lexical não especializada, as colocações especializadas têm como base uma unidade terminológica ou um termo, como é caso das colocações acima retratadas. Isto se deve pelo fato de o *corpus* tratar de temas da área médica, ou seja, é também composto por combinações convencionais empregadas por um grupo de especialistas.

3. COMPILAÇÃO DO *CORPUS*

Nosso *corpus* constitui-se pelas transcrições das legendas em inglês e em português de sete temporadas da série de televisão *House M.D.* Para compilá-lo, consideramos alguns aspectos importantes, tais como: autenticidade dos textos, propósito de pesquisa linguística, manuseio por computador e representatividade de uma dada variedade. Dessa forma, seguimos a definição de *corpus*, proposta por Berber Sardinha (2004):

Um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos, sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum dos seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise. (BERBER SARDINHA, 2004, p. 18)

Seguindo a mesma linha teórica de Berber Sardinha (2004), Frankenberg-Garcia (2008), ao tratar sobre compilação e uso de *corpora* paralelos, assinala que “um *corpus* é basicamente uma coleção extensa de textos naturais, selecionados de acordo com critérios específicos e armazenados em formato digital” (FRANKENBERG-GARCIA, 2008, p. 119). O *corpus* apresentado nesta pesquisa atendeu aos critérios acima mencionados. O *corpus* em inglês, por exemplo, totalizou 109 episódios. Segundo a classificação de Berber Sardinha (2004, p. 26), é um *corpus* considerado de médio-grande porte, pois é extenso, possuindo mais de um milhão de palavras (1.161.213), o que lhe garante maior representatividade. Pode também ser considerado específico (BERBER SARDINHA, 2004), uma vez que, embora contenha situações da língua geral, tem como conteúdo predominante a temática médica, possuindo, dessa maneira, léxico e vocabulário especializados. Além disso, é também um *corpus* paralelo bilíngue, ou seja, é formado por um conjunto de textos originais e suas respectivas traduções. De acordo com Frankenberg-Garcia (2008), um *corpus* paralelo “é uma combinação de pelo menos dois sub-*corpora* alinhados entre si. [...] de um lado, um sub-*corpus* composto de textos originais numa determinada língua (L1), e, do outro, um sub-*corpus* com os mesmos textos traduzidos para uma outra língua.” (FRANKENBERG-GARCIA, 2008, p. 119). Para a autora, tais textos, quando alinhados, permitem a extração de concordâncias, além de possibilitar pesquisas e traduções simultâneas.

Por meio do *software WordSmith Tools* (SCOTT, 2004), versão 4.0, fizemos o levantamento das colocações. Assim, salvamos as transcrições dos episódios do seriado em formato txt, o que permitiu o manuseio do *corpus* pelo referido *software*. Este possui três ferramentas básicas que possibilitaram a observação dos dados no *corpus*:

1. **WordList:** Esta ferramenta disponibiliza uma lista de palavras em ordem alfabética e por frequência, além de apresentar dados gerais do *corpus* de estudo. A partir dela também podemos estudar o tipo de vocabulário usado; identificar *clusters* de palavras comuns; comparar a frequência de uma palavra em arquivos de diferentes textos ou gêneros; comparar as frequências de palavras cognatas ou equivalentes de tradução entre línguas diferentes. A lista de palavras, salva em formato *lst*, também pode ser usada para gerar a lista de palavras-chave (*KeyWords*).

2. **Concord:** Esta ferramenta cria concordâncias das palavras de busca (listas de palavras em contexto), gera listas de *collocates* (listas das palavras que ocorrem à esquerda e à direita da palavra de busca selecionada, em ordem de frequência), e listas de padrões de colocados (frases comuns).
3. **KeyWords:** Esta ferramenta gera uma lista palavras-chave de um dado texto através de comparações entre listas de palavras do *corpus* de estudo e do *corpus* de referência (*British National Corpus*).

Por meio dessas ferramentas básicas, é possível a observação dos dados, mas cabe ao pesquisador interpretá-los e tirar conclusões para sua pesquisa. Uma das mensagens deixadas por Scott (2004) na janela de abertura do programa: “as ferramentas computacionais não pensam por você” (*Computer tools don’t do the thinking*), deixa claro como é indispensável a intervenção humana na interpretação dos dados. O computador, certamente, facilita o manuseio e a organização de grandes quantidades de dados, mas a seleção e o questionamento dos dados por parte do pesquisador é fundamental. Outras importantes ferramentas para a análise e interpretação das colocações levantadas foram os *corpora* de referência: *British National Corpus* (BNC), utilizado como referência para formarmos a lista de palavras-chave (*KeyWords*); *Corpus of Contemporary American English* (COCA), empregado para compararmos a frequência das colocações em inglês na língua em uso; o *Corpus do Português* e o *Corpus Brasileiro*, para compararmos as frequências na língua em uso das colocações em português levantadas no *corpus*. Além das ferramentas citadas, utilizamos o *WebCorp*, um conjunto de ferramentas que utiliza a *Web* como *corpus*, também para compararmos a frequência das colocações em inglês e em português; e o dicionário de colocações *Oxford Collocations Dictionary for Students of English* (MCINTOSH; FRANCIS; POOLE, 2009).

4. LEVANTAMENTO DAS COLOCAÇÕES

A partir das primeiras 500 palavras-chave (*KeyWords*), disponibilizadas pelo *software WordSmith Tools*, selecionamos aquelas que pertenciam ao léxico da medicina e que poderiam formar colocações da língua geral e colocações especializadas. Uma vez selecionadas as palavras-chave, o passo seguinte foi o levantamento das colocações com o auxílio da ferramenta *Concord*, por meio das abas *concordance* e *collocates*. A título de ilustração, tomaremos como exemplo o nódulo *blood* (sangue).

As linhas de concordância para o nódulo *blood*, geradas pela ferramenta *Concord*, mostraram-nos que ele era recorrente no *corpus*, com uma frequência de 1294. Pela frequência de *blood* no *corpus*, notamos que ele tinha potencial para formação de colocações. Dessa maneira, utilizamos a aba *collocates* da ferramenta *Concord* para identificarmos as palavras que apareciam com maior frequência à esquerda ou à direita da palavra de busca. A lista de *collocates* para *blood* mostrou 302 palavras, a partir das quais selecionamos aquelas pertencentes ao léxico da medicina e com possibilidade de formar colocações especializadas.

Apresentamos, para este artigo, as colocações especializadas formadas a partir dos nódulos *blood*, *brain*, *tumor* e *heart*. Cabe apontar que as colocações para estes nódulos estão separadas por sua formação morfológica, em que se destacam colocações substantivas: **noun + noun (substantivo + substantivo)**; colocações adjetivas:

adjective + noun (adjetivo + substantivo); e colocações verbais: **verb + noun (verbo + substantivo)**. Observamos também o equivalente da colocação em língua portuguesa, assim como exemplos extraídos do *corpus* em inglês e traduções para português, baseadas nas legendas do *corpus* em português, posto tratar-se de um *corpus* paralelo:

Quadro 1 – Colocações substantivas para o nóculo *blood*

BLOOD (Noun) + Noun	
Blood pressure	Pressão sanguínea
Could be an aneurysm. We've got to control her blood pressure . Start her on a calcium channel blocker.	Pode ser um aneurisma. Nós temos que controlar a pressão sanguínea . Iniciem bloqueadores dos canais de cálcio.
Blood flow	Fluxo sanguíneo
His blood's normal and his heart's pumping it up to the brain. But it's not all getting there, something's obstructing the blood flow .	O sangue está normal e o coração está bombeando para o cérebro. Mas não está chegando lá. Algo está obstruindo o fluxo sanguíneo .
Blood test	Exame de sangue
We have to do x-rays to make sure you don't have internal injuries. Blood test to make sure no infections. An EEG for neurological anomalies and biopsy a spinal nerve.	Temos que fazer raio-x pra ver se não tem ferimentos internos. Exame de sangue pra infecções, EEG para anormalidades neurológicas, e biópsia de um nervo da coluna.
Blood cells	Células sanguíneas
His body's not making new blood cells . His bone marrow's crashing.	O corpo não está produzindo novas células sanguíneas . A medula está se despedaçando.
Blood type	Tipo sanguíneo
How do you know what blood type I am?	Wilson: Como sabe de que tipo sanguíneo eu sou?
Blood loss	Perda de sangue
There's no reason for blood loss .	Não há razão para perda de sangue .
Blood vessels	Vasos sanguíneos
He has air in his blood vessels . If we don't open the artery to his intestines...	Ele tem ar nos vasos , se não abrímos a artéria até o intestino...
Blood disorder	Distúrbio hemorrágico
We cured his bleeding disorder, removed all the damaged cells we could find.	Nós curamos o distúrbio hemorrágico . Removemos todas as células danificadas que encontramos.

Blood sugar	Açúcar no sangue; nível de açúcar
Sometimes the blood sugar levels can fluctuate, make it difficult to diagnose.	Algumas vezes o nível de açúcar pode oscilar tornando difícil diagnosticar.
Blood clot	Coágulo sanguíneo; coágulo
We found a blood clot in your brain. We'd like to start you on heparin, it's a blood thinner.	Encontramos um coágulo no seu cérebro. Queremos dar-lhe heparina, um anticoagulante.
Blood sample	Amostra de sangue
We need a blood sample . We got to do the antientromere antibody test. It'll show House is right. It's scleroderma.	Precisamos de uma amostra de sangue . Precisamos fazer o teste de anticorpos antientrômero. Vai mostrar que House acertou, é esclerodermia.
Blood trasfusion	Transfusão de sangue
He needs blood transfusions every two weeks, which means he can dope all he wants. He's got a doctor's pass. It's medicinal. He got away with it.	Ele precisará de transfusões de sangue a cada duas semanas, o que significa que ele pode se dopar o quanto quiser. Ele tem receita médica. É medicinal. Ele escapou disto.
Blood thinners	Anticoagulantes
She only drinks when you give it to her. We put her on blood thinners . You can probably take her home tomorrow.	Ela só bebe quando lhes dão. Demos-lhe anticoagulantes . Provavelmente podes levá-la para casa amanhã.
Blood stream	Corrente sanguínea
Perforated intestine. If this thing started as normal bacteria living in the intestines but got into a blood vessel through a vascular anomaly in the bowel wall, then they would affect every organ through the blood stream . It screws up everything. For everyone.	Intestino perfurado. Se isso começou com uma bactéria vivendo no intestino, mas chegou a um vaso sanguíneo por uma anomalia nas paredes intestinais, então afetaria todos os órgãos pela corrente sanguínea . Arruinando tudo. Para todos.
Blood smear	Esfregaço de sangue
Do a peripheral blood smear . Look for acanthocytes.	Faça um esfregaço de sangue periférico. Procure acantócitos.
Blood culture	Hemocultura
Blood culture was negative for bacteria.	A hemocultura deu negativo para bactérias.

Quadro 2 – Colocações verbais para o nódulo *blood*

Verb + BLOOD (Noun)	
Draw blood	Recolher sangue
Let's keep him on the broad-spectrum antibiotics, and since he's displaying septic physiology, draw blood for adrenal and thyroid function.	Vamos mantê-lo com os antibióticos normais, e uma vez que apresenta uma fisiologia séptica, recolham sangue para testar as funções supra-renal e tiroidiana.
Lose blood	Perder sangue
Patient's losing blood faster than we can transfuse.	A paciente está perdendo mais sangue do que conseguimos repor.
Cough blood	Cough blood
We still don't know why Ryan coughed blood , and now he's got small hemorrhages in his eyes.	Ainda não sabemos por que Ryan tossiu sangue , e agora ele tem pequenas hemorragias nos olhos.
Donate blood	Doar sangue
Can I at least donate blood ?	Posso pelo menos doar sangue ?

A fim de verificar o uso e a frequência das colocações levantadas no *corpus* de estudo em relação a outros *corpora* anteriormente mencionados, apresentamos a frequência levantada da colocação *blood pressure* no *corpus* de estudo em inglês (item 1 da tabela 1), no *corpus* de estudo em português, ou seja de sua tradução para o português – “pressão sanguínea” (item 2 da tabela 1), no *Corpus of Contemporary American English COCA* (item 3 da tabela 1); *Corpus do Português* (item 4 da tabela 1) e *Corpus Brasileiro* (item 5 da tabela 1), igualmente verificando a frequência da colocação “pressão sanguínea”; e *WebCorp*, observando a frequência da colocação *blood pressure* em inglês (item 6 da tabela 1) e de sua tradução para o português (item 7 da tabela 1).

Tabela 1 – Estatísticas da colocação *blood pressure*

BLOOD (Noun) + Noun						
blood pressure - pressão sanguínea						
We've got to control her blood pressure .						
Nós temos que controlar a pressão sanguínea .						
1 – <i>Corpus</i> em inglês	2- <i>Corpus</i> em português	3 - COCA	4 – <i>Corpus</i> do Português	5 – <i>Corpus</i> Brasileiro	6 – <i>WebCorp</i> (inglês)	7 – <i>WebCorp</i> (português)
89	39	5616	30	13618	3253	295

Segundo podemos notar na tabela acima, tanto a frequência da colocação em inglês (*blood pressure*) quanto a frequência da colocação em português (“pressão sanguínea”) são bastante significativas, evidenciando que se trata de colocações comumente empregadas nas duas línguas investigadas. Além disso, podemos observar que a própria frequência das colocações no *corpus* de estudo paralelo, tanto no texto original quanto na tradução, é bastante expressiva nas duas línguas.

Quadro 3 – Colocações substantivas para o nódulo *brain*

BRAIN (Noun) + Noun	
Brain damage	Dano cerebral
She obviously has brain damage .	Ela tem dano cerebral .
Brain tumor	Tumor no cérebro
You seem to have a massive brain tumor .	Parece ter um enorme tumor no seu cérebro .
Brain biopsy	Biópsia no cérebro
Get a brain biopsy to confirm .	Faça uma biópsia no cérebro para confirmar.
Brain stem	Tronco cerebral
The brain stem controls the body's involuntary systems.	O tronco cerebral controla o sistema autônomo.
Brain cancer	Câncer cerebral
My Marisa was diagnosed with terminal brain cancer .	Minha Marisa foi diagnosticada com câncer cerebral terminal.
Brain surgery	Neurocirurgia
The patient just had brain surgery .	O paciente acabou de fazer uma neurocirurgia .
Brain cells	Células nervosas
There are molecular similarities between brain cells and tumor cells.	Existem similaridades moleculares entre as células nervosas e as tumorais.
Brain infection	Infecção cerebral
Well maybe he doesn't have cancer maybe he has a brain infection that's causing multiple abscesses.	Talvez ele não tenha câncer, talvez uma infecção cerebral que está causando acessos múltiplos.
Brain scan	Tomografia cerebral
Brain scan was completely clean.	A tomografia cerebral deu completamente normal

Brain problem	Problema cerebral
You made the right call. The problem was a brain problem . Without the procedure, House never notices the increased left-brain.	Você fez a coisa certa. O problema era um problema cerebral . Sem o procedimento, House nunca teria notado o aumento da atividade no lado esquerdo do cérebro.
Brain function	Funções do cérebro
If this isn't just ancient history then maybe it's something we can correct. Might even get some brain function back.	Se isto não é uma antiga história, talvez é algo que nós podemos corrigir. Pode até voltar algumas funções do cérebro .
Brain activity	Atividade cerebral
Brain activity isn't accelerating, it's slowing down.	A atividade cerebral não está acelerando, está lentificando.
Brain symptoms	Sintomas do cérebro
Brain symptoms that aren't brain symptoms, tumors that come and go?	Sintomas do cérebro que não são os sintomas do cérebro, Os tumores que vêm e vão?
Brain injury	Lesão cerebral
Vomiting means your brain injury is getting worse.	O vômito significa que a sua lesão cerebral está piorando.
Brain waves	Ondas cerebrais
Saw a very small diffuse abnormality in the brain waves .	Vi uma pequena e difusa irregularidade nas ondas cerebrais .
Brain disease	Doença cerebral
Could be anything from an environmental factor to an early sign of a degenerative brain disease .	Pode ser qualquer coisa desde um fator ambiental a uma doença degenerativa cerebral .
Brain hemorrhage	Hemorragia cerebral
Your son could die of a brain hemorrhage .	Seu filho pode morrer de uma hemorragia cerebral .
Brain death	Morte cerebral
Garden-variety EEG sucks compared to the in-brain variety, which is not gonna miss brain death .	EEG externo é uma merda comparado com EEG interno, mas não iria deixar passar morte cerebral .
Brain aneurism	Aneurisma cerebral; aneurisma
Her son didn't have a brain aneurism .	O filho não teve um aneurisma .
Brain herniation	Herniação do cérebro

He did a brain biopsy without CT mapping, causing brain herniation and death.	Ele fez uma biópsia cerebral sem tomografia, causando herniação do cérebro .
Brain angiogram	Angiograma cerebral
Steroids to treat, brain angiogram to confirm, EMG and nerve biopsy while you're at it.	Trate com corticoides, faça um angiograma cerebral . Eletroneuromiografia e biópsia neural, enquanto isso.
Brain dysfunction	Defeito no cérebro
No signs of spasm, impeded blood flow, or brain dysfunction .	Sem sinais de espasmos, obstrução na corrente sanguínea ou defeito no cérebro .
Brain stimulation	Estimulação cerebral
You said you wanted to try deep brain stimulation .	Você disse que queria tentar estimulação cerebral profunda.
Brain tissue	Tecido cerebral
We're gonna remove a small piece of brain tissue .	Temos que remover um pequeno pedaço do seu tecido cerebral .
Brain lesions	Lesão cerebral
Brain lesions would've shown up on the CT.	Lesão cerebral iria aparecer na tomografia.
Brain involvement	Envolvimento cerebral
Brain involvement gives us a new symptom.	Envolvimento cerebral nos dá um novo sintoma.
Brain abscess	Abcesso cerebral
The infection could be limited to a brain abscess .	A infecção poderia estar limitada ao abcesso cerebral .

Quadro 4 – Colocações adjetivas para o nódulo *brain*

Adjective + brain (noun)	
Left brain	Lado esquerdo do cérebro
Left brain has language, arithmetic, rationality.	O lado esquerdo controla a fala, aritmética, razão.
Right brain	Lado direito do cérebro
Hemispherectomy would completely stop the right brain seizure activity and he would no longer need to take his anticonvulsive medication.	Hemisferectomia poderia parar completamente a atividade de convulsão do lado direito do cérebro , e ele não precisaria mais tomar a medicação anticonvulsivante.

Tabela 2 – Estatísticas da colocação *brain stem*

BRAIN (Noun) + (Noun)						
Brain stem – tronco cerebral						
The brain stem controls the body's involuntary systems. O tronco cerebral controla o sistema autônomo.						
1 – <i>Corpus</i> em inglês	2 – <i>Corpus</i> em português	3 - COCA	4 – <i>Corpus</i> do Português	5 – <i>Corpus</i> Brasileiro	6 – <i>WebCorp</i> (inglês)	7 – <i>WebCorp</i> (português)
27	22	172	19	773	697	191

Observando a tabela acima, podemos concluir que a frequência para a colocação no *corpus* de estudo em inglês (item 1 da tabela 2) e no *corpus* de estudo em português (item 2 da tabela 2) é bem próximo, 27 e 22 ocorrências, respectivamente. Quanto aos *corpora* utilizados como referência, notamos menor frequência (19 ocorrências) da colocação no *Corpus do Português* (item 4 da tabela 2). Nos demais, os números para a colocação *brain stem* (tronco cerebral) são mais significativos.

Quadro 5 – Colocações substantivas para o nódulo *tumor*

TUMOR (Noun) + Noun	
Brain tumor	Tumor cerebral
MRI revealed no stroke or brain tumor .	A ressonância não indica nem derrame ou tumor cerebral .
Liver tumor	Tumor no fígado
All he does is stroke, blood clot, liver tumor .	Tudo que esse cara faz é derrame, coágulo sanguíneo, tumor no fígado .
Heart tumor	Tumor no coração
She didn't bleed out of her eye from a heart tumor .	Ela não sangrou pelo olho por causa do tumor.
Tumor cells	Células tumorais
There are molecular similarities between brain cells and tumor cells .	Existem similaridades moleculares entre as células nervosas e as tumorais .
Lobe tumor	Tumor no lóbulo
MRI was clean. No frontal lobe tumor .	A ressonância estava limpa. Sem tumores no lóbulo frontal .

Spinal tumor	Tumor na espinha
The spinal tumor wouldn't explain the liver or the lungs.	Um tumor na espinha não explicaria o fígado ou os pulmões.
Pituitary tumor	Tumor na hipófise
Pituitary tumor . Could also explain his libido and heart issues.	Tumor na hipófise . O que poderia explicar sua libido e os problemas cardíacos.
Pancoast tumor	Tumor de Pancoast
It's gotta be a Pancoast tumor . Patient's got dyspnea, seventh-nerve palsy. He's sweating on one side of his face.	Tem que ser tumor de Pancoast . O paciente tem dispnéia, Paralisia no sétimo nervo.

Quadro 6 – Colocações adjetivas para o nódulo *tumor*

Adjective + tumor (Noun)	
Small tumor	Pequeno tumor
A small tumor in his brain stem, causing paraneoplastic syndrome.	Pequeno tumor no tronco cerebral, causando síndrome paraneoplásica.
Benign tumor	Tumor benigno
A nine year old with terminal cancer gets an unrelated benign tumor growing in her heart, why?	Uma garota de 9 anos com câncer terminal tem um tumor benigno não relacionado ao outro, crescendo no coração, por quê?
Solid tumor	Tumor sólido
Clotting could be a problem well before a solid tumor forms.	A coagulação pode ser um problema desde antes da formação de tumores sólidos .
Cancerous tumor	Tumor cancerígeno
Cancerous tumor is still present, on the right lobe.	Tumor cancerígeno ainda presente no lóbulo direito.
Primary tumor	Tumor primário
There is a chance, by removing the eye, get the primary tumor , and three courses of radiation, that could...	Há uma chance que removendo o olho, que é o tumor primário , mais 3 sessões de radiação, isso poderia...
Biliary tumor	Tumor biliar
Biliary tumor causing paraneoplastic syndrome.	Um tumor biliar causando síndrome paraneoplásica.

Nasal tumor	Tumor nasal
I just said it wasn't a nasal tumor .	Acabei de dizer que não é um tumor nasal .
Carcinoid tumor	Tumor carcinóide
A corticotropin-producing carcinoid tumor of the intestine.	A corticotropina produz um tumor carcinóide no intestino.
Glomus tumor	Tumor glômico
Took a shot to the back of the neck. More specifically, to an abnormal growth of nerves caused by a glomus tumor .	Levou um soco na nuca. Mais especificamente, num crescimento anormal de nervos causado por um tumor glômico .
Neuroendocrine tumor	Tumor neuroendócrino
What about a neuroendocrine tumor ? Not without diminished mental capacity, loss of judgment. Disseminated intravascular coagulation, however, fits.	E um tumor neuroendócrino ? Não sem diminuição da capacidade mental, perda de julgamento. Coagulação intravascular disseminada, no entanto, se encaixa.
Pancreatic tumor	Tumor no pâncreas
The platelet dysfunction and cardiac arrest points to a pancreatic tumor .	A disfunção plaquetária e a parada cardíaca apontam para um tumor no pâncreas .
Thoracic tumor	Tumor torácico
That means it's probably something wrong with the blood itself, Leukemia, Von Willebrand's. Thoracic tumor is a better fit. Erodes into her airway and esophagus.	Isso significa que pode ser algo errado com o sangue. Leucemia, doença de von Willebrand. Tumor torácico encaixa melhor. Corroendo sua via aérea e o esôfago.
Vascular tumor	Tumor vascular
You're thinking vascular tumor on her spine?	Você está pensando em tumor vascular na espinha?
Nasolacrimal tumor	Tumor nasolacrimal
His present symptoms could mean autoimmune disease, nasolacrimal tumor , an infection or just a cold.	Os sintomas dele podem significar doença autoimune, tumor nasolacrimal , uma infecção ou só um resfriado.

Quadro 7 – Colocações verbais para o nódulo *tumor*

Tumor (noun) + (verb)	
A tumor shrinks	Um tumor encolhe
Early 1900s an Italian medical journal wrote up a woman with cervical cancer who was	Nos primeiros anos de 1900, tinha um jornal médico italiano escreveu sobre uma mulher

injected with a weak strain of rabies, I've no idea why they did that, but her tumor shrank .	com câncer cervical, que foi injetada com um tipo estranho de herpes fraco. Não tenho idéia porque fizeram isso, mas o tumor dela encolheu .
Remove the tumor	Remover o tumor
She'll need surgery to remove the tumor . Once it's gone, everything will get normal very fast.	Ela vai precisar de uma cirurgia para remover o tumor . Quando for removido, tudo vai voltar ao normal rapidamente.
Indicate a tumor	Indicar um tumor
Slight bowing in the wall of the fourth ventricle could still indicate a tumor .	A parede do quarto ventrículo está um pouco curvada ainda pode indicar tumor .

Tabela 3 – Estatísticas da colocação *primary tumor*

Adjective + TUMOR (Noun)						
Primary tumor – Tumor primário						
There is a chance, by removing the eye, get the primary tumor, and three courses of radiation, that could...						
Há uma chance que removendo o olho, que é o tumor primário, mais 3 sessões de radiação, isso poderia...						
1 - Corpus em português	2 - Corpus em inglês	3 - COCA	4 - Corpus do Português	5 - Corpus Brasileiro	6 - WebCorp (inglês)	7 - WebCorp (português)
1	1	76	0	613	564	33

A colocação *Primary tumor* (tumor primário) não é expressiva no *corpus* de estudo, aparecendo apenas uma vez. No entanto, apresenta frequências maiores no *Corpus Brasileiro* (item 5 da tabela 3) e no *WebCorp* (item 6 da tabela 3), com 613 e 564 ocorrências, respectivamente.

Quadro 8 – Colocações substantivas para o nódulo *heart*

Heart (Noun) + Noun	
Heart rate	Batimento cardíaco
We combine the increased heart rate with chest trauma by applying a little sweet science to the science.	Combinamos a subida do batimento cardíaco com o trauma no peito ao aplicar um pouco da doce ciência à ciência.
Heart attack	Ataque cardíaco
We induce a massive heart attack and kill off the extra muscle.	Causamos um ataque cardíaco e matamos o músculo sobressalente.

Heart failure	Insuficiência cardíaca
Q fever would cause congestive heart failure , not myocardial infarction.	Febre Q causaria insuficiência cardíaca congestiva, e não enfarte do miocárdio.
Heart problem	Problema no coração
A clot means a heart problem , we have to do an emergency angiogram.	Um coágulo indica um problema no coração . Precisamos fazer um angiograma de emergência.
Heart damage	Dano ao coração
Not all heart damage has to be structural. Drugs could have caused intermittent tachycardia.	Nem todo dano ao coração tem de ser estrutural.
Heart disease	Doença cardíaca
Any history of heart disease in your family?	Tem casos de doenças cardíacas na família?
Heart block	Bloqueio cardíaco
It'll fix the heart block . Bradycardia was me.	Isso vai corrigir o bloqueio cardíaco . Bradicardia foi comigo.
Heart condition	Problema no coração
He's got a heart condition that caused the arteries in his arm to constrict.	Ele tem um problema no coração que faz as artérias contraírem.
Heart valve	Válvula do coração
We screwed up his heart valve by making his cold worse.	Nós ferramos a válvula do coração dele ao fazer a gripe piorar.
Heart rhythm	Ritmo do coração
Hopefully it'll tell us whether his heart rhythm is abnormal.	Esperamos que ele nos diga o porquê do ritmo anormal do coração .
Heart arrhythmia	Arritmia cardíaca
Heart arrhythmia , kidney failure, and now bleeding lung.	Arritmia cardíaca , insuficiência renal, e agora sangramento pulmonar.
Heart muscle	Músculos do coração
The colchicine interferes with the ability of the heart muscle to contract pumping blood, lowering your blood pressure.	A colchicina interfere com a capacidade dos músculos do coração se contraírem e bombear o sangue fazendo a pressão arterial descer.
Heart surgery	Cirurgia cardíaca

I've booked our patient in for heart surgery .	Marquei cirurgia cardíaca para nossa paciente.
Heart defect	Defeito no coração
What about a congenital heart defect ? He has a midline anomaly on his head, makes sense he'd have a midline anomaly in his heart too.	E se ele tiver um defeito no coração ? Ele tem anomalia da linha mediana na cabeça, faz sentido que tenha a anomalia em seu coração também.
Heart palpitations	Palpitações no coração
I get skin rashes, and heart palpitations , and insomnia.	Eu tenho irritações na pele, e palpitações no coração , e insônia.
Heart transplant	Transplante de coração
She needs a heart transplant . We'll get her on the list immediately.	Ela precisa de um transplante de coração. Vamos coloca-la na lista imediatamente.

Quadro 9 – Colocações verbais para o nódulo *heart*

Heart (Noun) + Verb	
Restart sb's heart	Reiniciar o coração
Run the blood cultures. After we get confirmation, then we'll restart her heart .	Faça as hemoculturas. Depois que confirmarmos, reiniciamos o coração .
A Heart looks	Coração parece
The heart looked healthy at the last echo and an arrhythmia wouldn't cause clots in a healthy ventricle.	O coração parecia bem no último eco. E arritmia não causa coágulo num ventrículo saudável.
A Heart beats	Coração bate
We need to put in a cardioverter-defibrillator. It'll keep your heart beating while we keep looking.	Precisamos colocar um cardioversor. Manterá seu coração batendo enquanto continuamos procurando.
A Heart stops	Coração para
The only permanent solution is to cut into your brain and remove your pituitary gland, but the surgery's dangerous. Your heart stopped on the table once already.	A única solução permanente é cortar o seu cérebro e retirar sua glândula pituitária, mas a cirurgia é perigosa. Seu coração já parou na mesa uma vez.

Tabela 4 – Estatísticas da colocação *heart rate*

Heart (Noun) + (noun)						
Heart rate – batimentos cardíacos						
We combine the increased heart rate with chest trauma by applying a little sweet science to the science. Combinamos a subida do batimento cardíaco com o trauma no peito ao aplicar um pouco da doce ciência à ciência.						
1 – Corpus em português	2 – Corpus em inglês	3 - COCA	4 – Corpus do Português	5 – Corpus Brasileiro	6 – WebCorp (inglês)	7 – WebCorp (português)
97	10	2064	20	706	1319	149

A colocação *heart rate* (batimentos cardíacos) é recorrente no *corpus* de estudo em inglês, com 97 ocorrências (item 1 da tabela 4). Porém, observamos que a frequência da colocação para o *corpus* de estudo em português é muito menor, aparecendo apenas 10 vezes. Tal divergência poderia ser explicada pelo fato de, em português, a colocação ser sinônima de frequência cardíaca. Nos *corpora* de referência, excetuando o *Corpus do Português* (item 4, da tabela 4), as frequências para *heart rate* (batimentos cardíacos) são bastante significativas.

Uma possível explicação para a baixa frequência das colocações acima exemplificadas (*pressão sanguínea* – 30; *tronco cerebral* – 19; *tumor primário* – 0; *batimentos cardíacos* – 20) no *Corpus do Português*, poderia estar no fato do *corpus* ser composto de textos do português do século XIV ao século XX, restringindo, assim, o aparecimento de colocações com alta frequência. No caso do *WebCorp*, por exemplo, a frequência das colocações são bem maiores, justamente porque tal ferramenta faz uma varredura em diversos *web sites*, em busca da colocação.

5. CONCLUSÃO

Conforme os passos descritos acima, fizemos o levantamento das colocações e das colocações especializadas encontradas em nosso *corpus*. A próxima etapa de nossa pesquisa consiste na compilação de um glossário bilíngue de colocações e colocações especializadas, nas direções tradutórias inglês-português/português-inglês, com base no *corpus* paralelo bilíngue, constituído pelas transcrições das legendas em inglês e em português, dos episódios do seriado em questão. Este poderá servir como fonte de pesquisa para estudantes de Letras, de Tradução, e demais interessados no estudo das colocações e colocações especializadas da área médica, pois, muitas vezes, tais combinações não são encontradas em dicionários de língua geral, tampouco especializados.

Além disso, os dados resultantes de nossa pesquisa poderão também ser explorados em sala de aula, haja vista que as colocações, por prezarem pelo ensino contextualizado e fundamentado em formas recorrentes na língua em uso, colaboram para o enriquecimento do léxico mental do aluno, permitindo a produção de enunciados mais precisos e coerentes, de modo que este se comunique e interaja naturalmente, sem precisar utilizar a língua materna como base para construir seu discurso.

Ao organizar as colocações da língua geral e, especialmente, da língua de especialidade, observamos algumas dificuldades para classificá-las ora como colocações adjetivas, ora como nominais. Desse modo, enfatizamos a necessidade de o pesquisador se aprofundar nos estudos colocacionais, observar atentamente os dados, a fim de assegurar e facilitar a etapa classificatória das colocações e inseri-las de forma coerente na obra a ser compilada. Outro aspecto importante a ser considerado está relacionado à macro e microestrutura da obra. Faz-se necessário defini-las de forma bastante harmônica e coesa, para que se torne um material de consulta fácil e lógica.

REFERÊNCIAS

- BERBER SARDINHA, Tony. *Linguística de Corpus*. Manole, 2004.
- British National Corpus (BNC)*. Disponível em: <http://www.lexically.net/wordsmith/support/extras.html>. Acesso em: 20 jan. 2012.
- CORPUS BRASILEIRO. Disponível em: <http://corpusbrasileiro.pucsp.br/x/>. Acesso em: 19 jan. 2012.
- COWIE, A. P. (ed.). *Phraseology: theory, analysis, and applications*. Oxford: Oxford University Press, 1998.
- DAVIES, Mark. (2008-) *The Corpus of Contemporary American English: 425 million words, 1990-present*. Disponível em: <http://corpus.byu.edu/coca/>. Acesso em: 20 jan. 2012.
- _____; FERREIRA, Michael. (2006-) *Corpus do Português: 45 million words, 1300s-1900s*. Disponível em: <http://www.corpusdoportugues.org>. Acesso em 20 de jan. 2012.
- FRANKENBERG-GARCIA, Ana. *Compilação e uso de corpora paralelos*. In: STELLA, E. O.; TAGNIN, Oto Araújo Vale. São Paulo: Humanitas, 2008, p. 117-136.
- FILLMORE, Charles J. On fluency. In: FILLMORE, C. J.; KEMPLER, D.; WANG, W. S. Y. (Eds.). *Individual differences in language ability and language behavior*. New York: Academic Press, 1979, p. 85-99.
- HALLIDAY, Michael. Corpus studies and probabilistic grammar. In: AIJMER, K.; ALTENBERG, B. (eds.). *English corpus linguistics: studies in honour of Jan Svartvik*. London: Longman, 1991.
- HAUSMANN, Franz Joseph. Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen. *Praxis des neusprachlichen Unterrichts* 31, 1984, p. 395-406.
- HAUSMANN, Franz Joseph. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels'. In: BERGENHOLTZ, H; MEGDAN, J. (Orgs.). *Lexikographie und Grammatik*. Tübingen: Niemeyer, 1985.
- HEID, Ulrich; MARTIN, Willy; POSCH, Ilse. An overview of approaches towards the description of collocations. *Feasibility of standards for collocational description of lexical items. Eurotra 7- Report*, Stuttgart/Amsterdam, 1991.
- HILL, Jimmy. Revising priorities: from grammatical failure to collocational success. In: LEWIS, M. (eds). *Teaching collocation: Further developments in the lexical approach*. London: Language Teaching Publications, 2000, p. 47-69.
- HYPNOWEB. Disponível em: <http://dr-house.hypnoweb.net/episodes-.97.2/>. Acesso em: 15 abr de 2011.
- HOUSE MD – Guide. Disponível em: <http://www.housemd-guide.com/episodes.php>. Acesso em: 15 abr de 2011.

- KRISHNAMURTHY, Ramesh. *Keep good company: collocation, corpus, and dictionaries: corpus and dictionaries*. Cicle de Conferències 95-96 Lèxic, corpus I diccionaris, Institut Universitari de Linguística Aplicada, Universitat Pompeu Fabra, 1996, p. 31-56.
- MCINTOSH, Colin; FRANCIS, Ben; POOLE, Richards (Orgs.). *Oxford Collocations Dictionary for Students of English*. 2nd ed. Oxford: Oxford University Press, 2009, 963f.
- MEYER, Charles. F. *English Corpus Linguistics: an introduction*. Cambridge: Cambridge, 2004.
- NESSELHAUF, Nadja. *Collocations in a Learner Corpus*. Amsterdam: John Benjamins Publishing Company, 2005, p. 01-54.
- ORENHA, Adriane. *Unidades fraseológicas especializadas: colocações e colocações estendidas em contratos sociais e estatutos sociais traduzidos no modo juramentado e não juramentado*. 2009. 282f. Tese (Doutorado em Estudos Linguísticos) – IBILCE, Universidade Estadual Paulista, São José do Rio Preto.
- ORENHA-OTTAIANO, Adriane. *A compilação de um glossário bilíngue de colocações, na área de jornalismo de Negócios, baseado em corpus comparável*. 2004. 246f. Dissertação (Mestrado em Estudos Linguísticos e Literários), FFLCH/USP, São Paulo, SP.
- ORTIZ ALVAREZ, Maria Luiza. *Expressões idiomáticas do português do Brasil e do espanhol de Cuba: estudo contrastivo e implicações para o ensino do português como língua estrangeira*. 2000. 334f. Tese (Doutorado em Linguística Aplicada: Ensino/Aprendizagem de Segunda Língua e Língua Estrangeira) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.
- SAUSSURE, Ferdinand. *Curso de Linguística Geral*. São Paulo: Cultrix, 1973.
- SINCLAIR, John. *Corpus, concordance and collocation*. Oxford: Oxford University Press, 1991.
- SCOTT, Michael. *WordSmith Tools: version 4.0*. Oxford: Oxford University Press, 2004.
- TAGNIN, Stella Esther Ortweiller. *Collecting data for a bilingual dictionary of verbal collocations: from scraps of paper to corpora research*. In: PALC' 99 Practical applications in Language Corpora. Lodz: Lodz University Press, 1999.
- WEBCORP. Disponível em: <http://www.webcorp.org.uk/live/>. Acesso em: 21 jan. 2012.

Recebido em 01/05/2012

Aceito em 20/06/2012

Versão revisada recebida em 25/06/2012

Publicado em 30/06/2012

SPECIALIZED COLLOCATIONS IN THE MEDICAL FIELD EXTRACTED FROM THE *HOUSE M.D. CORPUS*

ABSTRACT: *In the light of Corpus Linguistics as well as Phraseology, it will be discussed some theoretical and methodological aspects, which allowed the extraction of collocations and specialized collocations from the medical field taken from the corpus compiled with the transcriptions of the House M.D. series. Moreover, the specialized collocations extracted from the nodes 'blood', 'brain', 'tumor' and 'heart' will be presented, in order to illustrate how the collocations will be organized in a bilingual glossary of specialized collocations in the medical field under compilation.*

KEYWORDS: *Corpus Linguistics; Phraseology; specialized collocations; bilingual glossary.*