

A Linguística de *Corpus* como recurso metodológico para investigar a metaforicidade dos *phrasal verbs*

Corpus Linguistics as a methodological resource to investigate the metaphoricity of phrasal verbs

Samanta Kélly Menoncin Pierozan*

Resumo

Este artigo demonstra a Linguística de *Corpus* como um recurso metodológico para efetuar a coleta de dados autênticos e representativos para uma pesquisa maior em Linguística Aplicada. Trata-se de uma proposta que possibilita identificar os *phrasal verbs* (PVs) mais frequentes no *corpus The Corpus of Contemporary American English* (COCA), apresentando os procedimentos realizados para a extração de concordâncias associadas. O procedimento aqui apresentado foi delineado de modo a atender objetivos previamente estipulados, no caso: a investigação da metaforicidade de PVs sob a ótica da semântica cognitiva, em especial os que dizem respeito às partículas *up* e *down*. A estrutura delineada neste trabalho serve como um tutorial para outras pesquisas de interesse comum.

Palavras-chave

Linguística de *Corpus*. COCA. *Phrasal verbs*. Linguística Aplicada.

Abstract

This article demonstrates the Corpus Linguistics as a methodological resource to collect authentic and representative data for a wide research in Applied Linguistics. It is a proposal to identify the most frequent phrasal verbs (PVs) in the *corpus The Corpus of Contemporary American English* (COCA), showing the stages followed for extracting associated concordances. The procedures designed go forward some previously aims, in this case: to investigate the metaphoricity of phrasal verbs from the perspective of cognitive semantics, especially those related to the particles 'up' and 'down'. The structure outlined in this study serves as a tutorial for other research of common interests.

Keywords

Corpus Linguistics. COCA. *Phrasal verbs*. Applied Linguistics.

* Doutora em Letras pela Universidade de Caxias do Sul (UCS).

Introdução

O presente trabalho refere-se ao procedimento metodológico, desenvolvido para contribuir com a dissertação de mestrado, intitulada *A metaforicidade dos phrasal verbs constituídos por up e down: uma investigação sob a ótica da semântica cognitiva* (PIEROZAN, 2015), a qual tem a Linguística Cognitiva (LC) como pilar. O procedimento aqui apresentado é decorrente de objetivos previamente estabelecidos na pesquisa. Entende-se que toda e qualquer metodologia deve atender questões norteadoras elaboradas a partir de um objeto de estudo, que nesse caso, diz respeito a *phrasal verbs*¹. Os PVs são ‘construções verbo-partícula’² muito utilizadas entre falantes nativos da língua inglesa. Essas construções têm sido consideradas tópicos problemáticos para aprendizes de inglês como língua estrangeira (LE), uma vez que se tratam de estruturas distintas às da língua portuguesa, causando dificuldade na compreensão de seus sentidos. O significado de um PV se dá pela construção de mais de uma palavra, verbo mais partícula(s). A partícula³ refere-se a um advérbio ou preposição, e tanto um quanto outro, além do verbo, contribuem com o significado total de um PV. Neste sentido, os PVs distinguem-se dos verbos preposicionados.

Por não ser um fenômeno comum na maioria das línguas, sendo mais característico às línguas germânicas, os PVs parecem desafiar estudantes de LE inglês. O maior desafio para compreender os PVs é devido a sua interpretação, uma vez que sua unidade de sentido não corresponde à soma do significado isolado de cada constituinte, pois a construção verbo-partícula forma uma única unidade lexical; tratam-se de composições semânticas em que a partícula contribui significativamente. Além disso, podem estar relacionados a mais de um significado, ou seja, o significado dos PVs pode variar, por vezes sendo mais literal e, por outras, mais figurativo/idiomático.

Apesar da pesquisa não estar relacionada diretamente ao ensino e aprendizagem de LE, a descrição do fenômeno ‘construção verbo-partícula’ facilita a compreensão dos aspectos envolvidos na prática de ensino, por exemplo, facilitando sua interpretação. Trata-se de um fenômeno que envolve diversas peculiaridades, as quais são vistas pela ótica da semântica cognitiva. Desse modo, entende-se que a

¹ Tratarei *phrasal verbs* pela sigla PVs ou PV, quando no singular.

² Usa-se aspas simples (') para expressões de destaque.

³ Opta-se pelo termo ‘partícula’ por se tratar de uma expressão neutra; a noção de partícula é variável, podendo se referir tanto a uma preposição quanto a um advérbio, ou ambos; o termo também é útil para evitar que se confunda *phrasal verbs* com verbos preposicionados.

Linguística de Corpus serve como suporte à investigação linguística, sustentada na semântica cognitiva, mais especificamente na Teoria da Metáfora Conceptual, fornecendo um ferramental fidedigno à extração de dados representativos.

Visto a complexidade da estrutura semântica das construções dos PVs, a contribuição semântica de suas partículas está relacionada a extensões metafóricas do significado. Sendo assim, destaca-se que as direções apresentadas para a extração de concordâncias no *corpus*, as quais contemplam a estrutura verbo-partícula, foram estabelecidas a partir de objetivos específicos que nortearam a pesquisa de mestrado já mencionada; são eles: (i) refletir sobre as propriedades gerais dos PVs, constituídos pelas partículas ‘*up*’ e ‘*down*’, em especial as que dizem respeito à idiomaticidade e metaforicidade; (ii) averiguar qual é a relação da idiomaticidade e metaforicidade com os PVs separáveis e não separáveis; e (iii) verificar em que medida a metáfora das partículas ocorrem e/ou se caracterizam.

Para tanto, teve-se como perguntas norteadoras: (i) quais são os PVs mais frequentes no corpus de estudo que contêm as preposições ‘*up*’ e ‘*down*’?; (ii) qual a relação da idiomaticidade e metaforicidade dos PVs?; e (iii) os PVs não separáveis são mais idiomáticos que os separáveis?. Contudo, ao artigo aqui apresentado, cabe somente a primeira pergunta (‘i’), uma vez que esta envolve todo o processo de coleta de dados e, as demais, por sua vez, dizem respeito ao estudo linguístico propriamente dito.

A resposta à pergunta ‘i’ se faz a partir de uma análise feita com base em dados autênticos, ou seja, dados provenientes de falantes nativos em discursos realísticos. Para isso, a Linguística de *Corpus* exerce um papel fundamental, pois através de um *corpus* é possível identificar concordâncias que apontam para ocorrências na língua.

A Linguística de *Corpus* como metodologia

Adota-se o a Linguística de *Corpus* como recurso metodológico, a fim de coletar os dados necessários para evidenciar o fenômeno investigado e atingir os objetivos propostos. O *corpus* de estudo escolhido, *The Corpus of Contemporary American English* (COCA), é considerado representativo e de fácil acesso; trata-se de um *corpus* que está integrado a um *software* gratuito, que não exige instalação de qualquer programa eletrônico e conta com ferramentas úteis para a extração de dados.

Considerando os objetivos que conduzem a pesquisa alvo, o presente estudo aplicado é apresentado em três momentos: (i) a Linguística de *Corpus*, (ii) o *corpus*

de estudo e (iii) os procedimentos metodológicos. No primeiro momento, faz-se uma apresentação da Linguística de *Corpus*, destacando pontos relevantes para a pesquisa. O segundo, o *corpus* de estudo, trata do *software* e *corpus* eletrônico de onde os dados serão coletados, concordâncias que contêm os PVs mais ocorrentes com as partículas ‘*up*’ e ‘*down*’, e que proverá as evidências necessárias para que a análise semântica seja realizada; faz-se sua apresentação geral e o modo como os recursos ferramentais estão disponíveis aos usuários. O terceiro momento diz respeito aos passos tomados no processo de pesquisa para a coleta e análise semântica dos dados; análise semântica que foca, em especial, a metaforicidade das partículas dos PVs. É a partir desta última etapa que, por meio da semântica cognitiva, é dirigida a análise linguística para compreender a metaforicidade das combinações (verbo + partícula).

Linguística de *Corpus*

Conforme Sardinha (2004), a Linguística de *Corpus* trata do uso de *corpora* computadorizados – coletâneas de textos, escritos ou de transcrições de fala, mantidas em arquivo de computador – ocupando-se da coleta e da exploração de *corpora*, além de dedicar-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador. A Linguística de *Corpus* como recurso metodológico para a extração de concordâncias conduz o analista a um *corpus* representativo e que possua ferramentas adequadas ao processo de coleta de dados, facilitando a busca por PVs.

No que diz respeito à representatividade, a linguagem é um sistema probabilístico, em que certos traços são mais recorrentes que outros e a probabilidade linguística está fortemente ligada à representatividade do *corpus*. De acordo com Tagnin (2011), ainda não há um consenso para a definição de ‘representatividade’, mas a autora garante que um *corpus* é representativo daquilo que se deseja estudar e que “cabe ao criador do *corpus* estabelecer os critérios que garantam essa representatividade” (TAGNIN, 2011, p. 360). De modo geral, diz-se que, quanto maior um *corpus*, mais representativo ele é, mas, conforme Sardinha (2004) é preciso questionar: representativo do quê e para quem. Neste sentido, sugere-se considerar a amostragem de um dado *corpus* a fim de responder às questões de pesquisa do usuário, ou seja, o *corpus* deve ser adequado à investigação do pesquisador. Por exemplo, para verificar a probabilidade de ocorrência de palavras raras, é adequado

ter um *corpus* com grande quantidade de palavras – tornando-o, nesse caso, representativo – pois “quanto maior a quantidade de palavras, maior a probabilidade de aparecerem palavras de baixa frequência” (SARDINHA, 2004).

Sendo assim, pode-se dizer que vários são os aspectos considerados ao se apontar um *corpus* como representativo ou não representativo. Um dos aspectos que está no pilar da formação de um *corpus* é a autenticidade, que, por consequência, acaba refletindo sobre questões que definem sua representatividade. De acordo com Sardinha (2004, p.19), a “autenticidade dos textos subentende textos escritos por falantes nativos” e em linguagem natural, ou seja, “os textos não podem ser produzidos com o propósito de serem alvo de pesquisa linguística, e não podem ter sido criados em linguagem artificial” (SARDINHA, 2004).

COCA: o *corpus* de estudo

Pensando na representatividade e adequabilidade, os motivos que levaram a escolha pelo *corpus* *The Corpus of Contemporary American English* (COCA) estão relacionados (i) ao seu livre acesso, (ii) à forma ampla e balanceada, (iii) a sua regular atualização, (iv) à possibilidade em comparar frequências entre períodos distintos⁴ (de 1990 até 2012, ano da última atualização), além da (v) facilidade de busca por palavras, frases e categorias gramaticais que sua arquitetura permite. Contudo, anterior aos motivos citados acima, o destaque dado ao *corpus* em algumas obras lidas – Tagnin (2013); Viana e Tagnin (2011); O’Keeffe e McCarthy (2010); e Sardinha (2004) –apontaram pressupostos atrativos em relação ao COCA.

O *corpus* foi criado por Mark Davies, da Brigham Young University (Utah – EUA), em 1990 e, até o momento da finalização da dissertação⁵, continha mais de 450 milhões de palavras. De acordo com Davis (2009), o COCA foi formulado para ser comparado ao British National Corpus (BNC), em termos de tipos de textos. Os textos do *corpus* estão distribuídos entre cinco variedades de gêneros: oral (20%), ficção (20%), revistas populares (20%), jornais (20%) e trabalhos acadêmicos (20%) – tornando-o um *corpus* balanceado⁶ – em que os textos contêm etiquetas

⁴ Apesar dessa ferramenta não ser utilizada na presente pesquisa, em termos de estudos futuros ela é considerada relevante.

⁵ Informação de 2015, ano em que a dissertação foi concluída.

⁶ Sardinha (2004) define ‘equilibrado/balanceado’ pelos seus componentes (gêneros, textos etc.) que são distribuídos em quantidades semelhantes, ou seja, o número de gêneros e textos se assemelham. Tagnin (2011, p. 357) aponta o balanceamento do *corpus* como “um processo pelo qual se garante que dois *corpora* sejam construídos de maneira similar quanto a origem, gênero, extensão, período de

morfossintáticas. Conforme Davis (2009), anualmente, os textos contidos em cada gênero são balanceados entre os seus subgêneros. Por exemplo, alguns dos subgêneros presentes na categoria 'revistas populares' são: *Time*, *Men's Health*, *Good Housekeeping*, *Cosmopolitan*, *Fortune*, *Christian Century*, *Sports Illustrated*, etc.

Conforme verificado no COCA, o *corpus* apresenta uma interface com outros *corpora*, inclusive o BNC, permitindo a busca por frases ou palavras exatas, coringas (*wildcards*), lemas, classes gramaticais e combinações. O *corpus* também dispõe de buscas com base semântica, possibilitando a identificação de frequências e distribuição de sinônimos, além de permitir que essas frequências sejam contrastadas em diferentes gêneros linguísticos.

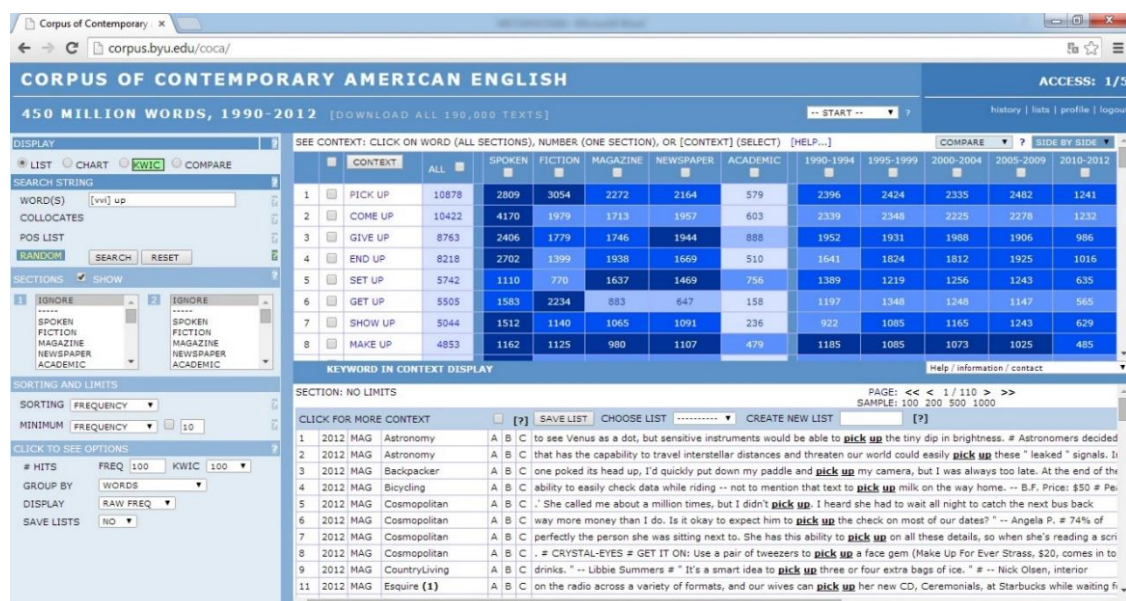
Trata-se de um *corpus* capaz de demonstrar a linguagem contemporânea, que está em uso pelos falantes nativos, bem como mudanças linguísticas ocorridas ao longo do tempo. Portanto, considerando suas características, o COCA apresenta autenticidade, representatividade e adequação para atender os objetivos da pesquisa – dissertação de mestrado.

A interface ferramental do *corpus*

O COCA possui uma interface *web* para consulta semelhante à maioria dos programas concordanceadores *on-line*. Como demonstração, a figura (1) abaixo se refere à busca por PVs constituídos pela partícula 'up' [verbo (forma infinitiva) + up]. Na parte superior e direita da tela, é possível visualizar uma lista de ocorrências/entradas colocadas por ordem de frequência, a qual também permite verificar as frequências relacionadas aos diferentes gêneros e períodos dispostos no *corpus*. Para analisar as concordâncias de uma determinada entrada da lista (e.g., 'pick up') basta clicar sobre a ocorrência desejada; dessa forma, as concordâncias serão apresentadas em KWIC (*Keyword in Context*), conforme parte inferior e direita da tela. A Figura 1 mostra as principais partes dessa interface.

produção dos textos, ou quaisquer outros critérios que sejam relevantes para a pesquisa a que se destinam”.

Figura 1: interface ferramental do *corpus*



Fonte: Corpus of contemporary American English

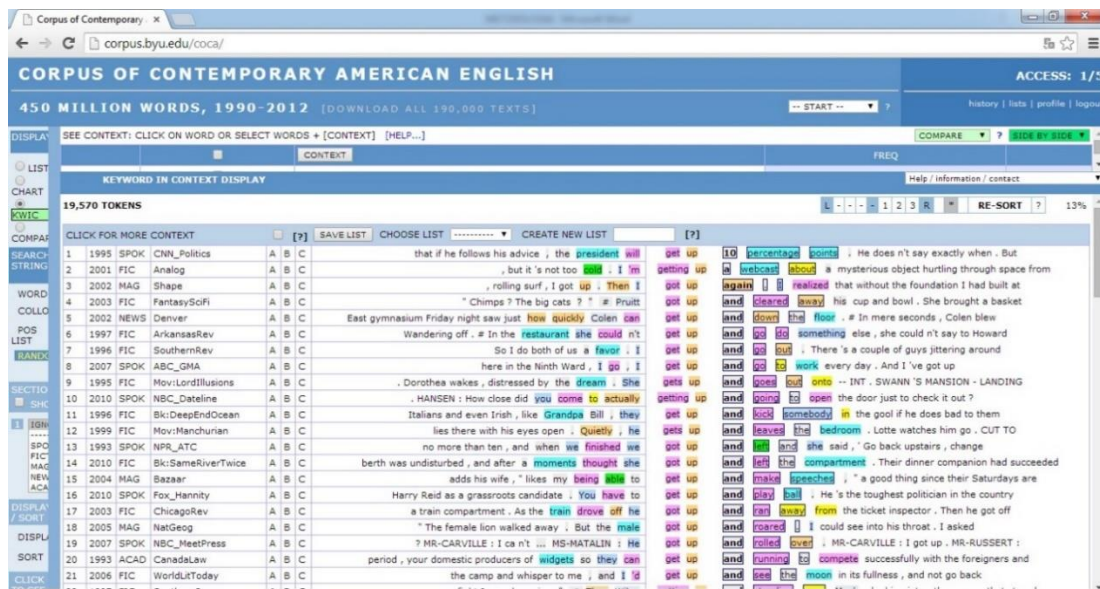
A parte esquerda da tela, que correspondente à Figura 2, diz respeito à forma de pesquisa, *'display'* – parte em que o usuário especifica os dados de busca, escolhendo o tipo de lista de frequência que deseja: (i) *list* (refere-se a uma lista de palavras ou colocações/ combinações de palavras); (ii) *chart* ('diagrama' – mostra, em forma de diagrama/gráfico, a comparação de frequências de uma palavra em diferentes gêneros ou período); (iii) *KWIC* ('palavra em contexto' – lista a palavra-chave em contextos com sua respectiva etiquetagem morfosintática; ideal para examinar classes gramaticais que seguem determinada palavra ou frase (ver Figura 3); e (iv) *compare* (utilizado para comparar diferenças de uso, permitindo identificar sinônimos ou mesmo diferenças definidas culturalmente entre duas palavras ou colocações).

Figura 2 - forma de busca



Fonte: Corpus of contemporary American English

Figura 3 - Keyword in Context (KWIC)



Fonte: Corpus of contemporary American English

Considerando que os PVs se confundem com verbos preposicionados, a ferramenta KWIC, que se refere ao formato das linhas de concordância evidenciadas em contexto, mostra-se relevante para a investigação, pois, através da etiquetagem morfossintática, pode-se distinguir PVs de ‘verbos preposicionados’ facilmente. Cada cor representa uma classe gramatical, conforme Quadro 1:

Quadro 1 - identificação de classes gramaticais em contexto

noun pronoun proper noun adjective verb adverb preposition

Fonte: Corpus of contemporary American English

Figura 4 - 'run up' – concordâncias com etiquetagem morfofossintática

DISP	SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]	COMPARE	SIDE BY SIDE
42	2003 MAG FieldStream A B C	to setting a pace, you ca n't be expected to	run up mountains but you are going to have to climb those
43	2005 FIC Callaloo A B C	the side of his mouth to Rabbit ; Fox tried to	run up on the once but he ai n't my type . He
44	2012 FIC Bk:JaneAustenMarriage A B C	'd loaned him , not to mention the debt he 'd	run up on my credit cards when his own were maxed out and
45	1995 FIC VirginiaQRev A B C	, looking at her dress again . " Some things you	run up on your sewing machine belong in the Stone Age . Why
46	1995 ACAD BioCycle A B C	sections , each sloping in opposite directions so that windrows	run up one side to crest and down the other side . Usable
47	1992 MAG Conservation A B C	a good number of lake trout in the river . Lakers	run up rivers in the fall and tend to concentrate in large pools
48	2010 NEWS Chicago A B C	up a ladder , pull hundreds of pounds of hose or	run up steps without getting winded . # For two months Tirado , Brown
49	2010 SPOK CNN_Brown A B C	has become an absolute cash cow for doctors . They have	run up the biggest bill in the country using your tax dollars .
50	2009 SPOK NPR_NewsNotes A B C	built-in deficit we 've ever had . The current year will	run up the biggest deficit we 've ever had . It will be
51	1990 FIC Mov:TotalRecall A B C	Richter and Helm , now behind him , shoot as they	run up the escalator Quaid hurls the remains of the corpse at
52	1999 FIC FantasySciFi A B C	gate , " Clare ordered , out of breath from her	run up the hill Men obeyed . She turned to her mounted
53	2007 SPOK Fox_Journal A B C	he may survive Iowa . GIGOT : So he has to	run up the middle basically in the other two decide to ...
54	2008 NEWS Chicago A B C	pass , the Bears call time out to set up a	run up the middle where Mokie is stuffed for no gain .
55	2001 SPOK ABC_GMA A B C	hand-done , from France . GIBSON : That 's going to	run up the price Ms-WALN : And this also has a detachable
56	1999 FIC FantasySciFi A B C	into the old crevasse . Carl stumbled and slipped trying to	run up the ravine straddling the widening crack . From behind he
57	1999 NEWS Houston A B C	quarter with a 44-21 lead because he did n't want to	run up the score # Jeers - To Saints coach Mike Ditka
58	1990 FIC Mov:FlightDragon A B C	of reporters , television people , photographers , etc . They	run up the steps past Peter as they head toward the entrance to
59	1992 FIC Mov:1492: Conquest of Parad... A B C	a large group of YOUNG MEN , shouting with excitement .	run up the street and brush past them . Then more people .
60	2000 FIC Storyworks A B C	shaming ! Let 's go . Storyteller 2 : The cats	run up the tree bite them . Storyteller 1 : Snowbell finds a
61	2006 FIC LiteraryRev A B C	passing cars on old Route 22 caused strips of	run up the walls to the rafters and disappear . Odd ideas flopped
62	2004 NEWS USAToday A B C	, " Dean said , Dean was " too busy to	run up the bank fan ; " Wright says in his book ,
63	2005 FIC Bk:Juv:Emperor A B C	of the long ladders to each unit . The soldiers could	run up the wall jump almost full speed , and it would
64	2010 SPOK PBS_NewsHour A B C	those families who have created businesses out of the garage ,	run up their credit cards , working evenings , years and years ,
65	2008 NEWS CSMonitor A B C	borrowers are likewise finding it hard to borrow and	run up their debt . * We have not seen a real downturn

Fonte: Corpus of contemporary American English

Para melhor compreensão: supondo um particular interesse pelo PV 'run up', como certificar-se se as concordâncias indicadas são PVs ou verbos preposicionados? Levando em conta que as partículas dos PVs exercem função adverbial, mesmo, por vezes, tratando-se de preposições, percebe-se, por exemplo, a diferença entre ambos na Figura 4, linhas 49 e 53: na linha 49, a concordância se refere a um PV e, na linha 53, a um verbo preposicionado, conforme aponta a etiqueta morfofossintática.

Quanto aos tipos de consulta, 'Search Strings', podem ser realizadas por (i) *word(s)*, (ii) *collocates* ou (iii) *pos list*. O primeiro tipo, 'word(s)', possibilita a busca por palavras ou frases e permite que esses termos sejam visualizados em formatos distintos. Para obter todas as formas sintáticas de uma palavra, a pesquisa pode ser por 'lema', bastando colocar a palavra entre colchetes [], por exemplo: os resultados da busca '[give]' são 'give (167155), given (94310), gave (88879), giving (48149), gives (39480), giveth (65) e gived (5)'; os números entre parênteses correspondem ao número de frequência. Para buscar 'sinônimos' de uma determinada palavra é necessário digitar '[=palavra]'; esse meio de consulta

não permite a busca de sinônimos por colocações. Não sabendo exatamente a palavra ou colocação de busca, a pesquisa também pode ser realizada por ‘classes gramaticais’, em que o sistema reconhece por curingas representacionais, como substantivos [n*], verbos [v*], adjetivos [j*], etc. Contudo, não é necessário memorizar esses símbolos, basta clicar no menu ‘*post list*’⁷ e selecionar a categoria gramatical que se deseja.

O segundo tipo se refere aos ‘colocados’⁸, que, nesta ferramenta, diz respeito a uma determinada palavra (escolhida pelo usuário) que ocorre na extensão da concordância entre até dez palavras antes ou depois da palavra ou frase pesquisada. Essa ferramenta permite que a pesquisa seja feita por PVs separáveis ou não separáveis. Dessa forma, ao considerar ‘*get up*’ como exemplo ilustrativo, para verificar suas devidas concordâncias, é necessário digitar ‘*get*’ no campo ‘*word*’ e ‘*up*’ em ‘*collocates*’, ou vice versa, a ordem não importa. Próximo ao menu ‘*collocates*’ há dois campos que permitem selecionar, em cada, um número – trata-se do alcance/extensão do colocado a ser pesquisado; o primeiro campo representa quantas palavras, antes da palavra ou frase pesquisada, se deseja que o colocado ocorra e o segundo campo representa quantas palavras, depois da palavra ou frase pesquisada, se deseja que a colocado ocorra. Neste sentido, para a busca ‘*word(s): get/ collocates: up*’, sugere-se usar o número três (3) no primeiro campo e o zero (0) no segundo; ‘zero’ porque não se deseja que ‘*up*’ apareça antes de ‘*get*’. Esse tipo de busca permite a verificação de ocorrências constituídas por PVs separáveis e não separáveis; entretanto, por esse ferramental, não se tem acesso às etiquetas morfosintáticas das concordâncias em contexto. Ver Figura 5:

⁷ O termo ‘*post list*’ equivale a ‘*parts of speech*’, conhecido na língua portuguesa como ‘classes/categorias gramaticais’.

⁸ Os ‘colocados’, ou ‘colocações’, também conhecidos por ‘*chunks*’ são chamados por Tagnin (2013) de ‘coligações’⁸; contudo, aqui não se faz distinção entre os termos, ambos são tratados como sinônimos.

Figura 5 - colocações – busca e resultado

The screenshot shows the COCA search interface. The search string is '[get]'. The results table shows the word 'UP' with a frequency of 30955. Below the table, a 'KEYWORD IN CONTEXT DISPLAY' section shows 10 example sentences with the word 'get' and its collocate 'up' highlighted in context.

1	CONTEXT	FREQ
1	UP	30955

KEYWORD IN CONTEXT DISPLAY

SECTION: NO LIMITS

1	2012	MAG	Astronomy	A	B	C	still be solid, but the ice would have melted to liquid and could be getting dredged up into the atmosphere slowly over time.
2	2012	MAG	Backpacker	A	B	C	the water you'll need, and gather your ingredients close at hand. Avoid getting up and down repeatedly, risking a spill. # *
3	2012	MAG	Backpacker	A	B	C	it's the perfect multitasking solution. " Downside? You'll get asked to get up and make coffee for everyone. \$450; 2 lbs. 7 oz
4	2012	MAG	ChristCentury	A	B	C	opening prayer. Kendra leads our singing, and Ryan preaches. Laura and Matt get up from their pews and walk to the table.
5	2012	MAG	ChristCentury	A	B	C	professor Diogenes Allen lecture on the sacred and secular for an hour. Afterward I got up and said to him, " Where I come
6	2012	MAG	ChristCentury	A	B	C	in the incarnation and our wholesale bemoaning of the way that the " secular " gets mixed up with the Christmas celebration
7	2012	MAG	Cosmopolitan	A	B	C	design consultancy called Red Antler, takes a different tack. " Sometimes, I get so hung up on trying to find a solution for or
8	2012	MAG	Cosmopolitan	A	B	C	brains respond to that type of structure. # Agree to a random date. Get set up with your neighbor's grandson or go with you
9	2012	MAG	Cosmopolitan	A	B	C	damage. # FACT # Sag happens naturally, thanks to gravity, and it gets ramped up after you've had a baby and begin brea
10	2012	MAG	Cosmopolitan	A	B	C	you owned the floor when it came to showing off. Fast-forward two decades, get yourself up in front of your friends, and gra

Fonte: Corpus of contemporary American English

Para verificar as linhas de concordâncias em contexto com suas devidas etiquetas morfossintáticas, basta digitar 'get up' no campo 'word(s)' para resultar em ocorrências constituídas por PVs não separáveis e na forma apresentada do verbo, neste caso, infinitiva. Para averiguar as diferentes lematizações do verbo, é necessário colocá-lo entre colchetes '[]'.

Finalmente, o terceiro tipo, 'post list', uma lista de comandos que se refere às classes gramaticais e representado por curingas. Para ativá-lo, é necessário posicionar o cursor no campo 'word(s)' ou no campo 'collocate', clicar em 'post list' e então escolher a categoria gramatical que corresponde ao curinga (em inglês, *wildcard*). Por exemplo: para buscar os verbos (forma infinitiva) que ocorrem com mais frequência com a partícula 'down', seguem-se os seguintes passos: posicionar o cursor no campo 'word(s)'; localizar e selecionar a parte do discurso que corresponde aos verbos na forma infinitiva ("verb.INF" ou "verb.INF/LEX"), na lateral do campo 'post list'; aparecerá, respectivamente "[v?i*]" ou "[vvi*]" no campo de busca por palavra; logo, basta digitar 'down'; o resultado é demonstrado na Figura 6:

Figura 6 - *post list* – busca por construções verbo-partícula

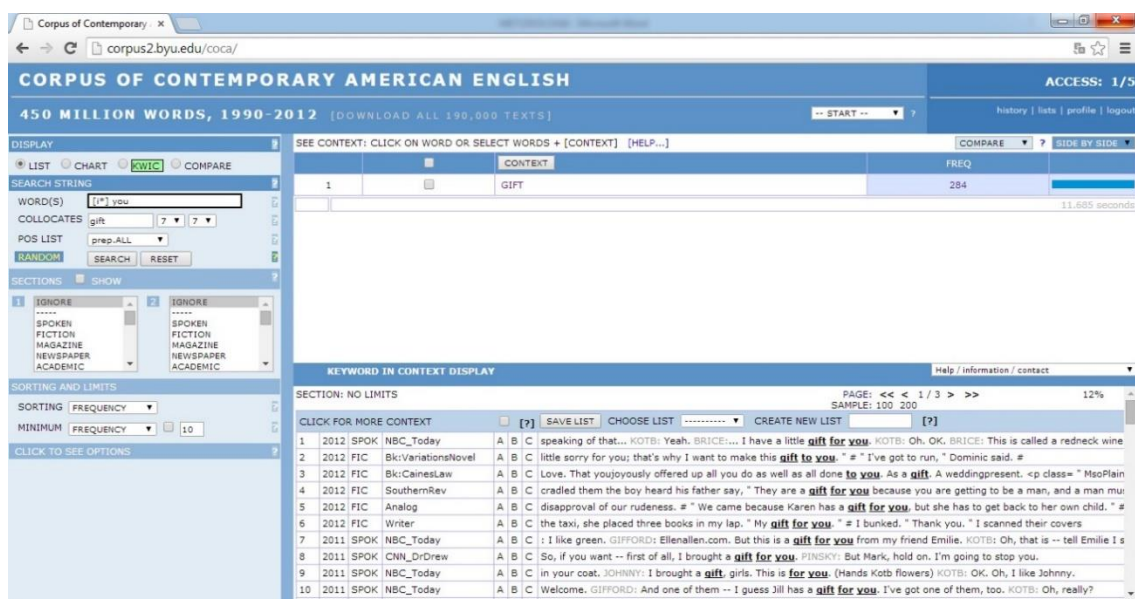


Fonte: Corpus of contemporary American English

A busca ilustrada na Figura 6 relaciona PVs não separáveis. Para efetuar a busca por construções em que o verbo e a partícula são separáveis, tem-se como uma das opções adicionar dois asteriscos entre o coringa e a partícula – os asteriscos devem estar entre espaços simples, isto é, '[vvi*] ** down'.

Um outro exemplo, muito útil na área do ensino e aprendizagem: imagine-se que se queira falar ou escrever 'um presente para você' em inglês – algo banal, mas que muitos alunos apresentam dúvidas em relação à escolha da preposição correta, 'to' ou 'for' – ver forma de busca na Figura 7:

Figura 7 - *post list* – outras variações de busca



Fonte: Corpus of contemporary American English

Procedimentos metodológicos

Os procedimentos metodológicos referem-se à coleta de dados, bem como sua análise e discussão. O processo de coleta envolve a busca por concordâncias que servirão como objeto de análise e evidências para descrever os PVs, conforme demonstrado na Figura 7. Acredita-se que as ocorrências, constituídas por PVs com as partículas ‘*up*’ e ‘*down*’, proverão as evidências necessárias para atingir os objetivos da pesquisa, bem como os números de ocorrências estimados e indicados durante as etapas do processo.

Dessa forma, de acordo com os objetivos, faz-se a identificação das ocorrências que serão analisadas em termos da semântica e metaforicidade. Para tanto, faz-se necessário seguir algumas etapas:

a) para cada partícula, verificar os 100 agrupamentos lexicais mais frequentes – 50 que precedem imediatamente a partícula e 50 que não a precedem de modo imediato;

b) para cada agrupamento encontrado, buscar e averiguar as primeiras 10 ocorrências da sequência de concordâncias através da ferramenta de busca ‘KWIC’ (que apresenta marcadores morfossintáticos), considerando todas as lematizações de cada verbo, identificar e descartar os agrupamentos lexicais que não são PVs e extrair as ocorrências selecionadas, armazenando-as em ‘.xlsx’ (Microsoft Excel).

A elaboração e cumprimento dessas duas etapas, resultaram em duas mil concordâncias, que, conforme descrito, metade corresponde a PVs separáveis e a outra metade a PVs inseparáveis. Essa divisão é feita com o objetivo de averiguar se os PVs não separáveis são realmente mais idiomáticos do que os separáveis, já que assim afirmou Rudzka-Ostyn (2003).

Fase I: *up*

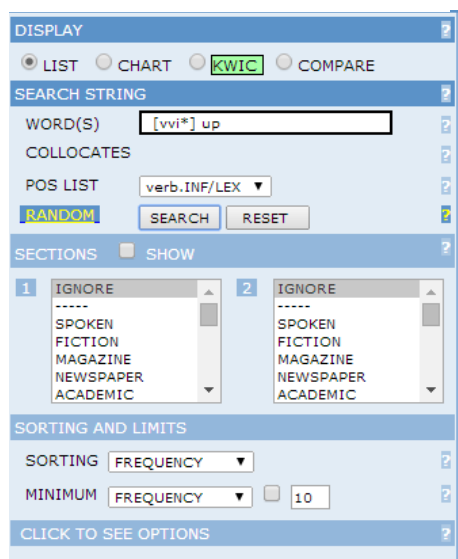
Esta seção considera as duas etapas (a-b) seguidas no processo de coleta de dados, mais precisamente, tratando de PVs com partícula '*up*'.

Fase I (up): etapa 'a'

Primeiramente, verificam-se os 50 agrupamentos lexicais não separáveis mais frequentes. Como critério para verificar uma ampla variedade de PVs, faz-se a busca por construções com verbos na forma infinitiva e a partícula '*up*' [verbo (infinitivo) + partícula (*up*)], ou seja, neste primeiro momento, desconsideram-se os PVs constituídos por verbos flexionados, evitando a repetição de um mesmo tipo de PV.

A partir da Figura 8 e Quadro 1 é possível verificar a forma de busca e o resultado:

Figura 8 - forma de busca



Fonte: Corpus of contemporary American English

Quadro 1 - lista de frequência dos PVs, não separáveis, mais comuns

Linhas	PVs	Frequência	Linhas	PVs	Frequência
1	PICK UP	10923	26	HOLD UP	1423
2	COME UP	10506	27	ADD UP	1371
3	GIVE UP	8830	28	SIGN UP	1366
4	END UP	8260	29	BLOW UP	1290
5	SET UP	5769	30	BACK UP	1256
6	GET UP	5594	31	TURN UP	1165
7	SHOW UP	5097	32	FOLLOW UP	1155
8	MAKE UP	4917	33	COVER UP	1067
9	STAND UP	4720	34	SHUT UP	1060
10	KEEP UP	4605	35	SIT UP	1027
11	GO UP	4204	36	SPEED UP	1014
12	WAKE UP	3454	37	LINE UP	865
13	GROW UP	3305	38	MOVE UP	850
14	PUT UP	3104	39	SPEAK UP	825
15	CLEAN UP	3056	40	HANG UP	760
16	TAKE UP	2975	41	STAY UP	757
17	CATCH UP	2953	42	SHORE UP	737
18	OPEN UP	2675	43	FILL UP	733
19	LOOK UP	2442	44	WARM UP	720
20	WIND UP	1903	45	PULL UP	716
21	LIVE UP	1784	46	RISE UP	682
22	BUILD UP	1732	47	THROW UP	681
23	BRING UP	1602	48	CALL UP	654
24	BREAK UP	1576	49	CLEAR UP	606
25	STEP UP	1529	50	DRESS UP	601

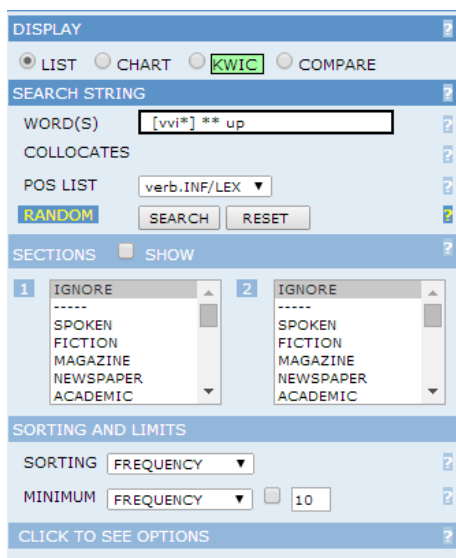
Fonte: elaborado pela autora a partir dos resultados obtidos no COCA

Posteriormente, verificam-se os 50 agrupamentos lexicais separáveis mais frequentes, considerando os verbos na forma infinitiva, e a partícula 'up' [verbo (infinitivo) + objeto + partícula (up)]. Neste caso, pode-se observar que os

PVs se repentem, pois a ferramenta de busca considera a variação dos objetos nas ocorrências.

Pela Figura 9 e Quadro 2 é possível verificar a forma de busca e o resultado:

Figura 9 - forma de busca



Fonte: Corpus of contemporary American English

Quadro 2 - Lista de frequência das coligações PVs, separáveis, mais comuns

Linhas	PVs	Frequência	Linhas	PVs	Frequência
1	PICK IT UP	1058	26	SHAKE THINGS UP	133
2	GIVE IT UP	543	27	BREAK IT UP	129
3	MAKE IT UP	499	28	GET IT UP	129
4	BRING IT UP	472	29	WAKE HIM UP	129
5	PICK HIM UP	448	30	BRING THAT UP	122
6	PICK THEM UP	428	31	GIVE THAT UP	120
7	PICK HER UP	410	32	SCREW IT UP	120
8	GET CAUGHT UP	381	33	WAKE ME UP	120
9	PICK YOU UP	380	34	SET YOU UP	109
10	PICK ME UP	352	35	PICK US UP	106
11	CLEAN IT UP	287	36	COME BACK UP	101
12	BACK IT UP	246	37	BRING THEM UP	99
13	SET IT UP	215	38	GIVE THEM UP	99
14	COVER IT UP	211	39	HELP CLEAN UP	99
15	LOOK IT UP	210	40	TAKE YOU UP	98
16	PUT IT UP	200	41	MESS IT UP	95
17	TAKE IT UP	197	42	WAKE HER UP	95
18	KEEP IT UP	191	43	BRING THIS UP	94
19	OPEN IT UP	166	44	KEEP YOU UP	94
20	WRAP IT UP	158	45	SUM IT UP	93
21	MIX IT UP	150	46	BRING US UP	91
22	GET BACK UP	148	47	LEAVE IT UP	90
23	GO BACK UP	144	48	HOLD IT UP	87
24	BRING YOU UP	143	49	GIVE HIM UP	86
25	WAKE YOU UP	141	50	BACK THEM UP	84

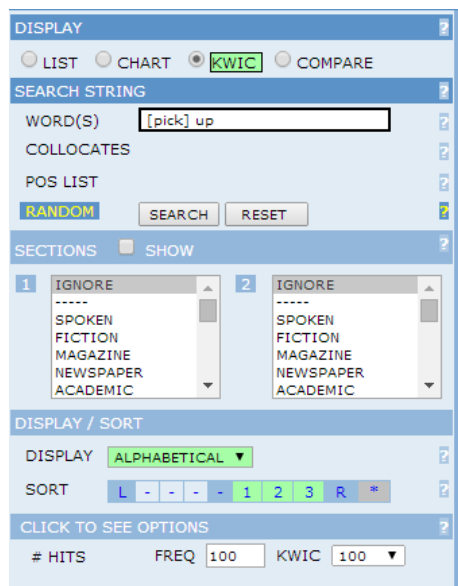
Fonte: Elaborado pela autora a partir dos resultados obtidos no COCA

Fase I (up): etapa 'b'

Considerando todas as formas de lematização dos verbos, nesta etapa, para cada agrupamento encontrado na etapa 'a', buscam-se e averiguam-se as dez primeiras ocorrências da sequência de concordâncias através da ferramenta KWIC, identificam-se e descartam-se agrupamentos lexicais que não se tratam de PVs e, então, extraem-se as ocorrências selecionadas, transferindo-as para uma planilha *Excel*.

Sendo assim, segue-se a busca por ocorrências, que estão apresentadas no Quadro 1 e Quadro 2. A subsequente Figura 10 demonstra a forma de busca para o PV mais frequente entre os não separáveis:

Figura 10 - forma de busca



Fonte: Corpus of contemporary American English

Os resultados da busca por '[pick] up', bem como dos dois próximos PVs mais frequentes estão apresentados no Quadro 3 – elaborado em *Excel*, trata da ilustração dos três primeiros PVs não separáveis mais frequentes. Para a dissertação em questão, entende-se que os três primeiros PVs foram os mais representativos.

Quadro 3 - as dez primeiras linhas de concordâncias dos PVs, não separáveis,

'pick up', 'come up' e 'give up'

	Ano	Gênero	Subgênero	Concordâncias: PVs não separáveis			
1. pick up	1	1993	MAG	MotherEarth	our home ? " On August 17 , 1992 , I	picked up	Hurricane Andrew 's image from a Russian satellite . I tracked
	2	1992	MAG	MotherJones	to death . " # Back in his office , Zelman	picked up	his messages . More congressional aides had called from
	3	2003	MAG	USCatholic	" # But are n't you curious ? " Barbara	picked up	a letter she 'd set aside . " Listen -- this
	4	2005	NEWS	Denver	his six-week rotation in psychiatry , for example . " they	picked up	all the patients here and moved them to Fitzsimons , "
	5	1994	FIC	Commentary	be alarmed ; he often dozed off after eating . She	picked up	the tray , put it on the night table , then
	6	1996	FIC	LiteraryRev	heard the invaders ? I had forgotten to listen . Lulwa	picked up	the knife , started chopping again . The onion fell away
	7	1991	FIC	KenyonRev	that evening and haggled the price down for us . She	picked up	our mail and delivered it herself , she got free medicine
	8	1994	FIC	Sagewoman	goddess , and I slept . # In the morning I	picked up	my astrology books and began to study . And over every
	9	2003	FIC	FantasySciFi	" Oh , are we there already ? " # Kelly	picked up	his book bag while the Navatar opened the door for him
	10	2007	FIC	FantasySciFi	and strength only carried him for a few minutes . He	picked up	the football and stepped outside , flinging it down the street
2. come up	1	1992	SPOK	CNN_Politics	attack on his character , which it was n't . He	came up	to me and tried to have a photo opportunity . I
	2	1999	FIC	FantasySciFi	it soon , whenever you are ready . " My hands	came up	to my chest and touched the tiny breasts there . It
	3	2005	MAG	AmSpect	Lake , basking in the good aura , and a couple	came up	to me a few minutes ago . They were a good-looking
	4	1999	SPOK	Ind_NewsForum	think that David Stern has moved off that number . He	came up	to 52 percent , and the players union was actually at
	5	1995	SPOK	NPR_ATC	here . Not surprisingly , the issue of moving the trial	came up	. Defense attorneys for Timothy McVeigh and Terry Nichols say
	6	2012	FIC	Analog	" while I was talking to my wife , and those	came up	. " # Susan loomed in . There were several about
	7	2004	SPOK	NPR_FreshAir	. GROSS : Can I ask you about the character you	came up	with of Fred Garvin : Make Prostitute ? Mr-AYKROYD : The
	8	2008	NEWS	Atlanta	in 2008 for the Dodgers and Padres . # " I	came up	here just to say thank you . " Maddux said to
	9	2004	FIC	BkConquistador	, then so be it . The SWAT troopers ' heads	came up	: something was going on , and they were getting the
	10	1997	MAG	Ebony	about it does n't help the work . When that question	came up	, I was like , Here we go . Let 's
3. give up	1	2005	MAG	Astronomy	I tried repeatedly to find it some years ago and almost	gave up	. The key turned out to be an OIII filter lent
	2	2002	NEWS	SanFranChron	. And again . Seven times . # Finally , I	gave up	and tried something else : Replacing the Ethernet cable that
	3	2000	NEWS	USAToday	marriages can compound the complications of caregiving . She	gave up	her career as a nurse and moved 900 miles with her
	4	1997	FIC	LiteraryRev	renounced writing , or at least publication . # When Prospero	gave up	his magic he left his island for the political and social
	5	2009	FIC	AntiochRev	there were three other apartments . But after a while she	gave up	sweeping and cleaning because a few hours later the dust and
	6	2001	MAG	Fortune	fell so hard last year . Sclavos disagrees : " We	gave up	40% of the company to get an asset that represented 60%
	7	2000	MAG	SportsIll	Kelly quit smoking . Then he quit drinking . Later he	gave up	his unofficial role as social secretary to his Heisman
	8	1997	MAG	AmSpect	far too early to tell whether Ireland will consider what it	gave up	worth what it gets . That 's the subject of Mary
	9	2000	NEWS	AssocPress	made his second rehab start Wednesday for Class A Huntsville and	gave up	three runs , four hits and two walks in 6 2-3
	10	2003	ACAD	AmerScholar	annexing Canada that it had cherished since 1812 . Eventually I	gave up	trying to reassure Canadians that this picture was wily

Fonte: Elaborado pela autora a partir dos resultados do COCA

Fase II: *down*

Nesta fase, respeitam-se e seguem-se as mesmas etapas apresentadas na seção de procedimentos metodológicos e demonstradas no processo de coleta de concordâncias que apresentam PVs com a partícula *'up'*. Entretanto, o foco da *'Fase II'* está em PVs constituídos pela partícula *'down'*.

Desse modo, assim como a etapa *'a'*, verificam-se os 50 agrupamentos lexicais relativos aos PVs não separáveis mais frequentes e, também, os 50 PVs separáveis mais ocorrentes. Como na fase anterior, a busca se dá por colocações constituídas por verbos na forma do infinitivo. Os resultados de busca podem ser visualizados no Quadro 4 e Quadro 5.

Quadro 4 - lista de frequência das coligações PVs, não separáveis, mais comuns

Linhas	PVs	Frequência	Linhas	PVs	Frequência
1	GO DOWN	5044	26	FALL DOWN	424
2	SIT DOWN	4625	27	TEAR DOWN	397
3	COME DOWN	3496	28	DRIVE DOWN	394
4	SLOW DOWN	2281	29	CLOSE DOWN	379
5	BREAK DOWN	1877	30	SHOOT DOWN	348
6	SHUT DOWN	1841	31	PAY DOWN	323
7	GET DOWN	1614	32	COOL DOWN	321
8	STEP DOWN	1149	33	KNOCK DOWN	305
9	BRING DOWN	1104	34	HUNT DOWN	297
10	CUT DOWN	1025	35	TONE DOWN	220
11	LIE DOWN	1002	36	SLIDE DOWN	211
12	TRACK DOWN	994	37	BURN DOWN	210
13	LOOK DOWN	930	38	MOVE DOWN	205
14	CRACK DOWN	856	39	NAIL DOWN	200
15	SETTLE DOWN	818	40	STRIKE DOWN	200
16	CALM DOWN	817	41	ROLL DOWN	192
17	PUT DOWN	797	42	WIND DOWN	190
18	WALK DOWN	715	43	PULL DOWN	178
19	LAY DOWN	674	44	CLIMB DOWN	177
20	WRITE DOWN	660	45	PLAY DOWN	169
21	TURN DOWN	611	46	DIE DOWN	167
22	TAKE DOWN	556	47	TRICKLE DOWN	167
23	PIN DOWN	448	48	SET DOWN	163
24	RUN DOWN	447	49	REACH DOWN	153
25	HOLD DOWN	435	50	STAND DOWN	151

Fonte: elaborado pela autora a partir dos resultados obtidos no COCA

Quadro 5 - lista de frequência das coligações PVs, separáveis, mais comuns

Linhas	PVs	Frequência	Linhas	PVs	Frequência
1	PUT IT DOWN	195	26	LET THEM DOWN	71
2	WRITE IT DOWN	191	27	SLOW YOU DOWN	71
3	BREAK IT DOWN	169	28	SLOW THINGS DOWN	68
4	SHUT IT DOWN	162	29	TAKE THEM DOWN	68
5	BRING IT DOWN	155	30	TONE IT DOWN	68
6	SLOW IT DOWN	149	31	SHUT THEM DOWN	67
7	TURN IT DOWN	135	32	LET ME DOWN	62
8	CALM HIM DOWN	125	33	PUT HIM DOWN	61
9	TAKE IT DOWN	125	34	BRING THEM DOWN	57
10	TRACK HIM DOWN	124	35	COME CRASHING DOWN	54
11	LET YOU DOWN	113	36	CALM THEM DOWN	53
12	GET IT DOWN	107	37	WRITE THAT DOWN	53
13	KEEP COSTS DOWN	107	38	BRING YOU DOWN	52
14	BRING HIM DOWN	106	39	PUT HER DOWN	51
15	GO BACK DOWN	102	40	GET BACK DOWN	49
16	TEAR IT DOWN	100	41	MAKE IT DOWN	49
17	TAKE YOU DOWN	97	42	WRITE THEM DOWN	49
18	GET BOGGED DOWN	95	43	PUT YOU DOWN	47
19	TRACK THEM DOWN	90	44	SLOW HER DOWN	47
20	SLOW HIM DOWN	87	45	CLOSE IT DOWN	46
21	SLOW THEM DOWN	84	46	CUT IT DOWN	46
22	CALM HER DOWN	83	47	GET HIM DOWN	46
23	COME BACK DOWN	83	48	KEEP IT DOWN	46
24	TAKE HIM DOWN	81	49	KNOCK IT DOWN	46
25	GET YOU DOWN	73	50	LET HIM DOWN	46

Fonte: elaborado pela autora a partir dos resultados obtidos no COCA

Com relação à etapa 'b', faz-se a busca pelas dez primeiras ocorrências mais frequentes, tanto para os PVs separáveis quanto aos não separáveis, averiguando-as por meio da ferramenta KWIC, identificando e descartando agrupamentos lexicais que não se comportam como PVs, se houver. Após, extraem-se as concordâncias, armazenando-as em uma planilha *Excel*. O resultado (parcial) da busca pode ser verificado no Quadro 6 e Quadro 7.

Quadro 6 - as dez primeiras linhas de concordâncias dos PVs, não separáveis, 'go down', 'sit down' e 'come down'

	Ano	Gênero Subgênero	Concordâncias: PVs não separáveis		
1. go down	1	1995 SPOK CNN_King	went up 2.1 percent and wages , salaries and benefits	went down	3 percent . There 's a widening gap , as well ,
	2	1991 SPOK CNN_Crossfire	you explain this guy at United Airlines who- their profits	went down	70 percent- Mr. BRINDISI 'That 's in one year . That
	3	1999 SPOK NBC_Dateline	icy . (Photo-of-Patrick-o) Mr-V-MARSH : Every time I 'd	go down	a hill that was , you know , fairly steep the sled
	4	1990 SPOK ABC_Jennings	say there 's no single trick to it , but	go down	a list of improvements and a car 's mileage will go up
	5	2003 NEWS Denver	# Weight-wise , she 's halfway there and has even	gone down	a shoe size . But emotionally , she 's on the road
	6	2006 NEWS Denver	school record in the decathlon . " You had to	go down	about eight stairs to get to the front door . The distance
	7	2002 MAG NaturalHist	point the tail must stop , reverse direction , and	go down	again . This reversal could be accomplished actively , through
	8	2010 ACAD PhysicalEduc	enough in the morning to realize that I need to	go down	and get it before they get there . Coordination and
	9	1996 NEWS Atlanta	his cows " getting out of the fences " and	going down	and grazing on the banks of the railroad . # " The
	10	2009 MAG MotorBoating	- almost . " Nash , I want you to	go down	and make sure everything looks good in the engine room ; check
2. sit down	1	2003 FIC Bk:SpiritsWhiteAs	, as are the rest of my teachers , " and	sat down	again . And at that blessed moment , the change-of-class
	2	1995 FIC FeministStud	machine . Drop nickels in . Nothing . I move .	Sit down	again . Jerry 's probably making out like a bandit at
	3	2007 SPOK NBC_Today	, girl . You look so good . Sit down .	Sit down	and join us . KATIE : Thank you . Thank you .
	4	1993 NEWS USAToday	material to fill air time . Then he comes in ,	sits down	and lets it happen . # He says he does n't
	5	2003 SPOK NPR_FreshAir	wills , you can see that husbands and wives very carefully	sat down	and planned and negotiated what the distribution of assets
	6	1993 ACAD AmerEthnicHis	knit . Needlework . If we had the paper , we	sat down	and read the Armenian paper . Letters , I used to
	7	1992 SPOK ABC_Special	's detrimental that we , as a generation , do n't	sit down	and read through a newspaper that we get things in bites
	8	1997 SPOK NPR_Saturday	sat down to write a funny song . I never could	sit down	and say , today I will write a funny song .
	9	2005 SPOK MSNBC_MeetPre	crisis then . This is seven years later . It 's time to	sit down	and talk about it . And I think we ought to
	10	2010 NEWS Denver	. The second thing is when people get a chance to	sit down	and talk to him , they 'll see his character ,
3. come down	1	2005 SPOK MSNBC_MeetPre	on future gasoline prices , are that the price is	coming down	a little bit ; that is , that we 'll have a
	2	2003 FIC LiteraryRev	Billy kicked him once in the stomach , and then	came down	across his face with the stick . Janet said she thought Billy
	3	2007 FIC Mov:EasternPromi	's voice ... STEPAN ... like potato sacks . Kirill	came down	after me and he hit me until I was bleeding . Then
	4	2003 FIC Bk:FatOlliesBook	listen while his aide talks into the mike , then	coming down	again and making sure the podium is draped right and the sign
	5	1999 SPOK NBC_Today	at all . As a matter of fact , he	came down	and did his traditional " Woo " like he always does upon
	6	2005 SPOK CNN_Brown	people who just ignored the bureaucracy and just decided to	come down	and do what you could , even though no one -- there
	7	1993 SPOK Ind_Limbaugh	the American Heart Association . And he asked me to	come down	and drive the -- the environmentally safe beverage cart around
	8	1999 SPOK NPR_ATC	there , you know . Mr-ROBINSON : Then the admiral	came down	and explained to us what our responsibilities were .
	9	2004 SPOK ABC_GMA	It was sort of an open call for folks to	come down	and help us out and tell us how we could really make
	10	2005 SPOK CNN_SatMorn	the hills . " A few days later , residents	came down	and returned to normal life , thankful that they minded island

Fonte: elaborado pela autora a partir dos resultados obtidos no COCA

Quadro 7 - as dez primeiras linhas de concordâncias dos PVs, separáveis, 'put it down', 'write it down' e 'break it down'

	Ano	Gênero	Subgênero	Concordâncias: PVs separáveis			
1. put it down	1	2000	FIC	Bk:NoGood-byes	. She picked up every framed photograph on the piano and	put it down	again in a slightly different place, Lonie was certain
	2	1991	FIC	Bk:ColdFire	the dresser, waited to see if he felt compelled to	put it down	again, and finally slipped it into his hip pocket
	3	1996	FIC	Bk:Nursery	! You remem-ber." He raised his cup, then	put it down	again. " You thought you dreamed it."
	4	1994	FIC	ContempFic	, picks it up, brings it to my lips,	puts it down	again. # " Retsina is bitter, " I
	5	1998	MAG	Parenting	were reading the most gripping book and someone ordered you to	put it down	and go to sleep, " says Berberian. Photograph
	6	2005	MAG	Backpacker	not reading. Intimidated by your camera's tech manual?	Put it down	and hit the backyard for an hour of trial and
	7	2005	SPOK	PBS_Tavis	be able to put it down and pick it up and	put it down	and pick it up. Tavis: Yeah. One
	8	1995	FIC	Mov:Casino	I am over here. Breaks open his blueberry muffin,	puts it down	and points to Green's. # ACE # Look at yours
	9	2004	ACAD	OrthoNursing	weary of carrying around the awful burden, and she simply	put it down	and walked away from it. # She did not
	10	2000	SPOK	NPR_FreshAir	, so you mix that stuff with ashes and you can	put it down	and you can build on that sort of thing after
2. write it down	1	1999	NEWS	Chicago	. But I did n't trust that bodyguard. So I	wrote it down	a second time and gave it to Mike. I
	2	1991	FIC	SewaneeRev	pencil as if it were loaded and cocked. # "	Write it down	about that squirrely eagle. # " Bill wrote,
	3	1992	FIC	Bk:DistinguishedGe	Make sure you clarify what you say. If necessary,	write it down	and be sure everyone agrees on the content. Never
	4	2011	SPOK	PBS_NewsHour	plan that he has been talking to John Boehner about and	write it down	and campaign on it, I think, if there
	5	2009	ACAD	Writer	play around with it a little. You may begin to	write it down	and discard the page. Another idea emerges, connected
	6	2006	SPOK	NPR_Saturday	are a story in a way and they never get to	write it down	and they never get to tell it. And when
	7	2012	ACAD	PracticeNurse	there something else that would come close? # Have you	written it down	and told other people what your goal is? #
	8	1996	SPOK	CBS_Morning	I think that if you sample something, take notes,	write it down	and -- and remember it. Ask your wine merchant
	9	2004	SPOK	CNN_King	Tennessee. They were doing a gag with me and I	wrote it down	as Kanoxville (ph). The boys in the
	10	1996	FIC	FantasySciFi	and got out some paper and a pen. " Please	write it down	before you lose it. It was beautiful, Terry
3. break it down	1	1999	NEWS	SanFranChron	mush, or evenly sized pieces, you just want to	break it down	a little. # Add the rice and salt to
	2	1992	SPOK	ABC_Special	this point JENNINGS Let me - let me go through -	break it down	and ask you to go through it a bit more
	3	1998	NEWS	WashPost	and smoked whole turkey. The gist of the answer:	Break it down	and reheat it until it reaches 165 degrees. #
	4	2006	ACAD	BioScience	. The organic matter sinks to the bottom, where bacteria	break it down	and release carbon dioxide. Bacteria thrive on
	5	2009	SPOK	CNN_Money	it's too expensive. That's how you sort of	break it down	and right now the number 46 million uninsured is the
	6	1991	NEWS	WashPost	often, but it can. Or they might need to	break it down	and self-rescue if the winds get strong. # "
	7	2003	SPOK	Ind_Oprah	WINFREY: Mm-hmm. Mr-MAC: And my wife, she	broke it down	and she came -- I knew something was wrong.
	8	2010	MAG	Sunset	dark place to store your oil. Heat and light will	break it down	and turn it rancid, so next to your stove
	9	2002	MAG	MensHealth	in fish oil, but adult humans seem to have trouble	breaking it down	and using it. However, it's linked to
	10	2006	SPOK	NPR_ATCW	more liberal direction on this particular issue. And if one	breaks it down	by age, the very youngest voters tend to be

Fonte: elaborado pela autora a partir dos resultados obtidos no COCA

Direcionamento da análise dos dados

Ao considerar as duas mil concordâncias coletadas – sendo um mil referente a construções verbo-partícula constituídas por 'up' (500 separáveis e 500 não separáveis) e as outras, um mil, correspondentes às construções com a partícula 'down' (500 separáveis e 500 não separáveis) – faz-se, primeiramente, o descarte das concordâncias formadas por combinações (verbo-partícula) que não correspondem ao fenômeno investigado, PVs.

Nesse ponto, se encerra a investigação via Linguística de *Corpus*, deixando o restante do processo – investigação semântica e metafórica – por conta da teoria que sustentará a análise. Ou seja, com os dados já coletados e armazenados em Excel, tem-se todo o *corpus* para se fazer a análise e discussão dos dados.

Conclusão

O fato de se valer da Linguística de *Corpus* como recurso metodológico para realizar a extração de ocorrências constituídas por PVs colaborou significativamente para a pesquisa. Essa parte aplicada da investigação atendeu imediatamente à primeira pergunta norteadora delineada para o desenvolvimento da dissertação em questão (quais são os PVs mais frequentes no corpus de estudo que contêm as preposições 'up' e 'down'?). As outras duas

perguntas não correspondem ao procedimento metodológico adotado via Linguística de Corpus.

Considerando os dados coletados e que serão submetidos à análise, o *corpus* escolhido, COCA, atingiu, de modo geral, as expectativas. Uma vez que o uso mais frequente de PVs se dá na linguagem cotidiada dos falantes nativos de língua inglesa, o *corpus* de estudo deve ser autêntico e, neste sentido, o COCA disponibiliza, além da sua autenticidade, dados balanceados e que contemplam diferentes períodos.

Atualmente, os PVs são utilizados nos mais variados gêneros discursivos, tanto em textos orais quanto escritos. Dessa forma, o *corpus* de estudo demonstrou ser equilibrado em virtude da variedade e distribuição de gêneros e subgêneros que envolve. Levando em conta essa característica, durante as buscas por colocações, consideraram-se todas essas variedades – oral, ficção, revistas populares, jornais e trabalhos acadêmicos – e os resultados foram apresentados em linhas de concordâncias em contexto, ou seja, a concordância foi demonstrada com informações relevantes ao seu contexto de uso, como: ano, gênero e subgênero. Ainda, por se tratar de *corpus* contemporâneo, não houve motivo para preocupar-se com a seleção de ocorrências mais atuais, já que os dados se referem ao período 1990-2012 – período próximo ao da elaboração da dissertação, iniciada em 2013 e concluída em 2015. Sendo assim, o COCA apresentou dados autênticos, representativos e adequados aos objetivos da pesquisa.

No que tange à necessária distinção que se deve fazer entre um PV e um verbo preposicionado, o ferramental do COCA colaborou significativamente; a etiquetagem morfológica facilitou essa distinção. Houve uma pequena restrição quanto ao tipo de busca para se chegar aos textos etiquetados, mas isso não prejudicou o sucesso da coleta de dados.

Em termos de relevância das linhas de concordâncias verificadas, elas evidenciaram o fenômeno investigado, isto é, os PVs como construções complexas, em que tanto o verbo quanto a partícula carregam significado. As concordâncias coletas serviram aos interesses da pesquisa focal, possibilitando uma relevante discussão acerca de seus objetivos.

Referências

DAVIES, Mark. *The Corpus of Contemporary American English: 450 million words, 1990-2012*. Disponível em: < <http://corpus.byu.edu/coca/>>. Acesso em: 14 fev. 2020.

DAVIES, Mark. The 385+ million word Corpus of Contemporary American English (1990-2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, v.14:2, p.159-190, 2009. Disponível em: <<https://benjamins.com/#catalog/journals/ijcl.14.2.02dav/details>>. Acesso em: 15 fev. 2020.

O'KEEFFE, Anne; MCCARTHY, Michael. *The Routledge Handbook of Corpus Linguistics*. 1.ed. New York: Routledge, 2010.

PIEROZAN, Samanta Kélly Menoncin. *A metaforicidade dos phrasal verbs constituídos por up e down: uma investigação sob a ótica da semântica cognitiva*. 2015.112 f. Dissertação (Mestrado em Linguística Aplicada) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Linguística Aplicada, São Leopoldo, 2015. Disponível em: <http://www.repositorio.jesuita.org.br/bitstream/handle/UNISINOS/4960/Samanta+K%E9lly+Menoncin+Pierozan_.pdf;jsessionid=2E606BDE3A07F051F75CCBAEE675F295?sequence=1>. Acesso em 15: Fev. 2020.

RUDZKA-OSTYN, Brygida. *Word Power: Phrasal Verbs and Compounds*. Berlin: Mouton de Gruyter, 2003.

TAGNIN, Stella E.O.; VIANA, Vander (Org.). *Corpora no Ensino de Línguas Estrangeiras*. São Paulo: HUB, 2011.

TAGNIN, Stella E. O. Glossário de Linguística de Corpus. In.: TAGNIN, Stella E.O.; VIANA, Vander (Org.). *Corpora no Ensino de Línguas Estrangeiras*. São Paulo: HUB, 2011.

TAGNIN, Stella E. O. *O jeito que a gente diz*. São Paulo: DISAL, 2013.

Recebido em: 25 fev. 2020

Aceito em: 18 abr. 2020