

Construções binominais quantitativas em perspectiva distintiva: uma análise colostrucional

Quantifying binomial constructions from a distinctive perspective: a colostrucional analysis

Karen Sampaio Braga Alonso^{*}
karensampaio@hotmail.com
Universidade Federal do Rio de Janeiro

Diego Leite de Oliveira^{**}
diegooliveira@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Nuciene Caroline Amphilóphio Fumaux^{***}
carol.fumaux@gmail.com
Universidade Federal do Rio de Janeiro

RESUMO: Este artigo investiga construções binominais quantitativas em uma perspectiva construcionista baseada no uso. Para isso, desenvolve, sobre os dados utilizados para a investigação, o método de análise colexêmica distintiva, um tipo de medida de associação entre lexema e construção, que compara duas ou mais construções alternantes. Nesse sentido, compreende-se que as construções envolvidas na análise apresentam certo grau de equivalência funcional, na medida em que expressam grande quantidade, porém apresentando diferenças sutis em termos de significado, explicitadas pelas preferências colocacionais de cada construção específica. O ponto central do artigo reside na reflexão sobre em que medida uma análise colostrucional de caráter distintivo pode contribuir para o entendimento mais profundo da organização do conhecimento linguístico do falante na forma de uma rede de construções inter-relacionadas.

PALAVRAS-CHAVE: Gramática de Construções Baseada no Uso. Construções Binominais Quantitativas. Análise Colexêmica distintiva. Um monte de SN. Uma enxurrada de SN.

ABSTRACT: This paper investigates quantifying binomial constructions from a usage-based constructionist perspective. For this purpose, it develops, on the data used for the investigation, a distinctive collexeme analysis, a type of measure of

^{*} Professora do Departamento de Linguística e Filologia da Faculdade de Letras da UFRJ e dos Programas de Pós-Graduação em Linguística da UFRJ e PROFLETRAS – UFRJ.

^{**} Professor do Departamento de Letras Orientais e Eslavas da Faculdade de Letras da UFRJ e do Programa de Pós-graduação em Linguística da UFRJ.

^{***} Doutoranda do Programa de Pós-graduação em Linguística da UFRJ.

association between lexemes and constructions, which compares two or more alternating constructions. In this sense, it is understood that the constructions involved in the analysis present a certain degree of functional equivalence, as they express a large amount, but having subtle differences in terms of meaning, explained by the collocational preferences of each specific construction. The central point of the paper is to reflect on the extent to which a distinctive colostrucional analysis can contribute to a deeper understanding of the organization of speaker's linguistic knowledge in the form of a network of interrelated constructions.

KEYWORDS: Usage-Based Construction Grammar. Quantifying Binominal Constructions. Distinctive Collexeme Analysis. Um monte de NP. Uma Enxurrada de NP.

Introdução

No decorrer dos últimos anos, diversos métodos de análise utilizados em Gramática de Construções Baseada no Uso (GCBU) têm se revelado bastante produtivos na descrição e explicação de construções das línguas naturais (cf. GOLDBERG, 2019; DIESSEL, 2015; PEREK, 2015). Nessa seara de métodos de investigação das construções linguísticas, interessa-nos, no presente artigo, a exploração de um conjunto de métodos conhecido na literatura como análise colostrucional (GRIES; STEFANOWISTCH, 2003, 2004).

O termo “colostrucional” consiste em uma espécie de neologismo, que envolve a semântica evocada pelas noções de lexema, colexema e construção. O conjunto de métodos em pauta considera o fato de que os lexemas observados em uma construção não ocorrem de modo aleatório, mas sim sistemático, sendo possível falar em preferências colocacionais dos lexemas (cf. HILPERT, 2014). Essa concepção fundamenta-se na ideia de que a identificação das preferências de ocorrência dos lexemas que podem preencher os *slots* de uma construção contribui para a descrição de aspectos sutis das construções sob análise e que, de outra forma, talvez não fossem observados pelo pesquisador. Assim, a metodologia em pauta revela-se interessante não só, mas também, quando se busca comparar construções que atuam em um mesmo domínio funcional e que, no âmbito do princípio da não sinonímia, espera-se apresentarem alguma diferença, seja do ponto de vista semântico, seja do ponto de vista pragmático (cf. GOLDBERG, 1995).

Este trabalho busca aplicar uma análise colostrucional a um conjunto de construções abarcadas pelo rótulo de construções binominais quantitativas do português brasileiro, que, em um nível de abstração mais esquemático, podem ser

representadas como “UM(A) N de SN”¹. Neste trabalho especificamente, são observados os padrões ‘UM MONTE de SN’ e ‘UMA ENXURRADA DE SN’ para expressar a semântica de grande quantidade. O objetivo central do presente artigo é contrastar os padrões construcionais mencionados, a partir do procedimento metodológico de análise colexêmica distintiva e refletir sobre em que medida o método em questão pode contribuir para o entendimento da organização das construções no *constructicon*, como o conjunto de conhecimentos que o falante tem de sua língua.

O artigo se divide da seguinte forma: primeiramente apresentaremos o arcabouço teórico que fundamenta esta pesquisa, para, em seguida, fornecermos uma breve descrição das construções investigadas, após o que delineamos a metodologia utilizada na análise empreendida no presente trabalho, para, finalmente, discutirmos os resultados a que chegamos através da análise construcional. Algumas considerações encerram o artigo.

1 Gramática de Construções Baseada no Uso

Com a virada construcionista nos estudos linguísticos (cf. FILLMORE; KAY; O’CONNOR, 1988; LANGACKER, 1987; LAKOFF, 1987; GOLDBERG, 1995, 2006, 2019; TRAUGOTT; TROUSDALE, 2013; HILPERT, 2014 e outros), a fronteira bem definida entre sintaxe (ou gramática), como o domínio das regularidades, e o léxico, como o domínio das idiossincrasias, deixa de existir como tal, em prol da integração entre os dois domínios como parte do conhecimento internalizado do falante sobre o funcionamento de sua língua. Em vez de se compreender o sistema linguístico como um conjunto de regras gramaticais, aplicáveis a um arcabouço lexical organizado em lista, a língua é compreendida como um inventário altamente estruturado na forma de uma rede de unidades simbólicas convencionalizadas, que pareiam, por um lado, forma linguística e, por outro, aspectos do significado, como propriedades semânticas, pragmáticas e discursivo-funcionais (cf. GOLDBERG, 1995, 2006, 2019; CROFT, 2001; HILPERT, 2014). Nesse sentido, sintaxe e léxico recebem um tratamento homogêneo, sendo ambos interpretados à luz do conceito de construção, considerada a unidade básica de análise gramatical (cf. GOLDBERG, 2006), ainda

¹ A notação UM(A) N DE SN é utilizada para descrever um padrão geral de construções binominais quantitativas e será mais bem explicada ao longo do artigo.

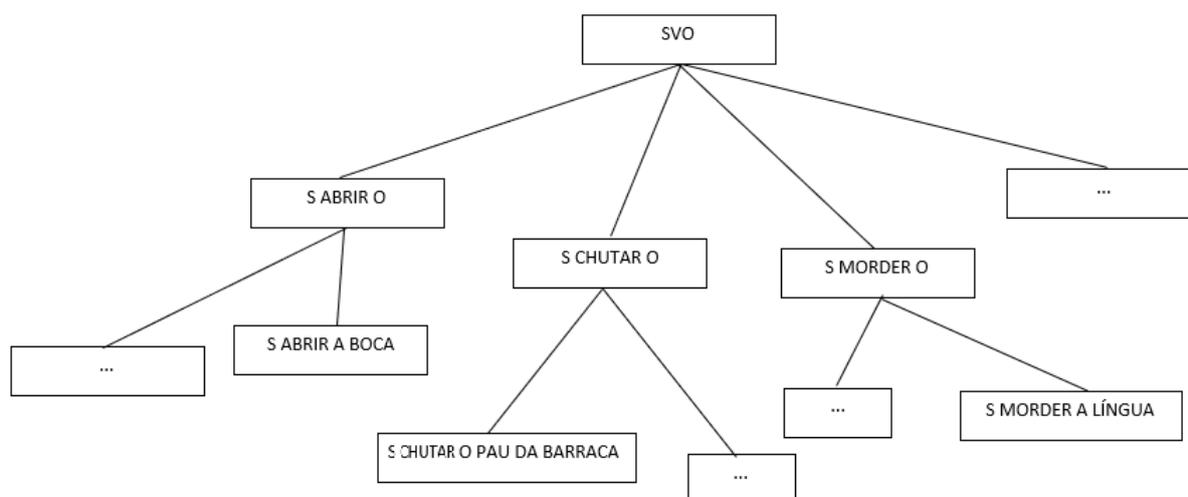
que as construções apresentem graus distintos de esquematicidade/especificidade, produtividade e composicionalidade (TRAUGOTT; TROUSDALE, 2013).

Como já dito, em uma concepção construcionista, as construções se organizam na forma de uma rede complexa de relações, nos planos taxonômico, associativo, sintático e lexical². Dos planos elencados, os que nos interessam de modo mais fundamental para o presente trabalho são os planos taxonômico e lexical, que passamos a descrever nas linhas que seguem.

Do ponto de vista taxonômico, as construções se relacionam entre si em níveis de maior ou menor especificidade ou abstração. Em uma perspectiva baseada no uso, a língua emerge a partir da experiência e, portanto, a partir do contato do falante com instâncias concretas de uso, mais especificadas e envolvidas em um contexto discursivo específico (BYBEE, 2010; 2013). Com o passar do tempo e a partir do contato do usuário da língua com instâncias diversificadas de um determinado padrão, tais instâncias de uso vão sendo comparadas e categorizadas, a ponto de padrões correlacionados serem construídos como representações mais abstratas, estas sim denominadas construções. Tais representações abstratas, à medida que os falantes vão se deparando com mais instâncias de uso da língua, podem assumir maior grau de abstração e generalização, de modo que nós mais abstratos podem ou não emergir na rede, representando o conjunto de conhecimentos que o falante tem de sua língua. Se pensarmos, por exemplo, em uma construção transitiva do tipo SVO do português brasileiro, talvez seja possível sugerir uma representação em rede com a organização taxonômica ilustrada abaixo:

² Para maior compreensão sobre as relações apresentadas, conferir o trabalho de Diessel (2015).

Figura 1. Proposta de organização da rede da construção transitiva SVO do português brasileiro



Na figura anterior, temos a representação das relações entre as construções transitivas em termos taxonômicos. Cada nó especificado refere-se a uma construção da língua, organizada de acordo com seu grau de especificidade (ou esquematicidade) e composicionalidade. No nível mais baixo da rede temos nós mais especificados e menos composicionais, em que o significado da construção como um todo não é refletido como o somatório de suas partes. Num nível mais alto, temos construções mais esquemáticas, com menor grau de especificidade e maior grau de composicionalidade e produtividade. Por fim, no nível mais abstrato, temos a construção transitiva SVO, totalmente esquemática, e com alto grau de produtividade em termos de extensibilidade (cf. BARDDAL, 2008).

Como já dito anteriormente, a distinção clássica entre sintaxe e léxico é abolida em uma perspectiva construcionista, de modo que as construções passam a ser consideradas as unidades básicas de análise gramatical. A diferença entre as construções, portanto, se dá em termos de especificidade fonológica, na medida em que é possível ter construções totalmente especificadas ou substantivas, como, por exemplo, em [CASA] ou totalmente não especificadas do ponto de vista fonológico e, portanto, mais esquemáticas, como, por exemplo, em [S V O], já exemplificada anteriormente. Construções substantivas e esquemáticas diferem entre si em termos de produtividade, no que se refere a sua extensibilidade, isto é, sua capacidade de atrair itens diversificados para preencherem os *slots* não fonologicamente especificados. Nesse caso, os itens a participarem de tais *slots* não preenchidos são

os lexemas que, por sua própria natureza, consistem em construções mais especificadas fonologicamente.

Para ilustrar a natureza das relações lexicais, ou seja, as relações que se dão entre os lexemas – construções mais especificadas e geralmente atômicas (cf. TRAUGOTT; TROUSDALE, 2013) – e as construções mais esquemáticas, tomemos a construção ditransitiva do português, tal como descrita por Furtado da Cunha (2017). A autora define a ditransitiva “como uma construção que consiste de um verbo ditransitivo, um argumento agente (A), um argumento recipiente (R) e um argumento paciente (P)” (FURTADO DA CUNHA, 2017, p. 116), em que um “agente faz com que o recipiente receba o paciente” (p. 117). De acordo com a própria autora, o significado de transferência expresso pela construção está associado a verbos de oferecimento, cujo protótipo é o verbo *dar*.

Nesse sentido, com base na GCBU (cf. DIESSEL, 2015), seria possível postular a existência de um *link* lexical forte entre a construção (ou lexema) DAR e a construção DITRANSITIVA do português brasileiro, dada a frequência com que tal lexema é recrutado para participar das instâncias da construção. Outros lexemas também recrutados pela construção são verbos como ‘presentear’, ‘entregar’, ‘passar’ etc., porém com um *link* lexical mais fraco, também por razões de frequência.

Seguindo, portanto, os princípios da GCBU, levaremos em conta, para a descrição das construções em foco neste trabalho, que: (1) os falantes têm acesso direto apenas a formas linguísticas concretas, ou seja, aos construtos; e (2) as construções, estruturas abstratas, recrutam preferencialmente nomes que se acomodam a elas por coerência semântica, o que nos leva a concluir que conhecer os lexemas mais atraídos por uma dada construção é conhecê-la.

As construções ‘UM MONTE DE SN’ e ‘UMA ENXURRADA DE SN’ fazem parte de um rol de diferentes construções binominais quantitativas abstraídas do uso e que servem, na língua portuguesa, para a multiplicação de referentes em termos de grande quantidade. Embora funcionalmente equivalentes, o uso de cada uma delas possui particularidades bastante interessantes, a partir das quais podemos fazer uma breve reflexão acerca da abstração de construções mais esquemáticas na língua.

Assim sendo, a questão que pretendemos provocar permeia os limites da metodologia de análise de dados de *corpus* para uma descrição adequada da rede

construcional. Na análise por nós empreendida, por exemplo, pudemos observar que a maioria esmagadora das construções binominais quantitativas apresenta um artigo indefinido seguido de um nome no singular, da preposição 'de' e de um outro nome no plural. Em função desse comportamento recorrente das construções binominais quantitativas no português, postulamos, que, em termos de realidade psicológica, o falante tenha armazenadas tanto as construções 'UM MONTE DE SN' e 'UMA ENXURRADA DE SN', como a construção mais abstrata 'UM(A) N de SN'.

Isso implica considerar que, embora haja dados no *corpus* em que, no lugar do artigo indefinido, observemos a ausência do determinante ou mesmo a presença de outros elementos, como, por exemplo, artigos definidos e demonstrativos, não postulamos que os falantes do português tenham armazenada uma construção ainda mais abstrata na rede, como 'SN de SN', por exemplo. Isso porque, em se considerando a metodologia adotada, não obtivemos resultados quantitativos suficientemente fortes para sugerir que o falante faça uma abstração para além daquela em que o determinante é preenchido por um artigo indefinido. A análise dos dados nos leva a avaliar, nos casos de ausência de determinantes ou presença de elementos diferentes de *um(a)*, uma motivação mais contextual do que a abstração de uma construção mais ascendente na rede.

Essa ausência do *um(a)* ou presença de elementos diferentes de *um(a)* é mais observada na construção 'UMA ENXURRADA DE SN' do que em 'UM MONTE DE SN', e isso muito provavelmente se deve ao fato de que 'UMA ENXURRADA DE SN' é uma construção mais recente na língua e apresenta maior grau de analisabilidade e composicionalidade, na comparação com 'UM MONTE DE SN'. Essa última, por sua vez, conforme já demonstrado em Fumaux (2018), foi deixando de recrutar elementos diferentes de *um* no decorrer de seu longo processo de formação, ao mesmo tempo em que se observava a preferência dessa construção por recrutar nomes cada vez menos coerentes com o sentido original de monte (formação geológica).

Análises colostrucionais, modeladas para medir a força de atração de itens e construção, podem, nos diferentes níveis taxonômicos, nos ensinar não apenas sobre o uso das construções, mas também sobre a representação mental que os falantes fazem delas. Essa relação entre os métodos de análise colostrucional e o nível taxonômico das construções será apresentada na seção 3 deste trabalho.

2 Construções binominais quantitativas do tipo *um N1 de N2*

Falantes do português têm à sua disposição um longo arsenal de construções gramaticais que servem para quantificar referentes. Construções de plural (como *Ns*), construções formadas por pronomes indefinidos seguidos de nomes (como *muito(s) N(s)*), com numerais cardinais e nome (como *Num Card N*), dentre outras, são algumas das alternativas de que eles podem lançar mão quando desejam expressar noções de quantidades de mais de um, quantidades subjetivas, como as de *pouco* ou *muito*, ou ainda precisas, como as de *cinco*, *dez*, entre outras.

As construções binominais focalizadas no presente estudo se caracterizam por ser um formativo nominal complexo constituído, preferencialmente, pela combinação de um artigo indefinido seguido de um nome (entendido como *quantificador*), da preposição *de* e de um SN (entendido como *quantificado*). Semanticamente, entende-se que é uma construção utilizada, de forma geral, no português, para quantificar (multiplicar) referentes contáveis. Por exemplo, em *um monte de pessoas*, o referente *pessoa* foi multiplicado em grande quantidade, ao se combinar sintagmaticamente o artigo indefinido *um*, o nome quantificador *monte*, a preposição *de* e o nome quantificado *pessoas*.

Além dessas propriedades gerais, construções desse tipo têm como característica o fato de que podem recrutar nomes quantificadores que não são originalmente relacionados com a ideia de quantificação. No exemplo anterior, foi possível observar o recrutamento dos colexemas *monte* e *pessoas* pela construção *um(a) N1 de N2*, em que *monte* é recrutado para preencher o *slot* do quantificador, embora originalmente refira-se ao sentido de formação geológica, ou seja, não se trata de um quantificador nato, mas de um nome que ganha esse sentido ao instanciar a referida construção quantitativa.

De fato, parece haver uma gama de itens que podem ser recrutados para preencher o primeiro *slot* da construção quantitativa 'UM(A) N1 DE N2', os quais compartilham com ela alguma afinidade semântica; no caso, algum traço a partir do qual se pode inferir a ideia de quantidade grande ou pequena. No caso de grande quantidade, podemos ter exemplos como os de *um mundo de ideias*; *um universo de possibilidades*; *uma chuva de críticas*; *uma enxurrada de e-mails*; *uma avalanche de problemas*; *uma nuvem de exemplares*; *uma montanha de livros*; *um monte de gente*; *uma pilha de meias*; *uma pá de coisas*; *um caminhão de mensagens*; *uma*

cambada de marmanjos; uma cacetada de provas; uma porrada de canetas; uma dezena de crianças; uma centena de participantes; um mar de lágrimas etc. (outros exemplos, bem como maiores detalhes dessas construções podem, ainda, ser encontrados em ALONSO, 2010 e BRODBECK, 2010).

A relação entre itens e construções vem sendo estudada na literatura da gramática de construções por diferentes autores (por exemplo, GOLDBERG, 1995; HILPERT, 2014; PEREK, 2015), e é uma preocupação constante da área. De fato, toma-se, por princípio, que nomes frequentemente recrutados por construções gramaticais fornecem pistas acerca do sentido dessas construções, uma vez que apresentariam maior coerência semântica com elas.

Seguindo esse raciocínio, entende-se que a construção quantitativa do tipo 'UM(A) N1 DE N2' vai preferencialmente recrutar nomes quantificadores clássicos ou potencialmente quantificadores (que podem gerar inferência de quantidade via, por exemplo, extensão metafórica) para o *slot* aberto notado como *N1*. Nesse mesmo sentido, no que tange ao *slot* aberto notado como *N2* e quando se trata de construções que quantificam referentes em termos de grande quantidade, como é o caso das construções abordadas neste artigo, espera-se que sejam recrutados nomes contáveis ou nomes abstratos pluralizados.

Assumindo como verdadeira a hipótese da coerência semântica da construção, também se espera que ela influencie a seleção dos nomes recrutados para os slots *N1* e *N2*. Assim, quanto maior a afinidade semântica entre o nome quantificador e o nome quantificado, mais chances eles teriam de ocorrer juntos na construção; e quanto menor a afinidade semântica entre os dois nomes, menor a chance de se apresentarem como colexemas na construção quantitativa.

A preferência por nomes contáveis ou tomados como contáveis, bem como a coerência semântica entre *N1* e *N2* foram corroboradas pelos resultados encontrados em Alonso e colegas (no prelo) e por Alonso e Fumaux (no prelo). Conforme já dito anteriormente, submetendo construções quantitativas de grande quantidade do tipo 'UM(A) N1 DE N2' (especificamente, *um monte de SN*, *uma montanha de SN*, *uma enxurrada de SN* e *uma chuva de SN*) a análises colexêmicas, as referidas pesquisas obtiveram como resultado o recrutamento preferencial de nomes que se acomodavam mais fortemente à semântica de cada uma das construções.

Ainda sobre a relação entre itens e construções, Talmy (2006) defende que uma dada sentença é formada por dois subsistemas: o gramatical e o lexical, em que a estruturação gramatical afetaria o conteúdo especificado no léxico, funcionando como, segundo o autor, um tipo de esqueleto imagético. A partir daí, o autor faz uma reflexão acerca do processo de quantificação (tomado de forma bastante geral), em que descreve as seguintes categorias gramaticais: dimensão, plexidade, estado de delimitação e estado de divisão.

No presente artigo, vamos nos deter à categoria da plexidade, uma vez que ela se aplica mais diretamente aos nossos dados. Segundo Talmy (2006), plexidade associa a ideia de quantidade ao conceito de massa (equivalente aos valores de singular e plural) – o qual nos interessa especificamente – ou ação. Sobre o conceito de massa, o autor ilustra suas ideias a partir dos seguintes exemplos:

(1) A flor se abriu.

(2) As flores se abriram.

No exemplo (1) – *A flor se abriu* – a entidade *flor* é conceptualizada como única (uniplexidade), enquanto que, no exemplo (2) – *As flores se abriram* – a entidade é copiada, multiplicada (multiplexidade). Analogamente, tomando um exemplo como *uma enxurrada de críticas*, entende-se que a construção quantitativa de grande quantidade *uma enxurrada de SN* copia a entidade *crítica*, reconfigurando-a, assim, em termos do processo de multiplexização. No caso, a forma plural – *críticas* – ilustra bem o fato de que essa construção prefere nomes contáveis ou tomados como contáveis (ainda que abstratos), coerentes com ela.

Abordando construções binominais no inglês, Zhan e Traugott (2015) oferecem uma taxonomia que procura rememorar as relações dentro da rede construcional, focando os diferentes níveis de esquematicidade em que essas construções podem ser analisadas. Essa taxonomia abarca desde níveis mais abstratos – a que as autoras chamam de esquemas – até níveis menos abstratos – denominados de microconstruções –, incluindo, ainda, os construtos, que correspondem às realizações dessas construções no uso da língua.

a. Esquemas: construções abstratas que são generalizações taxonômicas sobre diversas construções mais particulares. Ex. quantificação binominal [Ni of a Nj] ←→ [Dimensão*i* relacionada a SEM*j*].

b. Subesquemas: subgrupos de esquemas, menos abstratos, mas ainda assim construções esquemáticas. Por exemplo, entre os subesquemas dos

quantificadores binominais, estão: dimensão pequena (a bit of N) e dimensão grande (a lot of N).

c. Microconstruções: Construções do tipo individual. Membros de um subesquema pequeno: a bit/shred/jot/iota of N

d. Constructos: Instâncias de microconstruções, tokens de uso real. Ex: they are hacks without a shered of intellectual honesty.

No presente artigo, investiremos em uma análise colostrucional colexêmica distintiva, considerando as construções *um monte de SN*, *uma montanha de SN*, *uma enxurrada de SN* e *uma chuva de SN*, seguida por uma análise covariacional, a qual investiga construções binominais quantitativas, a partir de relações colexêmicas encontradas em um nível taxonômico mais alto, a saber *um N1 de N2*, em que os dois *slots* estão abertos.

Embora já tenham sido apresentados resultados de estudos anteriores acerca dessas construções, ainda não havia um trabalho em que se aplicasse sobre elas análises colostrucionais distintiva e covariacional, o que representa o ineditismo dessa pesquisa em relação ao que já foi feito anteriormente acerca dessas construções binominais quantitativas no português.

A seguir, na seção 3, passamos à descrição da análise colostrucional propriamente dita e ao delineamento do *corpus* sobre o qual essa análise foi efetuada.

3 Análise colostrucional

O conjunto de métodos intitulado análise colostrucional e proposto por Stefanowistch e Gries (2003), Gries e Stefanowistch (2004) visa a identificar as relações que se dão entre os lexemas e as construções em línguas naturais. Esse conjunto de métodos se coloca como uma contribuição interessante, na medida em que permite ao pesquisador identificar lexemas que são atraídos com mais frequência para os *slots* esquemáticos das construções, em diversas situações específicas: (i) quando o pesquisador está interessado na investigação de uma construção com grau relativo de especificidade, que apresenta apenas um *slot* esquemático, (ii) quando o pesquisador deseja contrastar construções desse tipo que, apesar de diferirem na forma, apresentam funções similares, e (iii) quando o

pesquisador deseja estudar construções mais abstratas que apresentam dois *slots* esquemáticos.

Para atender às três demandas mencionadas, são utilizados três métodos distintos, a saber: a análise colexêmica, que visa a observar o grau de atração entre lexemas e um *slot* de uma construção; a análise colexêmica distintiva, que visa a analisar o grau de atração entre lexemas e um *slot* em duas ou mais construções; e a análise colexêmica covariacional, que tem por objetivo analisar o grau de atração entre lexemas e dois *slots* de uma mesma construção, em um nível mais abstrato de análise. Passamos, nas próximas linhas, a descrever esses métodos mais especificamente.

A análise colexêmica geralmente é utilizada para identificar as preferências colocacionais dos lexemas em relação a apenas um *slot* de uma construção específica. Ela leva em consideração a frequência de ocorrência dos lexemas que ocorrem no *slot* esquemático, a sua frequência de ocorrência no *corpus* como um todo e o número de palavras do *corpus*, exibindo, com base em uma média triangulada entre esses dados, a força de atração entre o lexema e a construção. No que diz respeito ao estudo das construções binominais quantitativas, Alonso e colegas (no prelo), ao contrastarem línguas como o russo e o português brasileiro, aplicaram esse método com vistas a identificar as propriedades colocacionais da microconstrução 'UM MONTE DE SN'. Na análise em pauta, os autores puderam observar que, apesar de essa microconstrução permitir a ocorrência de um número expressivo de lexemas variados, ela atraía mais frequentemente os lexemas 'gente' e 'coisa', vocábulos que caracterizam entidades animadas e inanimadas, respectivamente, porém de cunho genérico.

A análise colexêmica distintiva, por seu turno, permite comparar duas ou mais construções e leva em consideração a frequência de ocorrência dos lexemas que ocorrem no *slot* esquemático contrastivamente entre as construções analisadas. Esse tipo de análise demonstra-se bastante profícuo para contrastar construções que se encontram no mesmo nível de abstração na rede, na mesma língua ou em línguas diferentes. Por exemplo, Alonso e colegas (no prelo) utilizaram esse método para contrastar construções binominais quantitativas do russo e do português, chegando à conclusão de que construções similares em ambas as línguas atraíam colexemas distintos, ainda que o elemento fixo na posição de N1 exibisse as mesmas propriedades lexicais. No presente trabalho, o método de análise

colexêmica distintiva será utilizado para contrastar o *slot* esquemático SN nas construções 'UM MONTE DE SN' e 'UMA ENXURRADA DE SN', no português brasileiro.

Por fim, a análise colexêmica covariacional, como já dito, permite que o pesquisador observe particularidades de construções em um nível mais abstrato da rede taxonômica, na medida em que analisa as preferências colocacionais de dois *slots* esquemáticos de uma mesma construção. Para aplicar esse método de análise ao fenômeno em pauta, é necessário subir um nível na rede e analisar a construção 'UM(A) N DE SN', identificando quais são os lexemas atraídos tanto para o *slot* 1 como para o *slot* 2. No entanto, por questões da seleção do *corpus* utilizado na pesquisa, bem como pelo interesse dos pesquisadores deste trabalho por apenas 2 microconstruções da rede de construções binominais quantitativas já mencionadas acima, o primeiro *slot* da construção 'UM(A) N DE SN' restringiu-se apenas aos lexemas *monte* e *enxurrada*, o que, em certa medida, restringe o escopo de possibilidades da análise colexêmica covariacional. Em um futuro próximo, pretendemos ampliar o escopo de investigação para um conjunto maior de lexemas para que esse tipo de análise possa ser realizado.

Para desenvolver a análise colostrucional pretendida neste artigo, recorreremos a dados extraídos do *corpus* CHAVE, que contém textos coletados dos jornais Público, de Portugal, e Folha de São Paulo, do Brasil, compilados entre os anos de 1994 a 1995. Cabe ressaltar que, para esta pesquisa, foram utilizados somente os dados do português brasileiro, correspondendo a um total de 35.699.765 palavras. Ressaltamos, além disso, que se trata da mesma metodologia contida em Alonso e Fumaux (no prelo). Após a organização dos dados em uma tabela, estes foram processados no programa de análises estatísticas R, utilizando-se o *script* fornecido por Stefan Gries (2014), para o desenvolvimento das análises colostrucionais. No *script* referido, seguimos o passo a passo para o desenvolvimento da análise colexêmica distintiva, cujos resultados serão apresentados na próxima seção.

4 Análise dos resultados

Como já mencionado na seção anterior, a análise colexêmica distintiva se difere das demais análises colexêmicas por permitir a comparação direta entre as preferências colocacionais de construções alternantes. Neste trabalho,

consideramos que as construções 'UM MONTE DE SN' e 'UMA ENXURRADA DE SN' exibem certo grau de alternância, na medida em que as duas apresentam equivalência funcional no rol das construções binominais quantitativas de grande quantidade. No entanto, com base no princípio da não sinonímia supramencionado, acreditamos que essas construções exibem especificidades semânticas, estilísticas e contextuais e, nesse sentido, uma análise colexêmica distintiva poderia nos ser útil ao apontar não somente os lexemas preferidos por cada construção para ocuparem o *slot* esquemático SN – o que já é possível observar com uma análise colexêmica simples – mas também aqueles que são repelidos pelas construções em uma perspectiva contrastiva. Ao observarmos os lexemas preferidos e repelidos por cada construção, podemos identificar as especificidades semânticas não observáveis diante de outros tipos de análise.

Antes de iniciarmos a análise dos dados, é preciso atentar para o fato de que a análise colexêmica distintiva aplicada às duas construções em foco no presente estudo está circunscrita às especificidades e à dimensão do corpus selecionado na pesquisa. Sendo assim, todos os resultados e inferências a que chegamos a partir da aplicação do método devem ser compreendidos dentro desse limite metodológico. A tabela 1 abaixo apresenta os resultados de uma análise colexêmica distintiva das construções investigadas neste trabalho. São apresentados os resultados relativos aos dez lexemas mais frequentes e mais relevantes que ocorrem no slot esquemático das construções em pauta:

Tabela 1 Análise colexêmica distintiva das construções UM MONTE DE SN e UMA ENXURRADA DE SN

Coll_Word	Enxurrada	Monte	exp_Enxurrada	exp_Monte	pbin_Enxurrada	pbin_Monte
dólares	21	1	7	15	9.26	-9.26
gente	1	46	14.94	32.06	-6.45	6.45
coisa(s)	0	38	12.08	25.92	-6.32	6.32
ações	7	0	2.23	4.77	3.48	-3.48
dinheiro estrangeiro	7	0	2.23	4.77	3.48	-3.48
capital	4	0	1.27	2.73	1.99	-1.99
pedidos	4	0	1.27	2.73	1.99	-1.99
votos	4	0	1.27	2.73	1.99	-1.99
informações	3	0	0.95	2.05	1.49	-1.49
processos	3	0	0.95	2.05	1.49	-1.49

Na tabela, são indicadas as frequências observadas, as frequências esperadas (considerando o conjunto total de dados), bem como a probabilidade de cada lexema ocorrer em cada construção aleatoriamente e, por fim, o grau de

significância de cada combinação. De acordo com Gries (2014), quanto maior o valor de P_{bin} , maior o grau de significância³.

Um breve olhar sobre os lexemas mais frequentes nas construções binominais quantitativas em pauta nos permite concluir que os itens mais atraídos pela construção 'UM MONTE DE SN' são os lexemas 'gente' e 'coisa (s)', que ocupam, respectivamente, na tabela, as linhas dois e três. Alonso e colegas (no prelo) já haviam demonstrado que 'gente' e 'coisa(s)' são os elementos mais atraídos pela construção 'UM MONTE DE SN' e, pelo princípio da coerência semântica, impactam fortemente a representação cognitiva que o falante tem dela – construção de quantificação de referentes genéricos. Sendo esses os usos prototípicos da construção 'UM MONTE DE SN', esse resultado de atração, mesmo que circunscrito ao *corpus* selecionado, parece bastante relevante e coerente com outros estudos acerca do tema, reforçando os resultados anteriormente obtidos. Complementarmente, o que a análise colexêmica distintiva traz de novo para a investigação da construção 'UM MONTE DE SN' é que os lexemas gente e coisa(s) são repelidos por 'UMA ENXURRADA DE SN', contribuindo para uma compreensão mais profunda da organização dessas construções na rede.

Para além dos nomes apresentados na tabela, outros lexemas sem significância estatística também puderam ser observados como colexemas da construção 'UM MONTE DE SN'. Esses lexemas formam um arcabouço diversificado de itens, que vão desde entidades animadas como 'amigos', 'caras', entre outros, até entidades inanimadas como 'filmes', 'assuntos' etc. Trata-se de itens que, embora semanticamente distintos, não se afastam muito da ideia de referentes genéricos prototipicamente recrutados pela construção.

Por esse mesmo raciocínio, lexemas como 'votos', 'pedidos', 'processos', também de referência genérica, seriam potenciais colexemas da construção 'UM MONTE DE SN'. Entretanto, neste *corpus*, esses lexemas, embora bastante produtivos, não só não apareceram na construção, como foram repelidos por ela. A

³ De acordo com o *script* proposto por Gries (2014), os valores são os seguintes: $p_{bin} > 3 \Rightarrow p < 0.001$; $p_{bin} > 2 \Rightarrow p < 0.01$; $p_{bin} > 1.30103 \Rightarrow p < 0.05$. Geralmente o teste multidimensional aplicado aqui é utilizado para um conjunto de três ou mais variantes de uma variável dependente. Porém, apesar de, neste trabalho, estarmos estudando somente duas construções alternantes, decidimos utilizar o mesmo teste aplicado a mais de duas variantes e não o teste exato de Fisher, mais indicado quando existem apenas duas variantes, porque consideramos o fato de que existem muito mais do que apenas duas microconstruções em competição no rol das construções quantitativas do português brasileiro. A diferença de significado a depender do teste aplicado não se demonstrou diferente do ponto de vista estatístico.

partir daí, alguém poderia perguntar por que isso acontece. A resposta parece não estar na avaliação de gramaticalidade de usos como ‘um monte de votos’ ou ‘um monte de pedidos’, por exemplo, mas numa investigação *a fundo* dos contextos em que os lexemas em questão foram encontrados instanciando construções binominais.

A natureza genérica dos lexemas atraídos pela construção ‘UM MONTE DE SN’ se distingue significativamente em relação aos lexemas mais atraídos por ‘UMA ENXURRADA DE SN’, a qual parece, ao menos em relação ao *corpus* utilizado, exibir maior especificidade, haja vista a maior preferência da construção em relação a lexemas como ‘dólares’, conforme se observa na primeira linha da tabela, bem como ‘ações’, ‘dinheiro estrangeiro’, entre outros.

Um olhar mais atento aos contextos de utilização de ‘UMA ENXURRADA DE SN’ nos permite identificar a recorrência do uso de tal construção em situações que envolvem movimento, como é possível observar nos trechos a seguir:

(3) Uma redução abrupta do compulsório, avaliam os técnicos, provocaria uma febre de operações de ACC no mercado e **uma enxurrada de dólares** entrando no país. (Coleção CHAVE)

(4) O livro consegue realizar com sucesso um dos objetivos básicos de seus idealizadores e organizadores: abrir, ou até escancarar, portas e janelas do universo da produção acadêmica de conhecimentos sobre os índios no Brasil para **uma vital enxurrada de informações** que deve (esperamos) invadir outro universo, o da educação escolar (Coleção CHAVE)

Em (3), a ocorrência ‘uma enxurrada de dólares’ ilustra um típico exemplo da construção ‘UMA ENXURRADA DE SN’, em que o contexto de movimento, sinalizado pelo verbo ‘entrar’, favorece a seleção de uma construção quantitativa que mantém com ele um grau alto de coerência semântica. Considera-se assim, visto que a noção de movimento é facilmente inferível do conceito de ‘enxurrada’. A recorrência do lexema ‘enxurrada’ em situações que envolvem deslocamento pode ser, inclusive, facilmente observada em outros contextos de uso, isto é, fora da construção quantitativa.

Entendemos que a construção ‘UMA ENXURRADA DE SN’ incorpora o sentido de movimento inferido do contexto e que exibe, por outro lado, maior analisabilidade (ou seja, o falante reconhece mais facilmente as partes componentes da construção) do que a que se observa em ‘UM MONTE DE SN’. Em sendo assim,

a referência de ‘enxurrada’ como um grande volume de água que se desloca depois de uma chuva, por exemplo, estaria mais disponível para o falante, em situação de uso em construção binominal quantitativa, do que o de ‘monte’ como formação geológica, se considerarmos seu uso na mesma situação. O efeito disso é o de que, nos casos de ‘UMA ENXURRADA DE SN’, o falante se encontra mais livre para apagar o artigo indefinido, substituí-lo por evento de natureza distinta, inserir elementos intervenientes etc.

O dado em (4) – ‘uma vital enxurrada de informações’ – ilustra exatamente essa situação em que são admitidos elementos intervenientes, no caso ‘vital’, entrando na construção. Esse exemplo reforça a argumentação em torno do fato de que a construção ‘UMA ENXURRADA DE SN’ é coerente com o sentido de movimento, tipicamente atrelado aos contextos em que é preferencialmente usada – nesse caso, o verbo ‘invadir’ sinaliza a ideia de deslocamento. Estamos argumentando, então, que, enquanto a construção ‘UM MONTE DE SN’ tem a sua semântica bastante atrelada a um conjunto de itens – mais destacadamente dois, ‘gente’ e ‘coisa(s)’ – de sentido genérico; a construção ‘UMA ENXURRADA DE SN’ parece preferir, ainda, se acomodar muito claramente à noção de movimento, que, ao contrário, não está nos itens recrutados propriamente ditos, mas na sua própria semântica e, por consequência, na sua recorrência em contextos de movimento nas diferentes situações de uso.

Observamos também, em relação a essa amostra, o uso do adjetivo ‘vital’ no que diz respeito à sua semântica. A seleção desse adjetivo modificando ‘enxurrada’ pareceria, fora da construção binominal quantitativa, no mínimo, inusitado; entretanto, na construção, o adjetivo ‘vital’ se coloca não em relação a ‘enxurrada’, tomando o lexema de forma isolada, mas em relação à ideia de muita informação expressa por ‘uma enxurrada de informações’, ou seja, é a quantidade grande de informações (relevantes) que é vital. Exemplos como esse nos mostram como uma futura análise qualitativa com maior detalhamento dos dados de UMA ENXURRADA DE SN se faz necessária.

Os lexemas preferencialmente recrutados por ‘UMA ENXURRADA DE SN’ são itens bastante específicos da área econômica, o que, nessa análise, atribuímos ao fato de que o *corpus* selecionado se compõe de textos jornalísticos e de que a maior parte dos dados foram encontrados em colunas da área da economia. A

ampliação dos dados poderá vir a revelar, no futuro, o quanto esse resultado reflete ou não a especificidade do *corpus*.

Considerações finais

Este artigo buscou refletir sobre de que modo uma análise colexêmica distintiva, como uma medida de associação entre lexemas e construções, pode contribuir para a identificação de aspectos semânticos sutis de construções alternantes, na rede de construções da língua. Mais especificamente, foi desenvolvida a análise colexêmica distintiva de construções binominais quantitativas, a saber, as construções ‘UM MONTE DE SN’ e ‘UMA ENXURRADA DE SN’.

Circunscritos a um *corpus* de língua escrita no domínio jornalístico, os resultados apontam para preferências colocacionais distintas, no que diz respeito a cada uma das duas construções estudadas neste trabalho. Se, por um lado, os lexemas atraídos pela construção ‘UM MONTE DE SN’ se caracterizam por sua natureza genérica, incluindo entidades animadas e inanimadas como ‘gente’ e ‘coisa’ e corroborando os resultados observados em trabalhos organizados pelos autores em outras ocasiões; por outro lado, a construção ‘UMA ENXURRADA DE SN’ tende a atrair lexemas inanimados de natureza mais específica e, no *corpus* analisado, demonstrou-se bastante recorrente com lexemas utilizados em textos que abordam a esfera econômica, tais como ‘dólares’, ‘ações’, ‘capital’, entre outros.

Ainda sobre a construção ‘UMA ENXURRADA DE SN’, observamos que a semântica de movimento ou fluxo, evocada pelo lexema ‘enxurrada’, parece contribuir para a semântica dessa microconstrução específica, corroborada pelo contexto em que a construção ocorre. É comum o uso da construção em contextos de verbos de movimento e o uso de preposições que expressam a direção da enxurrada, sinalizando a maior especificidade da construção ‘UMA ENXURRADA DE SN’ em relação à construção ‘UM MONTE DE SN’, em que não se observa um conteúdo semântico específico atrelado ao lexema ‘monte’, que, dado o seu percurso histórico (cf. FUMAUX, 2018), parece mais direta e exclusivamente associada à semântica de grande quantidade da construção.

A análise colexêmica distintiva permite observar, assim como na análise colexêmica simples, as preferências colocacionais da construção, na medida em que indica a frequência com que os lexemas ocorrem em seu *slot* esquemático. Porém,

para além do que a análise colexêmica simples nos fornece, a distintiva permite a comparação entre construções alternantes, contribuindo para a identificação do que é específico de cada construção e para a abstração dos aspectos comuns entre elas, que podem estar envolvidos em uma construção mais abstrata em um nível superior da rede. Se em um nível mais baixo na rede, as construções 'UM MONTE DE SN' e 'UMA ENXURRADA DE SN' apresentam especificidades semânticas e contextuais, a frequência com que ocorrem na língua associada à abstração das diferenças entre elas permite a postulação de um nível mais abstrato na rede, envolvido na expressão de grande quantidade e que pode ser notado como 'UM N DE SN', em que o *slot* esquemático 1 pode ser instanciado pelo lexema 'monte', que tende a atrair, para o *slot* esquemático 2, lexemas de cunho mais genérico e sem restrições de animacidade, estatisticamente observadas. Além disso, o mesmo *slot* esquemático 1 dessa construção mais abstrata também pode ser instanciado pelo lexema 'enxurrada', que tende a atrair, para o *slot* esquemático 2, lexemas com maior grau de especificidade, envolvidos em uma situação coerente com a semântica de movimento, e geralmente inanimados.

Evidentemente, a análise em pauta requer maior aprofundamento, com a ampliação do *corpus* para um conjunto mais amplo de gêneros textuais e domínios discursivos, incorporando-se também à análise textos da modalidade oral da língua, para que seja possível ter um panorama mais geral acerca do funcionamento das construções binominais quantitativas no conjunto de conhecimentos que o falante tem de sua língua, o que certamente será o tema de um trabalho a ser desenvolvido futuramente.

Referências

ALONSO, K. S. B. *Construções Binominais Quantitativas e Construção de Modificação de Grau: Uma abordagem baseada no uso*, 2010. Tese (Doutorado em Linguística), Programa de Pós-graduação em Linguística, Universidade Federal do Rio de Janeiro, 2010.

ALONSO, K. S. B. Quantificando nomes e diferenciando construções: o caso das microconstruções binominais do tipo um(a) N de SN, no prelo.

ALONSO, K. S. B.; LEITE DE OLIVEIRA, D.; FUMAUX, N. C. A, NASCIMENTO, G. F; SILVA, T. M. Quantifying binominal constructions in Portuguese and Russian: the case of *um monte de NP* and *kucha NPgen*, no prelo.

BARDDAL, J. *Productivity: evidence from case and argument structure in Icelandic*. Vol. 8. Amsterdam: John Benjamins, 2008.

BRODBECK, R. C. M. S. *Um monte de problemas gera uma chuva de respostas: um estudo de caso de desencontro na quantificação nominal em português*. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal de Juiz de Fora, 2010.

BYBEE, J. Usage-based Theory and Exemplar Representations of Constructions. In: HOFFMANN, T.; TROUSDALE, G. *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press, 2013, p. 49-69.

BYBEE, J. *Language, usage and cognition*. New York: Cambridge University Press, 2010.

CROFT, W. *Radical Construction Grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press, 2001.

DIESSEL, H. Usage-based construction grammar. In: DABROWSKA, E.; DIVJAK, D. (eds.). *Handbook of Cognitive Linguistics*. Berlin: Mouton de Gruyter, 2015, p. 296-321.

FILLMORE, C. J.; KAY, P.; O'CONNOR, M. C. Regularity and idiomaticity in grammatical constructions: The case of *Let alone*. *Language*, v. 64, n. 3, p. 501-538, 1988.

FUMAUX, N. C. A. *Construcionalização de 'um monte de SN': uma abordagem centrada no uso*. Dissertação (Mestrado em Linguística). Programa de Pós-graduação em Linguística, Universidade Federal do Rio de Janeiro, 2018.

FURTADO DA CUNHA, M. A. As construções de movimento causado e ditransitiva: elos de polissemia. *D.E.L.T.A.*, v. 33, n.1, p. 109-132, 2017.

GOLDBERG, A. E. *Constructions: A construction grammar approach to argument structure*. Chicago/London: The University of Chicago Press, 1995.

GOLDBERG, A. E. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press, 2006.

GOLDBERG, A. E. *Explain me this: creativity, competition and the partial productivity of constructions*. Princeton: Princeton University Press, 2019.

GRIES, S. Th. Coll.analysis 3.5. A script for R to compute perform colostruational analyses, 2014.

GRIES, S.; STEFANOWITSCH, A. Extending colostruational analysis. A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics*, v. 9, n. 1, p. 97-129, 2014.

LAKOFF, G. *Women, fire and dangerous things*. What categories reveal about mind.

Chicago: Chicago University Press, 1987.

LANGACKER, R. *Foundations of Cognitive Grammar*, v. 1: theoretical prerequisites. Stanford: Stanford University Press, 1987.

HILPERT, M. *Construction Grammar and its application to English*. Edinburgh: Edinburgh Textbooks on English Language, 2014.

PEREK, F. *Argument Structure in Usage-Based Construction Grammar*. Amsterdam: John Benjamins, 2015.

STEFANOWITSCH, A. E.; GRIES, S. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, v. 8, n. 2, p. 209-243, 2003.

TALMY, L. Grammatical construal: the relation of grammar to cognition. In: GEERAERTS, D. (org.). *Cognitive Linguistics: basic readings*. Berlin/Nova York: Mouton de Gruyter, 2006.

TRAUGOTT, E. C.; TROUSDALE, G. *Construcionalization and Constructional Changes*. Oxford: University Press, 2013.

ZHAN, F. & TRAUGOTT, E. C. The constructionalization of the Chinese cleft construction. *Studies in Language*, n. 39. John Benjamins Publishing Company, 2015.

Recebido em 31/08/2019

Aceito em 11/11/2019