



**DIFERENÇAS SEMÂNTICAS DE MICROCONSTRUÇÕES
QUANTIFICADORAS: O CASO DE *SN1 DE SN2***

**SEMANTIC DIFFERENCES OF QUANTIFYING
MICROCONSTRUCTIONS: THE CASE OF BRAZILIAN *NP1 OF NP2***

*Karen Sampaio Braga Alonso*¹
*Nuciene Caroline Amphilóphio Fumaux*²

RESUMO

Este trabalho consiste em um estudo de construções binominais quantificadoras do tipo *SN1 de SN2* no português brasileiro sob a ótica da Linguística Baseada no Uso (BYBEE, 2010; BARLOW; KEMMER 2000; BOYLAND, 2009), associando os pressupostos da Gramática de Construções (GOLDBERG, 2006; DIESSEL, 2015; HILPERT, 2014; PEREK, 2015) com o Modelo de Exemplares (BYBEE, 2006; 2010; 2013; PIERREHUMBERT, 2001). Com o intuito de descrever o modo como ocorre a distribuição dessas construções no uso da língua, nós comparamos quatro microconstruções quantitativas, a saber *DET monte de SN*, *DET montanha de SN*, *DET enxurrada de SN* e *DET chuva de SN*, em relação às quais foi realizada uma análise colexêmica (GRIES; STEFANOWITSCH, 2004). Os resultados indicaram que as referidas microconstruções, que se assemelham funcionalmente, possuem preferências colocacionais próprias, as quais foram associadas à maior ou menor coerência semântica entre os nomes das colocações encontradas. Por exemplo, enquanto a microconstrução *DET monte de SN* *recruta preferencialmente nomes mais genéricos*, *DET montanha de SN* *prefere SN* cujo referente apresentem a propriedade de ser mais empilhável ou que estejam organizados de forma aglomerada.

PALAVRAS-CHAVE: Construções binominais quantitativas; quantificação; Gramática de Construções Baseada no Uso

1 Doutora em Linguística, Universidade Federal do Rio de Janeiro. Professora Adjunta do Departamento de Linguística e Filologia da UFRJ, Professora do Programa de Pós-graduação da UFRJ e Professora do PROFLETRAS/ UFRJ. karensampaio@letras.ufrj.br.

2 Doutoranda do Programa de Pós-graduação em Linguística da UFRJ. carol.fumaux@gmail.com.



ABSTRACT

The aim of this work is to analyze binominal quantifier constructions (NP1 OF NP2) in Brazilian Portuguese, from the perspective of Usage-Based Linguistics (BYBEE, 2010; BARLOW ; KEMMER 2000; BOYLAND, 2009), focusing on the association between Construction Grammar (GOLDBERG, 2006; DIESSEL, 2015; HILPERT, 2014; PEREK, 2015) with exemplar representation for constructions (BYBEE, 2006; 2010; 2013; PIERREHUMBERT, 2001). In order to identify the most frequent collocations of some binominal quantitative constructions, we observed the behavior of four microconstructions: *DET monte de NP*; *DET montanha de NP*; *DET enxurrada de NP*; *DET chuva de NP*. The results achieved in the analysis demonstrate that, although those microconstructions are functionally similar, they differ in terms of their own semantics (considering the semantic coherence between the two referents within the construction). For example, *DET monte de NP* is more frequently associated to generic referents while *uma montanha de NP* recruits mainly referents which are piled or gathered.

KEYWORDS: Quantitative binominal constructions; quantification; Usage-Based Construction Grammar

Introdução

O presente artigo tem como objetivo a análise de construções binominais do tipo *SN1 de SN2*, a partir da verificação de como essas formas são usadas na língua. Nesse sentido, pretendemos comparar as microconstruções *DET monte de SN*, *DET enxurrada de SN*, *DET chuva de SN* e *DET montanha de SN*, com a finalidade de verificar como essas construções, aparentemente sinônimas, distribuem-se no português, no que tange às suas preferências colocacionais, avaliando, a partir daí, o grau de coerência semântica entre o referente recrutado como quantificador (*monte*, *montanha*, *enxurrada* e *chuva*) e o referente quantificado.

Construções binominais quantificadoras do tipo *SN1 de SN2* caracterizam-se por recrutar nomes não tradicionalmente tomados como quantificadores, como *enxurrada* ou *montanha*, para quantificar referentes numa dada língua. Por exemplo, o referente *enxurrada*, instanciando a construção *SN de SN*, produz resultados como *uma enxurrada de dólares*, por exemplo, em que *dólares* é quantificado em termos de grande quantidade. Por conta dessa base lexical, assumimos, por hipótese, que o sentido do referente recrutado como quantificador poderá influenciar em maior ou menor grau a seleção do referente tomado como quantificado, a depender do quanto ele é recuperável ao instanciar a construção *SN de SN*. Sendo assim, apostamos no fato de que as preferências colocacionais encontradas poderão nos informar sobre a coerência semântica entre quantificador e quantificado (avaliação da composicionalidade), ainda que todas as microconstruções estudadas compartilhem a mesma função de quantificar referentes (em termos de grande quantidade/volume).

Assume-se, portanto, que os itens quantificados preferencialmente para cada microconstrução poderão ser mais ou menos coerentes com a semântica do quantificador, e que isso indica graus distintos de composicionalidade, bem como tem relação com a maior

produtividade de cada microconstrução, dado que, por hipótese, quanto mais produtiva for a construção, menos o falante acessará o sentido de base do referente quantificador. Para dar conta das preferências colocacionais de cada microconstrução, será realizada análise colexêmica para cada microconstrução (GRIES ; STEFANOWITSCH, 2004), a qual compõe um conjunto de métodos que objetivam, de forma geral, medir a relação de atração entre lexemas e construções.

As construções *DET monte de SN*, *DET enxurrada de SN*, *DET chuva de SN* e *DET montanha de SN*, citadas anteriormente, mostram-se, no português do Brasil, como alternativas linguísticas para a quantificação de grande quantidade, grande volume em relação a algo: por exemplo, *um monte de coisas* expressa uma grande quantidade de *coisas* (iguais ou distintas entre si), assim como *uma enxurrada de dólares* exprime uma quantia relevante da moeda americana. Da mesma forma, *uma montanha de livros* é usada para se referir a uma quantidade considerável de livros e *uma chuva de críticas* expressa muitas críticas.

Tomando esses exemplos, então, é possível perceber que eles se assemelham tanto em termos formais – todos se estruturam por meio de uma sequência sintagmática que se caracteriza por apresentar um sintagma nominal (cujo núcleo é a base da quantificação ou quantificador), a preposição *de* e um sintagma nominal (quantificado) – quanto pela função que exercem na língua – a de expressar grande quantidade/ volume de algo. Dadas essas semelhanças, estamos considerando que tais exemplos ilustram microconstruções ligadas à construção mais geral *SNI de SN2*, com valor de grande quantidade/volume de algo, conforme demonstrado na figura 1, a seguir:

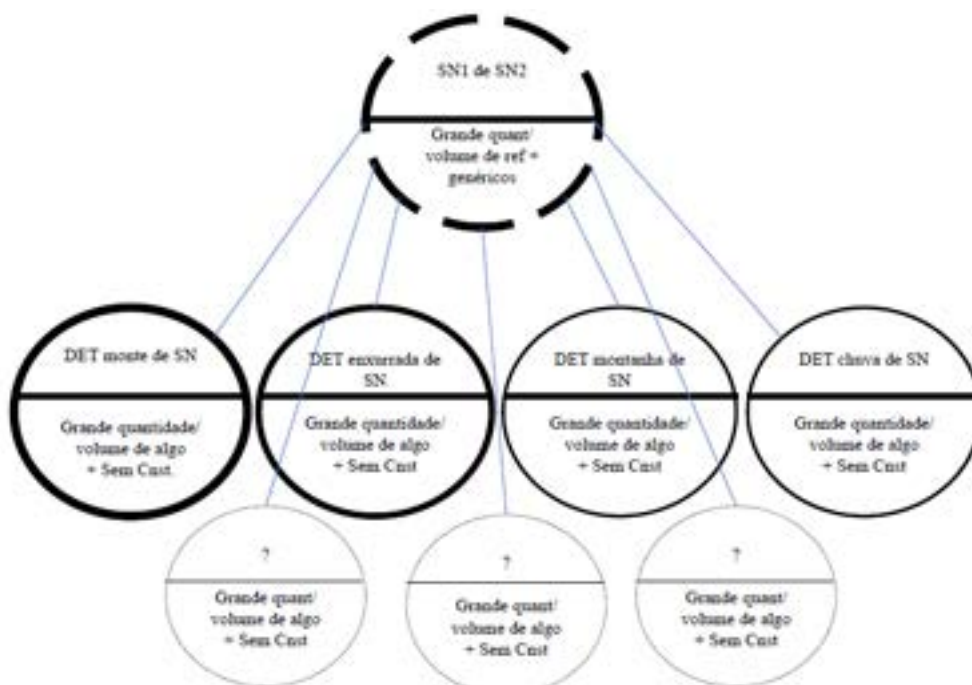


Figura 1: relação entre microconstruções e a construção *SNI de SN2*

A Figura 1 ilustra o fato de que essas quatro microconstruções são funcionalmente equivalentes, ou seja, compartilham a mesma função (pertencendo ao esquema mais geral

dos quantificadores, como se verá adiante), a saber: a de quantificar referentes em grande quantidade ou volume. Mas será que essa correspondência funcional equivale a dizer que são *perfeitamente sinônimas*? Perseguindo responder essa questão, o presente trabalho investe no sentido de mostrar como tais microconstruções se distribuem na língua e como se diferenciam em termos semânticos, dado os diferentes graus de composicionalidade em que se encontram. Como se vê, estamos forçosamente fazendo uma diferenciação entre a função de quantificador e o significado de cada microconstrução individual, associando-o ao grau de maior ou menor transparência do referente recrutado como quantificador, embora esses dois termos sejam, em diversos outros momentos, usados muito amplamente e indiferentemente e que, em última análise, estamos tratando do polo do sentido da construção.

A diferença explorada deliberadamente entre sentido e função se justifica na presente pesquisa pelo fato de que, primeiramente, estamos buscando associar as microconstruções a uma construção mais geral – *SN1 de SN2* –, como resultado de processos analógicos que resultam na percepção da semelhança formal e funcional entre elas (o falante percebe, portanto, que há diferentes construções estruturalmente parecidas e que *servem para quantificar elementos em grande quantidade/volume*). Em seguida, estamos igualmente investindo na análise da composicionalidade dessas construções, a partir da coerência entre a semântica do nome quantificador – *monte, enxurrada, montanha e chuva* – e a semântica do referente quantificado (SN2).

Sobre os tipos semânticos dos nomes recrutados por cada microconstrução, vemos o seguinte: de forma geral, a microconstrução *DET monte de SN* tende a se combinar com nomes genéricos (gente, coisa, caras) pouco coerentes com o sentido de *monte* (formação geológica, elevação vertical, etc.), mas condizentes com as mudanças construcionais observadas por FUMAUX (2018) e FUMAUX; ALONSO; CEZÁRIO (2017), quando trataram do surgimento da construção *DET monte de SN* ao longo da história do português. No âmbito do presente artigo, espera-se que as demais microconstruções tenham uma distribuição distinta da de *DET monte de SN*, já que temos, por hipótese, que o grau de coerência entre nome quantificador e nome quantificado afeta a escolha do falante por uma ou outra microconstrução.

Mais especificamente, esperamos demonstrar, a partir de análise de *corpus*, as diferenças e semelhanças distribucionais (em termos de preferências colocacionais) entre as microconstruções citadas anteriormente, assumindo, com base nos pressupostos teórico-metodológicos da Linguística Baseada no Uso, que:

- (i) as quatro construções estão ligadas, na rede construcional, por um elo de herança taxonômica mais alto – a saber, o da construção *SN1 de SN2* com função de quantificar grande quantidade/ volume;
- (ii) no nível microconstrucional, *DET monte de SN*, *DET enxurrada de SN*, *DET chuva de SN* e *DET montanha de SN* apresentam preferências colocacionais próprias, bem como compartilham contextos de uso;

- (iii) as quatro microconstruções competem entre si, de forma que a microconstrução mais frequente tende a impactar mais fortemente a representação da construção *SNI de S N2* na mente do falante.

A Linguística Baseada no uso

Apesquisa em que o artigo se baseia está ancorada nos pressupostos teórico-metodológicos da Linguística Baseada no Uso, que rechaçam a ideia de um conhecimento especificamente linguístico inato, defendendo, ao contrário, um conhecimento desenhado pelos processos cognitivos de domínio geral – como *chunking*, associação transmodal, memória rica, etc. (para mais detalhes, ver Bybee, 2010).

Nas palavras de Boyland (2009: 352-353),

Modelos [baseados no uso] propõem que o conhecimento linguístico é corporificado em processamento mental e representações mentais que são sensíveis a contextos e probabilidades estatísticas, as quais são tanto cognitivamente plausíveis como poderosas o bastante para dar conta da complexidade do uso linguístico real.³

Assume-se, nesse sentido, que, sob a ótica da Linguística Baseada no Uso, a gramática é moldada no uso, e a gramática é descrita em termos de uma rede de construções – pareamentos simbólicos de forma e sentido (GOLDBERG, 2006). Nesse sentido, rejeita-se o modelo de dicionário e gramática, em prol de um modelo de construções gramaticais organizadas num formato de rede – *constructicon* (para saber mais sobre as relações dentro do *constructicon*, ver Diessel, 2015).

De acordo com Traugott ; Trousdale (2013:3) as diferentes gramáticas de construções (incluem-se aí abordagens formalistas e funcionalistas) compartilham os seguintes postulados:

- (a) A unidade básica da gramática é a construção, a qual é um pareamento convencional de forma e sentido;
- (b) a estrutura semântica é mapeada diretamente na estrutura sintática superficial, sem derivações;
- (c) a língua, como outros sistemas cognitivos, é uma rede de nós e relações entre nós; as associações entre alguns desses nós tomam a forma de hierarquias por herança (relações taxonômicas capturando o grau em que as propriedades de construções de nível mais baixo são previsíveis a partir daquelas mais gerais);
- (d) variação translinguística (ou dialetal) pode ser explicada por vários fatores, dentre eles os processos cognitivos de domínio geral.

³ These models propose that linguistic knowledge is embodied in mental processing and mental representations that are sensitive to context and statistical probabilities, which are both cognitively plausible and powerful enough to account for the complexity of actual language use.

O entendimento da gramática como rede (premissa *c*) prevê uma série de relações entre as formas linguísticas – dentre as quais, relações horizontais e relações verticais (DIESSEL, 2015), que vão nos interessar mais especificamente na presente pesquisa. No caso deste artigo, as microconstruções tomadas como foco de análise estão, entre si, em uma relação horizontal e por uma relação taxonômica com a construção mais geral *SN1 de SN2*, verticalmente, por meio de elo taxonômico, conforme já indicado na figura 1. Tendo essas diferentes relações em vista, defenderemos que as microconstruções binominais estão competindo (em relação horizontal), de tal forma que a construção mais frequente seja aquela que vai impactar mais fortemente a representação cognitiva do falante em relação à construção mais geral, ou seja, mais alta na rede (PEREK, 2015).

Analogamente, no nível mais baixo, os usos (exemplares) vinculados às microconstruções também competem entre si e impactam a representação mental que o falante faz de cada microconstrução. Sendo assim, assumimos que, se nomes de referência genérica como *gente* e *coisa(s)* são mais recrutados pela microconstrução *DET monte de SN*, esse fato levaria os falantes a acessarem tal construção mais rapidamente quando quantifica uma grande quantidade de coisas tomadas genericamente (sejam animadas ou não). Esse viés de análise, que será retomado mais adiante, quando mencionarmos mais diretamente o Modelo de Exemplares (BYBEE, 2006; 2010; 2013; PIERREHUMBERT, 2001), mostra como o uso da língua é capturado e cada exemplar influencia a representação que o falante faz das estruturas gramaticais.

Essa relação entre gramática e uso é um dos pontos que justificam uma gramática de construções baseada no uso, uma vez que instâncias da língua moldam a gramática. Toma-se, portanto, o fato primordial de que sequências sintagmáticas advindas do uso – e, por regra geral, frequentemente repetidas – podem ser rotinizadas (*entrenched*) e armazenadas como construções gramaticais.

Traugott e Trousdale (2013:17) oferecem um esquema para representar a distribuição taxonômica (ou seja, considerando níveis distintos de esquematicidade) das construções linguísticas, que vai desde as microconstruções até o esquema mais geral:

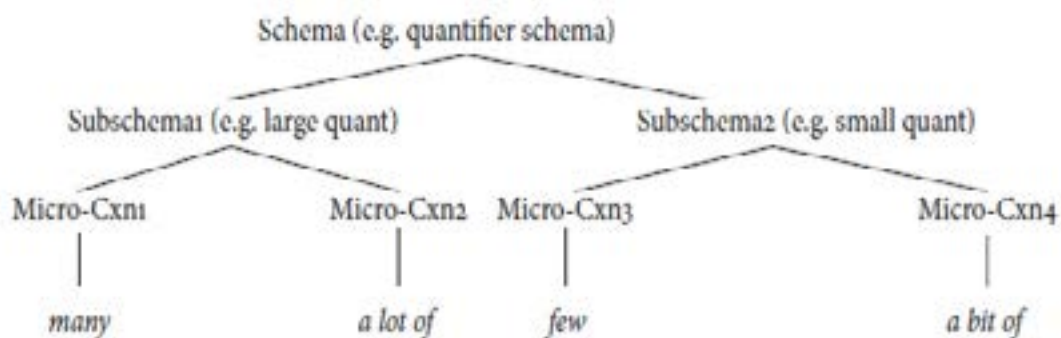


Figura 2: Esquema, subesquema e microconstrução

Na figura 2, além da relação hierárquica mencionada, os autores também destacam as relações horizontais (ou *horizontal links*, como chama DIESSEL, 2015) que se estabelecem na rede entre construções de mesmo nível, ou seja, com o mesmo grau de esquematicidade.

A aceção de esquema adotada por Traugott; Trousdale (2013) interessa a esse estudo, uma vez que objetivamos demonstrar, por meio de análise estatística com base em dados de *corpus*, de um lado, as especificidades semânticas de cada microconstrução; de outro, o impacto das instâncias de um conjunto de usos e/ou construções na construção de nível imediatamente acima. Assim, queremos observar o quanto uma construção como *DET enxurrada de SN*, por exemplo, impacta a construção mais geral *SN1 de SN2*, do mesmo modo que pretendemos explorar a diferença semântica entre *DET enxurrada de SN* e *DET monte de SN*, por exemplo.

As respostas para essas questões de pesquisa, como era de se esperar, estão longe de serem simples. Em relação ao impacto que as construções binominais exercem sobre a construção mais geral, por exemplo, temos de levar em conta que qualquer resultado de análise apresentada aqui incorre necessariamente em dois problemas: o primeiro, no fato de que não estamos analisando todas as construções binominais do tipo *um N1 de N2* que aparecem no nosso *corpus* – o que nos leva a relativizar o impacto real que cada microconstrução selecionada no estudo exerce sobre a construção mais geral; segundo, no fato de que, exatamente por termos um *corpus* e, portanto, um recorte da realidade de uso da língua, qualquer inferência está circunscrita a esse recorte. Ainda assim, acreditamos que os resultados nos apresentam *insights* interessantes acerca da representação mental que os falantes fazem das estruturas linguísticas, bem como da relação interdependente entre gramática e uso.

Sobre a diferença semântica de construções que compartilham função semelhante – no caso, a de quantificar elementos em termos de grande quantidade/ volume –, esbarramos novamente no recorte feito, o qual poderá ter impulsionado o uso de algumas construções e não de outras. Em pesquisa futura, com a ampliação da amostra analisada, poderemos avaliar, com mais precisão, o impacto do recorte escolhido para a rodada de dados abrangida neste artigo. Fato é que a quantidade de exemplos se apresenta bem mais significativa – acompanhando, portanto, uma maior ou menor capacidade de consolidação, via *corpus*, das análises oferecidas na pesquisa – em *DET monte de SN* e em *DET enxurrada de SN* do que em *DET montanha de SN* e *DET chuva de SN*.

Assumindo, portanto, os limites da análise, seguimos refletindo acerca de como os usos de sequências sintagmáticas binominais do tipo *SN1 de SN2* modelam a gramática, ao mesmo tempo que nos oferecem informações importantes para a investigação almejada nesta pesquisa – a de estabelecer preferências colocacionais e interpretações coerentes para a seleção de uma dentre as quatro microconstruções quantitativas estudadas, no português brasileiro.

Em um trabalho funcionalista clássico, Hopper (1987) defende a concepção de gramática emergente, que prevê que a gramática vem do discurso e por ele é moldada, não devendo ser

encarada, portanto, como um conhecimento anterior ao discurso, muito menos ser atribuída de forma idêntica a falantes e ouvintes. A gramática é, assim, o tempo todo negociada entre os interlocutores no momento da interação, estando necessariamente ancorada nos enunciados concretos da língua. Essa concepção de gramática constitui um dos fundamentos básicos para se pensar em uma estrutura gramatical que está em constante mudança e é, portanto, sensível a avaliações estatísticas sobre dados de uso.

Nesse sentido, a presente pesquisa incorpora o conceito de gramática emergente de Hopper, apostando que a arquitetura gramatical está o tempo todo sendo afetada pelos usos linguísticos. Faz sentido, então, entendendo a gramática dessa forma, investir em análise quantitativa e qualitativa de dados empiricamente atestados, como nos propomos a fazer neste trabalho e investigar a relação de retroalimentação entre gramática e uso. Retroalimentação, no sentido de que a gramática é estruturada no uso, mas também é estruturante, já que a forma como nós usamos a língua não é de todo inovadora, uma vez que lançamos mão de estruturas gramaticais (construções) já formadas.

O entendimento de língua como um sistema adaptativo complexo, tal como apresentado por Bybee et al. (2009), parece se conformar bem à noção de gramática emergente. Segundo os autores, é possível identificar nas línguas propriedades típicas de sistemas adaptativos complexos, tal como o fato de que uma língua envolve agentes múltiplos e interação entre eles; de que é um sistema adaptativo, que leva em conta o fato de que o comportamento dos falantes sofre influência de experiências passadas e atuais e que essas experiências são base para seu comportamento no futuro. Soma-se a isso o fato de que o comportamento do falante responde a fatores em competição e de que as estruturas linguísticas emergem de padrões inter-relacionados de experiência, interação social e processos cognitivos.

Uma reflexão mais aprofundada da visão de língua como um sistema adaptativo complexo pode ser encontrada em Ellis, Larsen; Freeman (2009). Segundo os autores (op. cit, p. 14-15):

A língua existe tanto em indivíduos (como idioletos) quanto numa comunidade de usuários (como língua comum). A língua é emergente nesses dois níveis distintos, porém interdependentes: um idioleto é emergente do uso linguístico individual, através das interações sociais com outros indivíduos na comunidade linguística, enquanto a língua comum é emergente como o resultado da interação dos idioletos. Distinção e conexão entre esses dois níveis é uma propriedade comum num SAC [sistema adaptativo complexo]. Padrões no nível coletivo (tal como bandos de pássaros, cardumes, economias) não podem ser atribuídos a uma coordenação global entre indivíduos; o padrão global é emergente, resultante de interações locais de longo prazo entre indivíduos.⁴

4 Language exists both in individuals (as idiolect) and in the community of users (as communal language). Language is emergent at these two distinctive but interdependent levels: An idiolect is emergent from an individual's language use through social interactions with other individuals in the communal language, whereas a communal language is emergent as the result of the interaction of the idiolects. Distinction and connection between these two levels is a common feature in a CAS.

Assumindo, então, as concepções de gramática emergente e de língua como um sistema adaptativo complexo, a presente pesquisa pretende, por meio de análise de dados em corpora, extrair inferências acerca da maneira como os falantes encaram as formas linguísticas. Nesse espírito, Boyland (2009:352) afirma que:

Quando falantes memorizam múltiplas instâncias de uso da língua em contextos particulares, eles desenvolvem uma base de conhecimento internalizado cada vez mais rica, a partir da qual eles podem, de uma forma cognitivamente realista, gerar generalizações cada vez mais sofisticadas sem recorrer necessariamente a regras gramaticais pré-existentes. A ideia de que processos baseados no uso e representações constituem conhecimento linguístico tem muitas implicações. Isso afeta nosso entendimento dos fenômenos sincrônicos, que são vistos normalmente como explicáveis por meio de regras, nosso entendimento de como se dá a mudança diacrônica e de que forma a estabilidade é mantida, como também nosso entendimento de padrões de variabilidade tanto dentro do indivíduo como entre falantes.⁵

Assumindo que o uso impacta as representações cognitivas, o presente trabalho adota o Modelo de Exemplares (BYBEE, 2006; 2010; 2013; PIERREHUMBERT, 2001), que prevê que o falante armazena todas as ocorrências de uma dada construção em termos de uma nuvem de exemplares, de forma que elementos mais fortemente relacionados – em termos de forma e/ou conteúdo – se encontrariam mais próximos na rede cognitiva. Da mesma maneira, elementos menos relacionados estariam mais distantes na rede.

Além disso, é importante ressaltar que, dentro de um modelo de exemplares (BYBEE, 2006; 2010; 2013; PIERREHUMBERT, 2001), usos mais frequentes afetam mais fortemente a representação mental que o falante faz de uma dada categoria, e categorias mais ricas, ou seja, com mais itens a ela associados, tendem a se fortalecer cognitivamente. Essa reflexão faz sentido no modelo de exemplares, uma vez que cada uso é armazenado como um item distinto, sendo as categorias, portanto, sensíveis às avaliações estatísticas de frequência das ocorrências na língua.

Metodologia

Patterns at the collective level (such as bird flocks, fish schools, or economies) cannot be attributed to global coordination among individuals; the global pattern is emergent, resulting from long-term local interactions between individuals.

⁵ When speakers record multiple instances of language use in particular contexts, they develop an increasingly rich implicit knowledge base from which they can, in a cognitively realistic way, generate increasingly sophisticated generalizations without necessary recourse to a priori grammatical rules. The idea that usage-based processes and representations constitute knowledge of language has manifold implications. It profoundly affects our understanding of synchronic phenomena that are currently seen as explainable by rules, our understanding of how change takes place diachronically as well as how stability is maintained, and our understanding of patterns of variability both within and across speakers.

Em termos metodológicos, o nosso objetivo é identificar as preferências colocacionais de cada microconstrução, a fim de obter informações acerca da contribuição do nome que ocupa o primeiro *slot* aberto e do grau de composicionalidade que apresentam. Para isso, realizamos, a partir dos dados coletados, uma análise de cunho colostrucional. Esse tipo de análise possui como objetivo medir a força de atração entre item e construção, isto é, medir a força colostrucional. O conjunto de análises sob o rótulo mais geral de *análise colostrucional* se subdivide em três métodos distintos, a saber: (i) a análise colexêmica, que verifica a força de atração de lexemas em um *slot* da construção; (ii) a análise colexêmica distintiva, que mede a força de atração de lexemas em um *slot* de duas ou mais construções; e (iii) a análise colexêmica covariacional, em que se observa a força de atração de lexemas em dois ou mais *slots* de uma construção (cf. Gries; Stefanowitsch, 2004).⁶

Neste trabalho, iremos realizar a análise colexêmica para verificar os itens mais atraídos por cada microconstrução, e, assim, descrever melhor os usos de cada uma, além de observar mais profundamente a competição entre essas construções quantificadoras. Para realizar a análise, coletamos os dados do *corpus* CHAVE, que reúne textos jornalísticos dos jornais Público de Portugal e Folha de São Paulo do Brasil reunidos de 1994 a 1995, dos quais foram utilizados somente a amostra do português brasileiro (Folha de São Paulo), que correspondem a um total de 35.699.765 palavras.

A análise colexêmica foi realizada no programa estatístico R e leva em conta, entre outros fatores, o número total de palavras do *corpus* selecionado, o número de vezes que um dado item aparece nesse *corpus* e no *slot* de uma construção, bem como o número de dados encontrados dessa construção no *corpus*. Esse método de análise fornece uma série de informações acerca da combinação de itens e construções em uma dada língua, dentre as quais a força colostrucional, que nos interessa mais fortemente neste artigo.⁷ Realizamos, assim, a análise colexêmica de *DET montanha de SN*, *DET enxurrada de SN* e *DET chuva de SN*. Para a microconstrução *DET monte de SN*, retomamos o resultado apresentado por Alonso et al. (no prelo)⁸.

A seguir, apresenta-se uma análise mais geral das construções binominais do tipo *SN1 de SN2*, seguida dos resultados da análise colexêmica aplicada a cada microconstrução.

⁶ Coll.strength>3 => p<0.001; coll.strength>2 => p<0.01; coll.strength>1.30103 => p<0.05. Forças colostrucionais acima de 3 são consideradas relevantes.

⁷ Optamos por apresentar nas tabelas as forças colostrucionais maiores do que 3, com exceção da microconstrução *DET enxurrada de SN*, que apresentou muitos itens com força de atração relevante; desta forma, por economia de espaço, apresentaremos um conjunto dos lexemas que foram mais atraídos pela construção.

⁸ Alonso et al. (no prelo) utilizou o mesmo *corpus* e somente a variante do português brasileiro, assim como fizemos neste trabalho; desta forma, podemos utilizar a sua análise em comparação com as realizadas nessa pesquisa.

Construções quantificadoras: considerações gerais

Para se ter um panorama geral do uso das construções quantificadoras do tipo *SN de SN*, toma-se, de início, o seguinte gráfico:

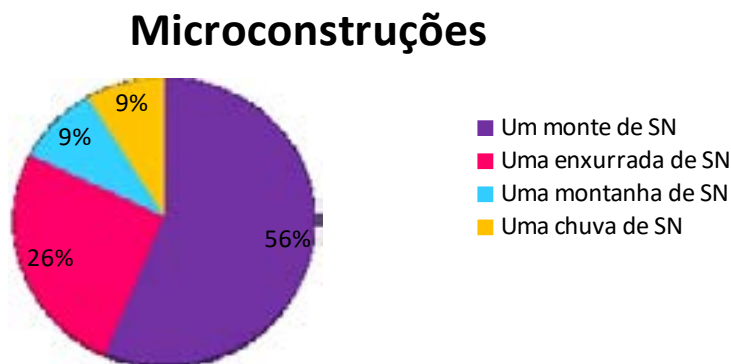


Gráfico 1: Microconstruções de *SN1 de SN2*

No gráfico (1), podemos perceber que, dentre as microconstruções binominais quantificadoras tomadas para estudo e considerando que os resultados estão circunscritos ao *corpus* selecionado, *DET monte de SN* apresenta um percentual maior de dados (56% do total, equivalente a 296 ocorrências). Ela é seguida de *DET enxurrada de SN* (com 26% do total, equivalente a 137 ocorrências), *DET montanha de SN* (com 9% dos dados ou 48 ocorrências) e *DET chuva de SN* (com 9% do total de dados coletados ou 46 ocorrências).

Considerando a amostra de dados, as microconstruções *DET monte de SN* e *DET enxurrada de SN* são mais fortes cognitivamente, em termos da representação mental que o falante teria da construção quantitativa mais geral *SN1 de SN2*, sendo *DET monte de SN* a construção que apresenta mais dados, ou seja, que possui uma nuvem de exemplares ainda mais rica. Parece, de fato, ser essa a situação atual da língua; entretanto, é preciso lembrar que o Modelo de Exemplares (Bybee, 2006; 2010; 2013; Pierrehumbert, 2001) prevê que, caso qualquer uma das demais microconstruções venha a acumular uma quantidade maior de exemplares em sua nuvem do que aquele que se atesta atualmente com *DET monte de SN*, a construção mais geral também terá sua representação cognitiva alterada, acompanhando a dinâmica da língua em uso.

Para falarmos um pouco mais sobre isso, vejamos as figuras 3 e 4, a seguir:

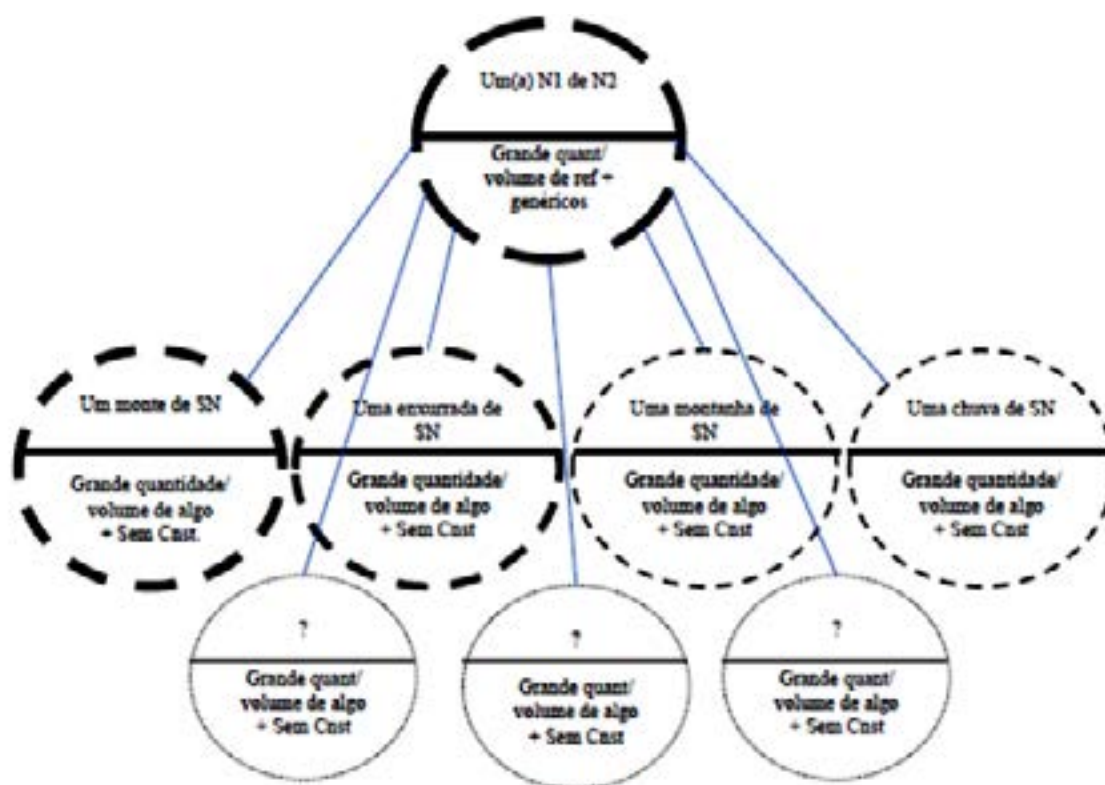


Figura 3: rede e exemplares para *um(a) N1 de N2*

A figura 3 pretende condensar algumas considerações feitas até aqui acerca da forma como inserimos as construções quantitativas binominais do tipo *SN1 de SN2* no aporte teórico-metodológico da Linguística Baseada no Uso. No alto da figura, encontra-se um círculo com a construção mais geral, cuja forma *SN1 de SN2* é associada à função de quantificar elementos +genéricos e de grande quantidade/ volume – uma vez que é mais fortemente impactada pela microconstrução *DET monte de SN*, que possuiria uma nuvem de exemplares mais rica. As linhas tracejadas indicam ser essa condição (função) provisória, uma vez que está necessariamente atrelada à dinâmica de forças das construções de nível mais baixo, a partir das quais a construção mais geral é abstraída a partir de processo analógico (ou seja, após o reconhecimento de similaridades entre os membros da categoria). Outras construções, ilustradas nos três círculos inferiores, podem estar ligadas a essa construção mais geral, a depender do recorte feito para análise. Na figura, procuramos apenas ilustrar possíveis casos de microconstruções com baixo impacto na construção de nível superior, fazendo uso, para isso, de uma linha pontilhada fina contornando o círculo. A configuração mais adequada dessa nuvem deverá ser capturada por meio de análise exaustiva de dados e olhando todas as microconstruções binominais quantitativas presentes no *corpus*.

As microconstruções estudadas no presente artigo, as quais serão mais bem discutidas na próxima seção – de um lado, compartilham a mesma função – propriedade representada também no nível construcional mais ascendente –; de outro, apresentam propriedades semânticas

específicas, as quais estão diretamente associadas aos seus respectivos exemplares, os quais nos ensinam sobre a contribuição do nome que ocupa o primeiro slot aberto da construção bem como sobre os diferentes graus de composicionalidade que apresentam. A gradação das linhas tracejadas nos círculos correspondentes às microconstruções mostram o quão rica é a nuvem de exemplares de cada microconstrução.

A seguir, propomos uma representação que revela a dinâmica de forças que envolve a relação gramática e uso, conforme a proposta teórica do Modelo de Exemplares. Nesse caso, especificamente, estamos deixando as colocações expressas.



Figura 4: exemplares para *DET monte de SN*

A figura 4 ilustra um conjunto de colocações observadas na análise baseada em corpus e procura, por meio de recursos de edição de texto, representar uma esquematização da nuvem de exemplares associada à microconstrução *DET monte de SN*. Procuramos dar mais destaque – combinando letras em caixa alta e negrito – a representações de exemplares mais frequentes no *corpus* selecionado para análise e menos destaque – caixa baixa sem negrito – para os menos frequentes. Assim, optamos por não replicar numericamente os exemplares em coerência à frequência desses no *corpus*. Essa figura traz outra novidade em termos de representação por exemplares que a figura anterior não trouxe: procuramos aproximar exemplares mais frequentes (assim como afastar os mais infrequentes) da construção mais esquemática.

Após essa primeira reflexão geral envolvendo a relação gramática e uso, a partir de considerações acerca das construções quantificadoras do tipo *SNI de SN2*, no português, a análise dos dados das quatro microconstruções tomadas para estudo, de forma a compreender melhor como é a distribuição dessas na gramática da língua.

Analisando os dados

Os resultados para a construção *DET monte de SN*, retomados de Alonso et al. (no prelo) são apresentados na tabela 1.

Tabela 1: Análise colexêmica da microconstrução *DET monte de SN*

Word	Freq. in <i>Corpus</i>	Freq. in <i>Cx</i>	Expect. freq.	Collostructional strength
Gente	5323	46	0.0441	121.7839
Coisas_coisa	17475	38	0.1449	77.7477
Dinheiro	14177	7	0.1175	10.2862
Caras	724	4	0.0060	10.2812
Amigos	3611	5	0.0299	9.7245
Cretinos	22	2	0.0001	7.8007
Saquê	47	2	0.0003	7.1305
Bandas	1297	3	0.0108	6.6923
Besteiras	80	2	0.0006	6.6647
Mentiras	188	2	0.0016	5.9197

Fonte: ALONSO ET AL (no prelo).

Esta construção possui um total de 296 dados e 174 *types* diferentes instanciando o *slot* da construção. De acordo com os autores, os lexemas mais atraídos pela construção são *gente* e *coisa(s)* com forças colostrucionais bem altas, (121.783) e (77.747) respectivamente, que parecem ser itens bem gerais e indefinidos. É o que se pode ver nos exemplos a seguir:

- (1) Esse tipo de trabalho é ótimo, pois você sente que tem **um monte de gente** no mundo inteiro preocupado com as mesmas coisas que você», diz. (ALONSO et al., no prelo)
- (2) **Um monte de coisas** acontecem aqui, mas são muito bem camufladas. (ALONSO et al., no prelo)

Além desses lexemas, os autores ainda apresentam as forças colostrucionais dos itens: *dinheiro* (10.286), *caras* (10.281), *amigos* (9.724), *cretinos* (7.800), *saquê* (7.130), *bandas* (6.692), *besteiras* (6.664) e *mentiras* (5.919).

Conforme apresentado no trabalho de Alonso et al (no prelo), a microconstrução *DET monte de SN* parece apresentar maior generalidade, pois recruta itens mais genéricos, o que significa que essa construção não recruta itens de um grupo semântico mais específico, que compartilhem entre si um traço ou propriedade, como por exemplo poder ser empilhado. Também não observamos coerência semântica entre os itens recrutados e o sentido de *monte* como formação geológica – indicando que a microconstrução *DET monte de SN* já parece bem acomodada à função quantificadora e que, portanto, o nome quantificador (*monte*) não interfere no tipo de lexema que é atraído pela construção.

No caso de *monte*, trabalhos anteriores como o de Fumaux (2018) demonstraram que esse nome preferencialmente se combina, dentro de uma estrutura binominal do tipo *SN1 de SN2*, nas primeiras fases de seu uso no português, com itens que possam ser empilháveis ou amontoados, em um formato que remete à forma de um *monte*⁹. Contudo isso não se confirma nos resultados do século XX, o que é também consoante ao trabalho de Fumaux (2018), que atesta a construcionalização da microconstrução *DET monte de SN*, e demonstra que não há a recuperação do sentido literal de *monte*, que se tornou bastante opaco na construção, recrutando, assim, itens que não se conformam à semântica de formação geológica e que nem mesmo precisam estar empilhados ou aglomerados.

Os lexemas mais atraídos pela microconstrução *DET monte de SN* – a saber, *gente e coisas* – Considerando os resultados da tabela 1, é possível concluir que a microconstrução *DET monte de SN* passou, ao longo do tempo, a preferir nomes de sentido +genérico. A generalidade dos itens é evidenciada também por nomes como *dinheiro, caras, amigos, cretinos*, assim como por itens mais abstratos do tipo *besteiras e mentiras*.

A seguir, discutiremos o resultado da análise colexêmica para *DET montanha de SN*. Observemos a tabela a seguir:

⁹ Como no exemplo: “Havia **um considerável monte de cobre e alguma prata**, no meio da mesa, e montes parciais, mais ou menos bem providos, ao lado de cada jogador” FUMAUX (2018).

Tabela 2: Análise colexêmica da microconstrução *DET montanha de SN*

Palavra	Frequência no <i>Corpus</i>	Frequência na construção	Frequência esperada	Força Colostrucional
1- Dinheiro	15.008	5	0,021789	10.654
2- Papel (is)	9.859	4	0,013255	8.950
3- Recursos	13.127	3	0,017649	6.071
4- Conexões e chips	1	1	0,000001	5.871
5- Cartas	1994	2	0,002681	5.454
6- Páginas	3	1	0,000004	5.394
7-13 trilhões	3	1	0,000004	5.394
8- Dívidas	2275	2	0,003058	5.340
9- Anáguas	6	1	0,000008	5.093
10- Dólares	3193	2	0,004293	5.045
11- Malfeitores	13	1	0,000017	4.757
12- Livros	4523	2	0,006081	4.743
13-Quinquilharias	43	1	0,000057	4.237
14- Culpas	71	1	0,000095	4.020
15-Diagnósticos	95	1	0,000127	3.893
16- Figurantes	148	1	0,000198	3.701
17- Mentiras	197	1	0,000264	3.577
18- Músculos	248	1	0,000333	3.477
19- Sapatos	489	1	0,000657	3.182
20- Convites	540	1	0,000726	3.139
21- Promessas	671	1	0,000902	3.044

A microconstrução *DET montanha de SN* resultou em um total de 48 dados e 36 tipos diferentes de colocações. Ao compararmos esta microconstrução com a anterior, podemos perceber o quanto a primeira é mais produtiva, apresentando maior frequência de dados e de tipos. Considerando os dados apresentados na tabela 2, podemos observar que as maiores forças colostrucionais correspondem a colocações com os itens *dinheiro* (10.654) e *papel* (8.950), que apresentaram frequência esperada de 0,021789 e 0,013255 (considerando o número de vezes que elas aparecem no *corpus*), mas ocorreram 5 e 4 vezes, respectivamente.

Percebemos nos dados que a construção em foco parece estar relacionada mais fortemente a lexemas amontoáveis/ empilháveis (como *dinheiro*, *papel(is)*, *cartas*, *páginas*, *livros*, etc.) ou que podem ser aglomerados, reunidos física ou abstratamente (como *recursos*, *dívidas*). Vimos, então, que de um total de 21 dados com força colostrucional mais altas, 10 correspondem a itens que se acomodam, em algum grau, ao sentido de *montanha*, já que recuperam deste a ideia

de verticalidade e de ajuntamento, por exemplo. São eles: *dinheiro, papel, cartas, páginas, 13 trilhões* (dinheiro), *anáguas, dólares, livros, quinquilharias* e *sapatos*. Vejamos os exemplos a seguir:

- (3) *F940411-039-537*: No ano passado, os anunciantes norte-americanos inundaram o país com a emissão de 310 bilhões de cupons. Se *empilhados*, formariam **uma montanha de papel** várias vezes maior do que o Everest. (*Corpus CHAVE*)
- (4) *F940320-017-161*: Semana que vem, volto ao assunto se não ficar *soterrada debaixo da montanha de páginas da Folha que os leitores certamente vão me enviar*. (*Corpus CHAVE*)

Os exemplos 3 e 4 ilustram que os lexemas recrutados pela construção apresentam maior coerência semântica com o nome quantificador na microconstrução do que aqueles instanciados pela construção *DET monte de SN*. O nome quantificador *montanha* parece ser menos opaco do que o nome *monte* para o falante, já que, no primeiro exemplo, com o uso de “se empilhados” ao falar de “papel” e, no segundo exemplo, com o uso de “soterrado”, o que o aproxima da ideia de *montanha*. Logo, essa construção tenderá a privilegiar referentes que guardam traços associáveis a uma formação geológica vertical decorrente de ajuntamento de uma dada matéria-prima.

A diferença apontada na comparação das microconstruções comentadas até o presente momento não implica que ambas não possam recrutar o mesmo item. Ou seja, estamos considerando que, embora preferencialmente elas se combinem com itens que compartilhem traços semânticos diferentes, esses traços podem ser opacificados em determinados contextos, em que a função quantificadora ganha proeminência (esteja mais saliente no discurso) sobre os traços mais específicos dos itens. Há, portanto, espaço para contextos de variação entre as microconstruções estudadas.

Uma questão que pode surgir ao observarmos a tabela 2 dá conta do fato de alguns lexemas possuírem força colostrucional alta, ainda que apareçam apenas uma vez na construção, no *corpus* selecionado – caso de *conexões* e *chips*. Certamente, é sempre preciso avaliar todos os fatores considerados na análise colostrucional (por exemplo, além da força colostrucional, é necessário confrontá-la com o grau de confiança gerado para cada combinação) e também relativizar resultados, dada a abrangência do *corpus* e número de dados obtidos e computados pelo programa.

Os itens *conexões* e *chips*, por exemplo, apareceram uma única vez no *corpus*, no contexto da microconstrução *DET montanha de SN*. Esse fato resulta em uma força de atração alta – já que a palavra, que não é recorrente no *corpus*, só apareceu quando instanciou a construção. Por outro lado, resultados com essa característica, precisam ser tratados com cautela e, preferencialmente, reavaliados em momento posterior (olhando outros fatores dentro do mesmo

corpus ou ampliando o *corpus* analisado), pois não é possível afirmar, com segurança, que a construção atrai um determinado item, quando não se tem ocorrências suficientes desse item no *corpus* para concluir que não foi uma combinação isolada, ocasional (para mais detalhes, consultar LEVISHINA, 2015).

Veremos a seguir os dados de *DET enxurrada de SN*.

Tabela 3: Análise colexêmica da microconstrução *DET enxurrada de SN*

Palavras	Freq. no <i>Corpus</i>	Freq. na construção	Freq. esperada	Força colostrucional
1- <i>Dólares</i>	3074	21	0,011796	60.936
2- <i>Dinheiro Externo</i>	91	7	0.000341	28.071
3- <i>Ações</i>	6274	7	0.024076	15.109
4- <i>Recursos Externos</i>	265	3	0.001016	9.771
5- <i>Investimento Estrangeiro</i>	16	2	0,000061	8.775
6- <i>Votos</i>	4495	4	0,072498	8.458
8- <i>Capital</i>	9745	4	0,037397	7.120
9- <i>Representações</i>	286	2	0.001097	6.225
10- <i>Reclamações</i>	687	2	0.002636	5.463
11- <i>Pieguices e Cafonices</i>	1	1	0,000003	5.415
12- <i>Estatizações</i>	1	1	0,000003	5.415
14- <i>Fotos</i>	1	1	0,000003	5.415
15- <i>Processos</i>	1	1	0,000003	5.415
16- <i>Pedidos de Aposentadoria</i>	1	1	0,000003	5.415
17- <i>Cartas e Telefonemas</i>	1	1	0,000003	5.415
18- <i>Comedias e Filmes</i>	1	1	0,000003	5.415
19- <i>Músicas</i>	1	1	0,000003	5.415
20- <i>Vales Transportes</i>	1	1	0,000003	5.415

Coletamos 137 dados da construção *DET enxurrada de SN* e 95 *types* diferentes no *corpus* analisado. Identificamos forças de atração bem altas de alguns lexemas, tais como *dólares* (60.963), *dinheiro externo* (28.071), *ações* (15.109). Esses itens, assim como *ações*, *recursos externos*, *investimento estrangeiro*, *capital* e *estatizações* parecem compor uma categoria semântica relacionada a capital, economia etc. Muito provavelmente, esse resultado tem relação direta, entre outros fatores, com o *corpus* jornalístico tomado para análise.

A semântica de *enxurrada* implica movimento, deslocamento físico na direção horizontal. Desta forma, é coerente o uso dessa microconstrução preferencialmente em contexto nos quais se fala de coisas que oscilam, como valor de uma moeda estrangeira ou que são deslocadas espacialmente, como os casos de entrada e saída do capital estrangeiro.

Portanto, assim como na microconstrução *DET montanha de SN* e diferentemente da microconstrução *DET monte de SN*, a semântica do item que ocupa o primeiro *slot* aberto – nesse caso, *enxurrada* – parece influenciar mais fortemente o tipo de item mais atraído pela construção, e, conseqüentemente, parece ser mais recuperável e transparente para os falantes (maior composicionalidade).

Embora, como já dito anteriormente, o *corpus* escolhido para análise tenha levado, provavelmente, ao recrutamento de itens da área econômica (como dólares, por exemplo), acreditamos que a ideia de movimentação é preferida pela construção e que provavelmente será observada com a ampliação futura do *corpus*. Percebemos que a construção costuma aparecer em contextos com verbo de movimento, como *entrar* e *frear*, nos exemplos abaixo:

- (5) Paralelamente a estes empréstimos às exportações, há **uma enxurrada de dólares** que *entra* no mercado brasileiro exatamente para lucrar com as incríveis taxas de juros reais (*Corpus CHAVE*)
- (6) F940930-027-272: Esses bancos apostam na vitória de FHC, mas não sabem que medidas poderão ser tomadas depois do pleito para *frear* eventual **enxurrada de dólares**. (*Corpus CHAVE*)

A semântica de *enxurrada* como grande volume de água em um contexto geralmente negativo, por ser frequentemente devastador, também pode ser evidenciado em alguns dos contextos de uso encontrados para essa microconstrução. Veja-se o exemplo a seguir, em que há um “jogo linguístico” com *águas de janeiro*:

- (7) F940116-118-1057: Graças à Copa São Paulo de Juniores, que, no gramado, é um encanto; fora dele, uma tragédia. Será possível que teremos todo ano de viver esse tormento? Junto com as *águas de janeiro*, **essa enxurrada de incompetência**, do lado da organização do evento, e de selvageria, do lado dos torcedores. (*Corpus CHAVE*)

Em (7), a associação com *águas de janeiro*, mês marcado, em algumas regiões do Brasil, ano após ano, por um período de chuvas fortes e grandes desastres delas decorrentes, é recuperado para se falar da grande incompetência por parte dos organizadores da Copa São Paulo de Juniores. O contexto negativo se estenderia também à avaliação da selvageria dos torcedores. Nesse caso, aparentemente, parte da construção foi suprimida.

Percebemos, então, até o momento, que a microconstrução *DET monte de SN* recruta preferencialmente nomes de referência genérica, enquanto *DET montanha de SN* prefere lexemas cujos referentes admitem ser amontoados ou empilhados, e *DET enxurrada de SN* recruta itens cujos referentes preferencialmente sofram algum tipo de transferência, oscilação, embora contextos de avaliação negativa também possam ser recrutados, quando aparentemente são devastadores. Veremos, então, a seguir, os resultados da análise colexêmica para *DET chuva de SN*.

Tabela 4: Análise colexêmica de *DET chuva de SN*

Palavra	Freq. no <i>Corpus</i>	Freq. na construção	Freq. esperada	Força Colostrucional
1- <i>Pétalas de Rosa</i>	13	3	0,000016	15.242
2- <i>Papel/Papel picado</i>	7641	4	0,009845	9.468
3- <i>Denúncias</i>	569	2	0,000733	6.581
4- <i>Ovos/Ovos E Tomates</i>	714	2	0,000920	6.383
5 - <i>Flores</i>	719	2	0,000926	6.377
6- <i>Cuspes E Xingamentos</i>	1	1	0,000001	5.889
7- <i>Cartas</i>	1994	2	0,002569	5.491
8- <i>Criticas</i>	2963	2	0,003817	5.148
9- <i>Bolhas de Sabão</i>	6	1	0,000007	5.111
10- <i>Canivetes</i>	23	1	0,000029	4.528
11- <i>Flechas</i>	37	1	0,000047	4.321
12- <i>Regalias</i>	48	1	0,000061	4.208
13- <i>Farpas</i>	77	1	0,000099	4.003
14- <i>Balões</i>	133	1	0,000171	3.766
15- <i>Capitais Externos</i>	146	1	0,000188	3.725
16- <i>Fogos de Artifício</i>	156	1	0,000201	3.696
17- <i>Escândalos</i>	177	1	0,000228	3.641
18- <i>Paulo Coelho</i>	191	1	0,000246	3.608
19- <i>Cinzas</i>	201	1	0,000258	3.586
20- <i>Medalhas</i>	232	1	0,000298	3.524
21- <i>Camisinhas</i>	325	1	0,000418	3.378
22- <i>Estrelas</i>	410	1	0,000528	3.277
23- <i>Correspondências</i>	678	1	0,000873	3.058
24- <i>Referências</i>	775	1	0,000998	3.000

Encontramos no *corpus* estudado 46 dados e 35 tipos para *DET chuva de SN*. As maiores forças de atração para *DET chuva de SN* são com os lexemas *pétalas de rosa* (15.242) e *papel/papel picado* (9.468), seguidas de denúncias (6.581), *ovos/ovos e tomates* (6.383) e flores (6.377). Como se vê, os itens preferencialmente recrutados compartilham o traço mais contável, mantendo a ideia de um conjunto de partes individualizadas, com alta coerência com a ideia de *chuva*. Nesse mesmo sentido, observamos que há uma prevalência dos dados em contexto de movimento – assim como ocorreu com *DET enxurrada de SN* –; contudo, uma movimentação com a direção de cima para baixo na maior parte dos dados, em conformidade com o movimento da chuva.

Algumas colocações (a exemplo daquelas que têm nomes como *pétalas de rosa, papel/papel picado, ovos/ovos e tomates, flores* preenchendo o segundo slot aberto) foram incluídas nos dados, no momento da análise colexêmica, ainda que esses dados sugiram apenas uma relação imagética com *chuva*, não caracterizando um típico exemplo da construção. Essa decisão, que não se fez necessária nos casos analisados para as microconstruções anteriores (porque não ocorrerão casos desse tipo), mostrou que, de fato, a microconstrução *DET chuva de SN* tem um comportamento um pouco diferente das demais, uma vez que, quando deliberadamente incluímos dados com inferência de quantidade na análise quantitativa, percebemos que eles são, inclusive, os preferidos em termos colocacionais. Portanto, compreendemos, que, embora se possa dizer que *DET chuva de SN* é uma construção na língua portuguesa, como se pode ver em exemplos como (10), a seguir, há ainda um maior conjunto de dados de uso em que a noção de ideia de quantidade parece ser inferida, enquadrando-se como contextos críticos (cujo impacto em termos do uso de *DET chuva de SN* pôde ser capturado quando inserimos esses dados na análise estatística).

Vejamos os exemplos abaixo:

- (8) *F940115-136-1018*: Nas imagens seguintes vê-se **uma chuva de estrelas**, na verdade bombas que em segundos transformam a cidade num deserto. (*Corpus CHAVE*)
- (9) *F940831-060-632*: Três helicópteros carregados com uma tonelada de pétalas de rosas de todas as cores despejaram ontem **uma chuva de flores** em frente ao Masp para divulgar a abertura da 13ª Expoflora, que acontece hoje na Cooperativa Agropecuária Holambra. (*Corpus CHAVE*)

Nesses exemplos, é possível observar que os itens recrutados configuram-se como unidades individuais, em uma provável associação aos pingos da chuva, em um movimento descendente, como o que se vê em (8), com as estrelas (bombas) que caem sobre a cidade e, em (9), quando pétalas de flores são jogadas do alto, a partir de um helicóptero. Em termos da coerência semântica entre item e *DET chuva de SN*, parece muito proeminente a configuração de referentes individualizados e um contexto de movimento descendente, ambos remetendo claramente a traços da semântica de chuva.

Percebemos, quando analisados mais detidamente os casos dessa microconstrução, contextos que geram inferência negativa, a partir de um grupo semântico formado por lexemas como: *xingamentos, denúncias, críticas, regalias, farpas e escândalos*. Esses dados refletem um viés negativo que remonta à metáfora RUIM É PRA BAIXO (LAKOFF; JOHNSON, 1980), como o que se pode ver no exemplo a seguir:

- (10) *F940324-182-1630*: Você dificilmente vai ver um outro filme protagonizado por Sharon Stone e William Baldwin. Os dois brigaram tanto durante as filmagens que

o diretor Phillip Noyce e o roteirista Joe Eszterhas tiveram que lembrá-los da multa que pagariam caso interrompessem os trabalhos. As filmagens continuaram, mas sob **uma chuva de farpas**, principalmente da parte de Stone, que reclamava da falta de cavalheirismo de Baldwin (e de seu mau hálito).

Considerando os resultados apresentados nessa seção, percebemos que, embora as quatro microconstruções estudadas funcionem na língua para quantificar referentes em termos de grande quantidade/ volume, cada microconstrução apresenta suas próprias características, evidenciadas pelas preferências colocacionais obtidas por meio das análises colexêmicas realizadas. Reforçamos, com isso, a aposta de que as quatro microconstruções são funcionalmente equivalentes, mas semanticamente distintas.

Considerações finais

A partir das análises de um conjunto de dados de escrita, procuramos entender as preferências colocacionais de cada microconstrução selecionada, ao realizarmos um estudo para verificar os itens mais atraídos por elas. A análise colexêmica permitiu *insights* interessantes acerca da interferência do item quantificador (*monte, montanha, enxurrada e chuva*) na escolha por uma ou outra microconstrução por parte dos falantes do português, demonstrando uma relação clara entre gramática e uso. Nesse sentido, observou-se o grau de coerência semântica entre os elementos quantificador (SN1) e quantificado (SN2) da construção, dentro do contexto de microconstruções quantitativas do tipo *SN1 de SN2*.

Sendo assim, ao retomarmos os resultados de outros trabalhos sobre *DET monte de SN* (FUMAUX, 2018, entre outros) e ao compará-los com os resultados das análises das outras microconstruções, percebemos que *DET monte de SN* é preferencialmente combinado com itens de referência mais genérica, apresentando pouca transparência em relação à semântica do item recrutado para o *slot* de N1 (quantificador) da construção e, associada a isso, baixa coerência semântica entre esse item e o referente expresso no segundo *slot* aberto.

Por outro lado, o resultado das análises colexêmicas realizadas na presente pesquisa revelam que as microconstruções *DET montanha de SN*, *DET chuva de SN* e *DET enxurrada de SN* possuem uma maior coerência semântica com os lexemas recrutados, apresentando, dessa forma, uma menor opacidade semântica do item recrutado para o primeiro *slot* aberto da construção, em comparação com o que se observou para a microconstrução *DET monte de SN*.

No caso de *DET montanha de SN*, há uma preferência por itens que possam ser empilhados ou aglomerados, enquanto *DET chuva de SN* privilegia contextos que envolvem movimento de cima para baixo, envolvendo referentes contáveis tomados de forma individualizada. Especificamente com relação a *DET chuva de SN*, foram encontrados contextos com inferência negativa, o que pode ser explicado pela metáfora RUI É PRA BAIXO (LAKOFF; JOHNSON,

1980). No caso de *DET enxurrada de SN*, embora envolva uma clara ideia de movimento e deslocamento, parece haver uma menor literalidade em relação ao movimento de referentes, em comparação ao que temos para *DET chuva de SN*.

Atestamos, assim, que essas microconstruções, embora pareçam sinonímias perfeitas, em um primeiro momento, apresentam graus de composicionalidade distintos, considerando o grau de coerência entre os referentes que ocupam os slots de N1 e N2 da construção, bem como fatores que extrapolam uma análise colocacional e levam em consideração, por exemplo, contextos específicos, como o de deslocamento – o que pôde ser visto nas microconstruções *DET enxurrada de SN* e *DET chuva de SN*).

Referências

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)

BARLOW, M.; KEMMER, S. (orgs.). *Usage based models of language*. Stanford, California: CSLI Publications, 2000.

BYBEE, J. From usage to grammar: the mind's response to repetition. *Language*, v.82, n.4, p. 711-733.

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

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

BYBEE, J.; BECKNER, C.; BYBEE, J. A. A usage-based account of constituency and reanalysis. in: Ellis, n. c.; Larsen-freeman, d. (eds.). *Language as a complex adaptive system*. boston, ma: wiley-blackwell, 2009. p. 27-46.

BOYLAND, J. T. Usage-based models of language. In: EDDINGTON, D. (ed.). *Experimental and Quantitative Linguistics*, Munich: Lincom, 2009, p. 351-419.

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

ELLIS, N. C.; LARSEN-FREEMAN, D. (eds.) *Language as a Complex Adaptive System*. *Language Learning*, n. 59, Suppl. 1, 2009.

FRANCIS, E. J., MICHAELIS, L. A. (orgs.). *Mismatch: form-function incongruity and the architecture of grammar*. Stanford, CA: CSLI Publications, 2003.

FUMAUX, N. C. A. *Construcionalização de DET monte de SN: uma abordagem centrada no uso*. Dissertação (Mestrado em Linguística) do Programa de Pós-graduação em Linguística, Faculdade de letras, UFRJ, Rio de Janeiro, 2018.

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

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

GRIES, S.; STEFANOWITSCH, A. Extending collostructional analysis. A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9:1, 2004, p. 97-129.

HILPERT, M. *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press, 2014.

HOPPER, P. J. Emergent Grammar. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society, 1987, pp. 139-157.

LAKOFF, G., JOHNSON, M. *Metaphors we live by*. Chicago/London: University of Chicago Press, 1980

LEVISHINA, N. *How to do Linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins, 2015.

PIERREHUMBERT, J. Exemplar dynamics: Word frequency, lenition and contrast. In: BYBEE, J.; HOPPER, P. (ed.) *Frequency effects and the emergence of linguistic structure*, p. 1-19. Amsterdam: John Benjamins, 2001.

PEREK, F. *Argument structure in usage-based construction grammar: Experimental and corpus-based perspectives*. Amsterdam: John Benjamins, 2015.

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. In: *Studies in Language*. N. 39. John Benjamins Publishing Company, 2015.