

**QUANTIFYING BINOMINAL CONSTRUCTIONS IN PORTUGUESE
AND RUSSIAN: THE CASE OF *UM MONTE DE NP*
AND *KUCHA NPGEN***

CONSTRUÇÕES BINOMINAIS QUANTIFICADORAS EM PORTUGUÊS E RUSSO:
O CASO DE *UM MONTE DE SN* E *KUCHA SNGEN*

Karen Sampaio Braga Alonso
Lattes | karensampaio@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Gabrielle de Figueira do Nascimento
Lattes | gfigueira.nascimento@gmail.com
Universidade Federal do Rio de Janeiro

Diego Leite de Oliveira
Lattes | diegooliveira@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Thiago Moreira da Silva
thiago.letras.ufrj@gmail.com
Universidade Federal do Rio de Janeiro

Nuciene Caroline Amphilóphio Fumaux
Lattes | carol.fumaux@gmail.com
Universidade Federal do Rio de Janeiro

Abstract: This paper reports on a contrastive study of Brazilian Portuguese *um monte de NP* and Moscow Standard Russian *kucha NPgen* quantifying binominal constructions, under a usage-based constructionist approach. To identify semantic specificities in both constructions, an empirical investigation based on collostructional analysis was conducted. The results indicate that despite some similarities between Portuguese and Russian in terms of cognitive operations involved in the conceptualization of the binominal quantifying constructions *um monte de NP* and *kucha NPgen*, there are specificities concerning each construction in each language in terms of semantic compatibility of lexemes co-occurring in them.

Keywords: Quantifying binominal constructions. Usage-based construction grammar. *Monte. Kucha.* Collostructional analysis.

Resumo: Este trabalho consiste em um estudo contrastivo das construções binominais quantificadoras *um monte de SN*, do português brasileiro, e *kutchá SNGen*, do russo, em uma perspectiva construcionista baseada no uso. Para identificar especificidades semânticas em ambas as construções, foi realizada uma investigação empírica baseada em análises colostrucionais. Os resultados indicam que apesar de algumas similaridades entre português e russo em termos de operações cognitivas envolvidas na conceptualização das construções binominais quantitativas *um monte de SN* e *kutchá SNGen*, há especificidades relativas a cada construção em cada língua no que se refere à compatibilidade dos lexemas que nelas ocorrem.

Palavras-chave: Construções binominais quantitativas. Gramática de construções baseada no uso. Monte. Kutchá. Análise colostrucional.

1. Introduction¹

One of the most interesting phenomena in natural language concerns the ability of some linguistic expressions to convey new meanings and/or functions over time in several domains. A number of scholars under the grammaticalization framework have already demonstrated paths of change ranging from lexicon to grammar. In order to cite one example, Bybee (2010, p. 199) supports such lexicon-to-grammar paths through a diachronic analysis of periphrastic future with *will* in English, reinforcing the importance of crosslinguistic comparison to understand mechanisms underlying language change.

The fact that (new) forms may be associated to new functions is observed in several domains, including quantification manifested in binominal constructions (TRAUGOTT, 2008; ALONSO, 2010; BREMS, 2012; VERVECKKEN, 2015; RAKHILINA, 2009, among others). A large group of nouns within quantifying binominal constructions – such *monte* ('heap'), *chuva* ('rain'), in Portuguese, or *kutchá* ('heap'), *more* ('sea'), in Russian – could be included in the class of quantifiers. Thus, a point to be considered is that not only classical quantifiers (such as *liter*, *kilo*, etc.) can be part of the fixed portion of such a construction.

Comparison among languages demonstrates that it is possible to make use of concrete referents to talk about abstract concepts that are related to subjective judgements of quantity (in terms of small or large quantity). That is the case of *monte* ('heap') which shows up within a binominal construction, i. e., *um monte de NP* to indicate 'a big amount

¹ This paper congregates the main preliminary results of a more comprehensive research project regarding binominal constructions in different world languages, with special attention to Brazilian Portuguese.

of something' in contemporary Brazilian Portuguese, a Romance language, as well as *kucha* ('heap') which was recruited in Russian, a Slavic language, as the fixed element instantiating the quantifying binominal construction *kucha NPgen*.

In this work, we are interested in the comparability between functionally similar constructions in Brazilian Portuguese and Moscow Standard Russian (hereinafter referred simply as Russian) and the crosslinguistic relevance of this analysis to a usage-based framework. We assume that, first, the comparison among different languages can reveal cognitive motivations underlying surface linguistic structure and, second, binominal constructions provide good evidences for that.

The main purpose here is to offer the results of a contrastive collocation analysis of quantifying binominal constructions in Brazilian Portuguese and Russian following Gilquin (2015), whose work demonstrated the feasibility of such methodology applied crosslinguistically. Differently from Gilquin (2015), our focus relies not only on constructions which are semantically equivalent at first glance, but on the comparison of constructions displaying a relevant degree of productivity² in each language. That seems to be the case of *um monte de NP* and *kucha NPgen*, since both, despite dissimilarities in terms of semantic traits, display high levels of entrenchment, with high extensibility and token frequency. In addition, while Gilquin is interested in academic writing, we focus on a more comprehensive type of writing such as the journalistic genres.

This paper is basically divided threefold: first, we will discuss collocation analysis family of methods under a usage-based constructionist approach; second: we will analyze Portuguese and Russian separately, focusing on the results of a collexeme analysis for *um monte de NP* (Portuguese) and *kucha NPgen* (Russian). Third, we will briefly show the results obtained by the distinctive collexeme analysis for *um monte de NP* and *kucha NPgen*.

2. Quantifying Binominal Constructions: a collexeme analysis

This section aims at providing a brief discussion of the quantifying binominal construction associated to the methods of collocation analysis often employed under a usage-based constructionist perspective on language. The next subsection discusses the way the quantifying binominal construction is understood under a usage-based constructionist perspective, which conceive language as a network of interrelated constructions, as form-function pairings. Subsection 2.2 shows how the specific collocation analysis methods of collexeme and distinctive collexeme analysis were applied in this investigation.

² For more details about productivity, see Barddal (2008).

2.1. Quantifying binominal constructions and the collostructional analysis methodology

Quantifying Binominal Constructions congregate a different set of form-meaning pairings coined in a number of languages as grammatical alternatives for quantifying referents. They frequently perform a figurative function due to the assumption that speakers' recruitment of non-quantity nouns to instantiate a quantifying construction reveals the nature of the cognitive processes (such as metaphor or metonymy) underlying the surface syntactic structure.

This analysis assumes a usage-based perspective on language and the understanding of grammar as a construction network (see GOLDBERG, 1995, 2006, 2019; DIESSEL, 2019). Since it is a functional-cognitive approach, it points out that the close relationship between semantic-pragmatic and morphosyntactic aspects of the language units will be considered here.

Assuming a constructionist approach to language, we will follow the network configuration which can be found in the background illustrated in the figure 1 below, quoted from Traugott and Trousdale (2013, p. 17) as reference for the levels of constructions schematicity.

Figure 1 – Gradience of hierarchic relationships among constructions

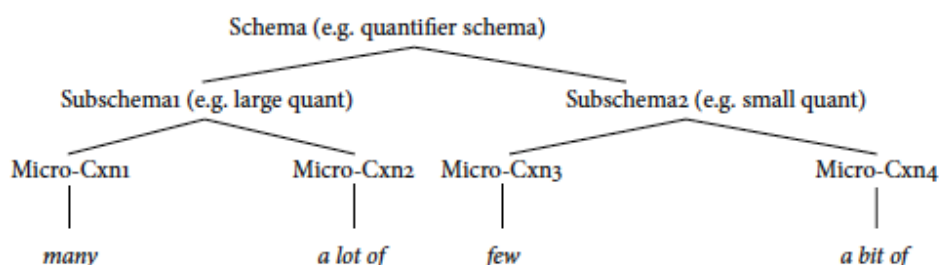


Figure 1 brings out the quantifier schema to English, reinforcing the constructionist crucial argument in favor of the arrangement of constructions in a network. Thus, it illustrates nodes and links among quantifying constructions, focusing on taxonomic links (see GOLDBERG, 1995 and DIESSEL, 2015, 2019 for a more detailed range of constructions links). In this paper, both Brazilian Portuguese and Russian analysis address selected binominal micro-constructions (e.g. *um monte de NP*, for Portuguese, and *kucha NPgen*, for Russian) assumed as nodes linked to the sub-schema 1 (large quantity).

Considering that both [*um monte de NP*] and [*kucha NPgen*] constructions display a fixed slot – *um monte de* and *kucha*, respectively – and a schematic slot requiring

NPs without specification on declension (as in Portuguese) or NPs inflected in genitive case (as in Russian), one of the researcher's tasks is to identify what kinds of lexemes can fill the NP schematic slot in each language. According to the general principles of Construction Grammar approach, the candidates to fill the open slot in the quantifying binominal construction would exhibit some kind of semantic compatibility to the construction, that is, a word may occur in a given construction if its meaning is compatible with the meaning of the construction (cf. STEFANOWITSCH; GRIES, 2003, p. 99). Thus we would expect that in *um monte de NP* and *kucha NPgen* constructions, in addition to quantifiable referents, would attract lexemes whose referents could be arranged in large quantity so that they form a concrete or abstract (metaphorical) heap. In the case of *um monte de NP* construction, due to its development in the history of Brazilian Portuguese, resulting in an expressive semantic expansion of the construction regarding the semantics of the candidates to fill the open slot (cf. FUMAUX, 2018), a question may be raised: what are the collocational preferences of *um monte de NP* construction, considering this semantic expansion. The identification of collocational preferences may provide some insights on the status of the construction in a given synchrony.

A crosslinguistic analysis of the quantifying binominal construction mentioned above may reveal interesting points concerning quantification in language and the cognitive processes which it involves. Based on this assumption, we use quantitative methods applied to corpus analysis, which may constitute a powerful tool in studies under a usage-based constructionist perspective.

2.2. Collostructional analysis as a feasible method to understand constructions' semantic properties

In this paper, in order to understand some specific semantic aspects of the constructions under analysis in Brazilian Portuguese and Russian, we have used collostructional analysis – a family of association measures methods, including: (i) a simple collexeme analysis, which measures the force of attraction of lexemes to one slot in a given construction; (ii) a distinctive collexeme analysis, which measures the force of attraction of lexemes to one slot in two or more constructions; (iii) a covarying collexeme analysis, which measures the force of attraction of lexemes to two slots in one construction (see GRIES; STEFANOWITSCH, 2004, for more details).

In order to analyze the force of attraction of lexemes to *the um monte de NP* construction in Brazilian Portuguese and to *kucha NPgen* in Russian, we have recurred to a

simple collexeme analysis for each language separately. Likewise, to compare the force of attraction of lexemes in both constructions, a distinctive collexeme analysis was applied. The covarying collexeme analysis was excluded from this research. According to Gries and Stefanowitsch (2004), in constructionist approaches the combinations of lexemes occurring in construction slots is conceived in terms of semantic compatibility. For the authors:

A word may occur in a given construction if its meaning is compatible with the meaning of the construction; it may 'alternate' between two constructions if (or to the degree that) the word's meaning is compatible with meanings of both constructions. In the context of alternating pairs, a focus on constructional semantics and semantic compatibility raises several questions: first, what exactly are the (often seemingly tenuous) semantic differences between the members of such a pair; second, how productive is the 'alternation' in actual usage, i.e. which verbs/nouns occur freely in both constructions and which have strong biases towards one of them and third, is a constructional, non-derivative approach plausible given the answers to the first two questions. (GRIES; STEFANOWITSCH, 2004, p. 99).

Considering the questions raised by Gries and Stefanowitsch (2004), we seek to understand which lexemes are compatible with the constructions studied in this research, in order to understand a little bit more about these constructions as pairing of form and meaning. Also, we aim to identify similarities and differences between them. In order to do that, first we proceed to a simple collexeme analysis carried out separately for *um monte de NP* as well as *kucha NPgen* constructions. This kind of analysis considers (i) the size of the corpus (in number of words); (ii) the token frequency of the construction in that corpus; and (iii) the token frequency of each lexeme in the construction in order to obtain the force of attraction of each lexeme in each of the analyzed constructions (for more details, see GRIES; STEFANOWITCH, 2004; GRIES, 2014).

In order to make a feasible comparison between the two languages, we had controlled the sample data for each one. As this is a preliminary study aiming at verifying the compatibility of a collostructional analysis to the object being investigated, we started from a limited sample. As for Brazilian Portuguese data, we have considered a 35,699,765 word subcorpus of journalistic writing sampled from the basis of Linguateca (www.linguateca.pt), covering data from 1994 to 1995 of Folha de São Paulo newspaper. By this analysis, we have collected a total of 296 instances of *um monte de NP*. As for Russian data, we have considered a 33,547,720 words subcorpus of journalistic writing custom-

ized on the basis of Russian National Corpus (available at www.ruscorpora.ru), covering data from 2007 to 2010 of Trud newspaper. 391 instances of *kucha* NPgen were collected.

In the analysis, the R software for statistical computing was employed in its 3.5.3 version, as well as the R Studio, a more “user friendly” software, in its open source version, available at www.rstudio.com/products/rstudio/. After downloading both softwares, a collexeme analysis was performed, using R Studio software and the script (<http://www.stgries.info/teaching/groningen/readme.txt>) made available by Stefan Gries at <http://www.stgries.info/teaching/groningen/readme.txt>. We ran the program providing the required information specified above, that is, the subcorpus total number of words, the token frequency of the construction under analysis and the token frequency of each lexeme co-occurring within the construction (the token frequency is provided by elaborating an Excel file with a 2x2 table indicating the number of times that each lexeme is found in the construction and the overall occurrences of each lexeme in the corpus. After processing the collected data, the software provides the colostruational strength of each lexeme in relation to the analyzed construction. The greater the colostruational strength, the greater the level of lexeme attraction for construction. When the analysis is performed, a Fisher’s exact test – a statistical significance test for contingency tables – is applied, in order to verify the statistical significance of the obtained results. The significance p-value is provided for the whole contingency table independently of the number of rows³.

After performing the collexeme analysis for each construction, we proceed to a distinctive collexeme analysis of both constructions in Brazilian Portuguese and Russian, which considers their token frequency and the lexemes occurring within them. As in this paper we are comparing similar constructions in two different languages, we followed the methodological steps proposed by Gilquin (2015) to the analysis of causative constructions with *make* in English and *faire* in French. So, we had created an Excel file containing each token of [um monte de NP] construction and each token of [kucha NPgen] construction. To have a common basis for comparison, we used the translations of Russian lexemes to Portuguese as input for the analysis. The distinctive collexeme analysis considers only the token frequency of each construction and the token frequency of each lexeme occurring in each construction. As in collexeme analysis, the distinctive collexeme analysis provides significance p-value for the whole contingency table independently of the number of rows.

In the next section we provide a brief description of quantifying binominal constructions in Portuguese and Russian, with special emphasis for the *um monte de NP* and

³ In order to know more about Fisher’s exact test of significance, see Levshina (2015).

kucha NPgen constructions, presenting a separate simple collexeme analysis for both constructions.

3. Quantifying binominal constructions in Brazilian Portuguese and Russian

In this section, we provide a description of binominal constructions in Portuguese and Russian, giving special attention to *um monte de NP* construction in Brazilian Portuguese and *kucha NPgen* in Russian, which yield a large number of instances in each of the respective languages. In addition, the results of a simple collexeme analysis for Portuguese and Russian are presented and discussed.

3.1. The case of Brazilian Portuguese *um monte de NP*

Binominal constructions like NP1 of NP2 can be used to quantify referents. To briefly discuss that, take the examples below, which illustrate some usages directly related to quantifying binominal micro-constructions in Portuguese:

- (1) *Comprei uma dúzia de ovos.*
buy a dozen of eggs
'I bought a dozen eggs.'
- (2) *Acabou de passar uma dúzia de crianças ali.*
finish of pass a dozen of children there
'Just passed a dozen children there.'
- (3) *Terá um monte de gente na rua amanhã.*
have a heap of people in the street tomorrow
'There will be crowds of people on the street tomorrow.'

Considering the semantics of the examples listed above, a question that arises is about its source. We could say that the source of quantity interpretation of these examples is basically twofold: conventionalized usages whose quantifying meaning is expressed by a more canonical quantifier (examples 1 and 2); conventionalized usages whose quantifying meaning is obtained by a quantifying inference from a prior non-quantifying usage, as we see in example 3.

In examples 1 and 2, the noun *dúzia* ('dozen'), which refers to a very specific amount of something, fills *NP1* slot in the construction. Other quantifiers could be inserted in this set such as *grupo* ('group') which also implies necessarily quantity interpretation,

but not as precisely as *dúzia* does, since groups can vary in relation to numbers of their members. Despite this difference, what all of those nouns fitting NP1 slot (i.e. *dúzia*, *quilo* and *grupo*) have in common is the fact that they can be considered typical quantifiers in Portuguese, since they can be associated to quantity meaning throughout all Portuguese constructions they can be fitted with.

Example number 2 can be interpreted in Portuguese in two different ways. Firstly, it means that exactly twelve children have passed there, and then it would be considered an instance of the quantifying binominal construction. Secondly, and apparently much more frequently, example 2 means that an indeterminate (usually, expressing hyperbolic sense to a certain amount or a rough quantity of something and presenting a stress on *NPI*) number of children have passed there. In this case, the meaning seems to be derived from an invited inference from the quantifying binominal construction that matches the first interpretation.

Example 3 differs from those just mentioned due to the fact that it did not function as quantifiers in the past synchronies of Portuguese. *Monte* ('heap') was not originally used within a quantity construction neither functions as quantifier primarily but allows a quantity inference towards a metaphorical process. Portuguese, as many other languages, seems to conventionalize quantification in terms of scales (such as *monte*) as a good way to express quantity meaning, including those usages within binominal constructions.

Indeed, Brazilian Portuguese offers a large group of non-canonical quantifiers showing up within quantifying binominal micro-constructions. As partially mentioned before, nouns such as *monte* ('heap'), *mar* ('sea'), *enxurrada* ('flash flood'), *nuvem* ('cloud'), *tempestade* ('storm') illustrate how diverse is the set of nouns which are entrenched in a quantifying binominal construction, such as *(Det) monte de NP*, *(Det) mar de NP*, *(Det) enxurrada de NP*, *(det) nuvem de NP*, *(Det) tempestade de NP*. Thus, the whole schema highlights the multi-word chunk, i.e., the part of the micro-construction which is often repeated and made up of a neuromotor routine (BYBEE, 2006).

Quantifying binominal constructions in Brazilian Portuguese have already been studied by other linguists (ALONSO, 2010; BRODBECK, 2010; FUMAUX, 2018; TAVARES, 2014). Based on a corpus analysis, Alonso (2010) presented some formal and semantic-pragmatic properties of Brazilian Portuguese quantifying binominal constructions. Considering her findings, the constructions to be analyzed here can be, to some extent, associated to the "multiplexing" process, defined by Talmy (2006) as an operation in which "an original solo referent is, in effect, copied onto various points of space or time" (TALMY, 2006, p. 79).

Taking the multiplexing process into account, usages such as *um grupo de pessoas* ('a group of people') are typical cases associated to it, since 'group' implies the idea of the existence of various copies of something. In this case, a certain number of people corresponding to a group, which refers necessarily to more than one person, probably more than three. In addition, *group* frequently occurs when *NP2* is a countable noun, what favors the multiplexing interpretation.

Alonso argued in favor of the idea that multiplexing process yields a large number of instances of the quantifying binominal construction, including those that express more subjective evaluations of quantity. Referents such as *monte* (lit. 'heap') implies quantity interpretation, which is construed more subjectively (individually).

Constructions vary in terms of degrees of entrenchment and, as far as quantifying binominal constructions are concerned, a good test to measure how entrenched the parts of the constructions are is to verify if they work separately in a given linguistic context. Considering examples such as *Ela comprou um quilo de açúcar* (lit. She bought a kilo of sugar/ 'She bought a pound of sugar'), *um* ('a') can be individually a good answer to the question *How many pounds of sugar did she buy?* The same is not true to examples like *Há um monte de cadeiras aqui* (lit. There is a heap of chairs here/ 'There are a lot of chairs here'), since considering the question *How many chairs are there here?*, the answer must be *um monte*. Separately, *um* as the answer to the former question creates an ungrammatical structure.

Specifically related to the subject of this section, Fumaux (2018) described the process of constructionalization (TRAUGOTT; TROUSDALE, 2013) of *um monte de NP* in Portuguese. In this case, differently from Figure 1, quoted from Traugott and Trousdale (2013, p. 17), the open slot – *NP* – is now being considered, because of the crucial role played by the second *NP* in the construction representation through the history of Portuguese. It bears pointing out that the architecture of an emergent grammar is shaped by use and then synchrony must not be sharply separated from diachrony. This line of investigation, then, seeks to explain how language use shapes grammar.

The discussion presented by Fumaux (2018) about the formation of the Portuguese quantifying binominal construction *um monte de NP* is guided by the following basic question: *How did um monte de NP become a construction through the history of Portuguese?* To answer the question just pointed out, let's analyze the examples below:

(4) *Um monte de Lisboa.*
A heap of (loc.) Lisbon
'A heap in Lisbon.'

(5) *Um monte de coisas.*
A heap of things
'A lot of things.'

On one hand, example 4 represents a very typical instance of the specifying interpretation of the binominal construction, in which *um monte de NP* leads to a more compositional interpretation, i. e., *a heap in Lisbon*. On the other hand, example 5 could not be interpreted in the same way. In this case, the meaning of the whole formation is not the sum of the meaning of each part of it, but the idea of much/many of something.

Indeed, based on corpus analysis, we could consider that initially the most prominent usage of *um monte de NP* leads to the interpretation that usages such as *um monte de Lisboa* meant, and, in fact, it still means a heap in Lisbon, that is, a heap belongs to Lisbon. Besides that, other usages of *monte* within the specifying binominal construction yielded metonymic inferences, such as *um monte de neve* (a heap of snow/ a lot of snow). Fumaux (2018) argued that inferences like that triggered the constructionalization process. Assuming these cases as the source of constructionalization, we should consider that in examples like that *monte* maintains the semantic traits of a heap (geological formation), and so the binominal construction recruits preferably nouns which make reference to natural elements, such as snow or sand, whose meanings are well accommodated to the meaning of the other parts of the construction, exhibiting, thus, a high level of compositionality.

Over the centuries, the syntagmatic string *um monte de NP* became gradually more frequent and increased its range of collocations with nouns from different natures instantiating *NP* slot. So, it undergoes increase of type frequency. In this case, what we observe from data analysis is a gradual decrease of compositionality and the formation of a chunk.

According to Fumaux (2018), *um monte de NP* specifying binominal construction has undergone a gradual loss of the occurrence of intervenient elements (affecting analyzability) within the construction, a gradual fixation of the indefinite article as the preferable determinant of *monte*, an increase of countable nouns instantiating the construction (second noun typically occurs in the plural) and a gradual change of the head of the construction. Consequently, a new construction emerges, i.e. a quantifying binominal construction *um monte de NP*.

Fumaux (2018) argued in favor of the emergence of a quantifying binominal construction – *um monte de NP* – adding a node in the schema of quantifiers in Portuguese. She points out that *um monte de NP* undergoes change in meaning (quantification) and form, considering that *monte* initially instantiated a more compositional construction [[NP] of NP] and then the syntagmatic string *um monte de NP* was neoanalyzed and *monte* came to be part of a new chunk – [um monte de [NP]]. The results found in Fumaux (2018) lead to the interpretation that *um monte de NP* is a quantifying construction which exhibits a high level of entrenchment and has at first glance few restrictions in terms of the lexemes it recruits.

Changing of head can be considered a crucial argument to argue for the emergence of a new construction (quantifying construction) and will be detailed in another text, being structurally supported by evidences such as the following: instances such as *um monte*_{masc.sg.} *de pessoas*_{pl.fem.} (lit. ‘a heap of people’ / ‘a lot of people’) are anaphorically recovered by *elas* (*them*_{pl.fem.}) instead of by *ele* (lit. ‘he’ / ‘it’); the whole string *um monte* can be substituted by a degree word like *muito* (‘much’); verb agrees preferably with *pessoas*, instead of with *monte*. However, there is no sufficient evidence to affirm that *um monte de NP* completely ran its turns yet.

Another potential argument, which was not provided by Fumaux’s results due to the fact that some quantifying micro-constructions entrenched referents such as *enxurrada* (flash flood), which apparently have been recruited by a more schematic quantifying binominal construction – [NP of [NP]] – since the unique interpretation yielded from that binominal construction instantiated by *enxurrada* was the quantifying one. We defend that the recruitment of *enxurrada* would have been motivated by analogy and enabled by the coercion of the already existent quantifying binominal construction upon the lexeme.

Although Fumaux’s research has already provided very good insights concerned to that subject, we believe that the methodology offered by collocation analysis can take another step forward on what we know about the collocational preferences of *um monte de NP* in Portuguese and what it reveals about the cognitive representation of the construction in the speakers’ mind.

The above considerations lead us to conduct a collexeme analysis to verify which lexemes are preferably recruited by *um monte de NP*. Table 1 below provides the results that point out to the top ten lexemes attracted to the micro-construction *um monte de NP* in Brazilian Portuguese.

Table 1 – Collexeme analysis of *um monte de NP* micro-construction.

Word	Freq. in Corpus	Freq. in Cx	Expect. freq.	Collostructional strength
Gente ⁴	5323	46	0.0441	121.7839
Coisas_coisa ⁵	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

Coll.strength>3 => p<0.001; coll.strength>2 => p<0.01; coll.strength>1.30103 => p<0.05.

Analyzing the lexemes presented in Table 2, we can observe that, although different items can be attracted by the construction, *gente* (people) and *coisa(s)* (thing(s)) shows the highest rates of collostructional strength in comparison to the others. These results corroborate the considerations pointed out by Fumaux (2018), since *um monte de NP* can be mostly combined with lexemes that present no semantic coherence with the idea of a *monte* (heap). In addition, collexeme analysis revealed that items such as *gente* and *coisa(s)* are very strongly attracted by the construction. Such collexeme-analysis findings allow us to suggest that *um monte de NP* is associated to the quantification of more generic referents, whatever their nature. The examples below illustrate such cases:

(9) *Esse tipo de trabalho é ótimo, pois você sente que tem **um monte***

This kind of work is great because you feel that have **a heap**

de gente no mundo inteiro preocupado com as mesmas coisas que você.

of people in the world all worried with the same things that you.

‘This kind of work is great because you feel like there are **a lot of people** around the world worried about the same things that you do.’

(10) *Descobri **um monte de coisa** naquele verão.*

discovered **a heap of thing** in that summer.

‘I discovered **a lot of things** in that summer’

(11) *Um monte de coisas acontecem aqui, mas são muito bem camufladas.*

A heap of things happen here but are very well camouflaged.

'**A lot of things** happen here but they are very well hidden.'

The collexeme analysis to *um monte de NP* was a very important tool on the understanding of how language usage impacts constructions in terms of mental representation. While the idea of quantity of the first uses of *um monte de NP* were more contextually dependent and associated to invited inferences triggered by semantic mismatching (FRANCIS; MICHAELIS 2003). Thus, the diachronic analysis took the semantic coherence between *monte* and *NP* in consideration. In the contemporary Brazilian Portuguese, the analysis of *um monte de NP* micro-construction goes in a different direction, though. Collexeme analysis for current Brazilian Portuguese demonstrates the relevance of describing the semantic coherence among exemplars fitting SN slot, since on one hand the idea of a heap (considering a quantifying binominal construction) seems not to be available anymore and, on the other hand, most of the top ten lexemes are more generic referents.

In the next section, we will present an overview about quantifying binominal constructions in Russian followed by the results of a collexeme analysis applied to *kucha NPgen*.

3.2. The case of Russian *kucha NPgen*

As well as in Portuguese, in Russian the variety of nouns used for quantification is very expressive. According to Rakhilina (2009, p. 353) these nouns constitute a class of about thirty lexical items. A few interesting examples of such nouns are the lexemes *kopna* (haycock), *gora* (mountain), *morie* (sea) and *kucha* (heap), which originally mean entities that may be found in nature. Often, they don't function as predicates themselves exhibiting relational function but as predicate arguments, functioning as a head, e. g., *bolshaia kopna* ('big haycock'), *vysokaia gora* ('high mountain'), *sinieie morie* ('blue sea') and *krasivaia kucha* ('beautiful heap').

However, when occurring in the genitive binominal construction (*NP NPgen*), such nouns may assume a predicate or relational function, being reinterpreted as quantifiers, by means of semantic coercion. In those cases, the noun itself doesn't designate an entity, as it originally did, but express a big quantity of a thing – and in a few cases, a small one (see *kaplia*, 'drop', RAKHILINA, 2009, p. 350). So, the above-mentioned lexemes *kopna*, *gora*, *morie* and *kucha*, as well as many other nouns in Russian may instantiate the *NP NPgen* construction – as in *kopna NPgen*; *gora NPgen*; *morie NPgen* and *kucha NPgen*,

respectfully – to express a big quantity of something.

While nouns as *kopna*, *gora*, *morie* and *kucha* may be used in a quantifying binominal genitive construction, being related to the sub-schema of large quantity, the quantity itself expressed by those nouns is permeated by qualitative nuances indicating subtle semantic differences among them, in terms of which items they can be combined with. Such subtle differences may be an evidence that such combinations as the above-mentioned *kopna NPgen*, *gora NPgen*, *morie NPgen* and *kucha NPgen* may be considered more specific constructions in the constructional network of quantifying binominal constructions in Russian and from now on they will be called here micro-constructions, according to Traugott and Trousdale (2013) terminology mentioned before.

In a work in progress, we have empirically observed through corpus analyses that although *gora* may be combined with a number of different referents in the quantifying genitive construction, it tends to quantify more frequently solid entities that may be grouped together to form a big mass of something, whereas *morie* is more often used to quantify uncountable and abstract entities like pleasure, joy and other emotions. *Kucha* as well as *gora* may be combined with a lot of nouns designating different referents, but according to Levontina (2004) there is some kind of difference between these two: while *kucha* refers to elements which are messily grouped together, *gora* refers more to the form and the extension of this grouping of elements. *Kopna* in turn displays semantic constraints on the elements that may be used in the quantifying genitive construction, allowing only a few numbers of items to co-occur with it. The following examples illustrate these assumptions:

- (12) *on khudoshav-ogo tieloslozheni-a i s kopnoi sedykh volos na golovie.*
3S thin-GEN build-GEN and with **haycock gray-GEN hair** on head-PRE
'he is of thin build and with **a shock of gray hair** on his head.'
- (13) *ona stoia-l-a u plity, vypeka-ia gor-u blin-ov.*
3S.F stand-PST-F at stove bake-GER **mountain-ACC pancake-GEN.PL**
'she stood at the stove, baking **a mountain of pancakes**.'
- (14) *Kazhdy zaniel ogromnuiu nichu v moiei zhizni, ostavil morie emotsii.*
Everyone occupied huge niche in my life left **sea emotions-GEN**
'Everyone occupied a huge niche in my life, left a sea of emotions.'
- (15) ... *kucha narod-u sidit po tiurmam ni za chto.*
heap people-GEN be sitting prisons NEG for nothing
'... **a lot of people** are in prison for nothing.'

Taking *kucha* in a closer detail, Rakhilina (2009) argues that this item exhibits the greatest degree of grammaticalization in comparison with other lexical items used for quantification, since it almost doesn't present constraints of combination⁶. From a usage-based constructionist perspective, it should be argued that the micro-construction *kucha NPgen* displays a high degree of entrenchment, given its productivity in terms of extensibility, so that it almost does not exhibit semantic constraints on which candidates are eligible to fill the open slot in the micro-construction. It may combine with nouns designating quite different types of referents as illustrated by the examples below:

- (16) *V rezultat-e kucha vrem-eni trati-tsa*
in result-PRE **heap time-GEN** waste-3S
na obsuzhdenie tekush-ikh vopros-ov
on discussion current-GEN questions-GEN
'As a result **a lot of time** is wasted on the discussion of current questions.'

- (17) *I dobavil: u parlamentar-iev kucha rabot-y*
and added in parliamentarians-GEN **heap work-GEN**
v et-om god-y
in this-PRE year-PRE
'And added: the parliamentarians have **a lot of work** this year.'

- (18) *Moi rabotodatel potratil kuchu denieg na poiezdku.*
1POS employer spent **heap money** On trip
'My employer spent **a lot of money** on the trip.'

- (19) *No kak vy ponimaetie mojno natvorit' kuchu*
but as 2PL understand-2PL possible do **heap**
bied, *okazavchis v chetnadsat liet v tuchzhoi*
trouble being in sixteen year in foreing-PRE
stranie biez vzroslykh
country-PRE without adults-GEN
'But as you understand it is possible to do **a lot of trouble** being at the age of sixteen without adults in a foreign country.'

However, although *kucha* may co-occur with a great number of referents in the quantifying genitive construction, there seems to be some constraints on its co-occurrence with both liquids (*kucha vody*, lit. 'heap of water' / 'a lot of water', and *kucha krovi*, lit. 'heap of blood' / 'a lot of blood') and referents which may be metaphorized as liquids (*ku-*

⁶ *Kopna*, in turn, may be use only with elements displaying some kind of visual similarity, as hair, for example, what made Rakhilina claim that this is the item with the minimum degree of grammaticalization in the quantification cline.

cha udovolstvia, lit. ‘heap of pleasure’/ ‘a lot of pleasure’), what may be seen in Li Su Hyoun (2005, apud RAKHILINA, 2009) and Rakhilina (2009). According to Li Su Hoyou hypothesis, there is some kind of semantic conflict between the homogeneity of liquids and the meaning of a messy grouping expressed by *kucha*, as pointed out by Levontina (2004). On the other hand, despite the claims made by Rakhilina and Li Su Hoyou about the constraints on the combination of *kucha* with liquids, it is possible to find real instances of this use in Russian in a search on Google, as shown in (20), (21) and (22) below:

- (20) *Protekla trubu v vannoi, kucha vody na*
 leaked pipe in bathroom heap water-GEN on
- polu. Vytierla poluchilos 4 vedra vody. Velika*
 Floor-PRE wiped turn out 4 bucket-GEN water-GEN Big
- li veroiatnost tchto zatopilo sosiediei?*
 would likelihood that flooded neighbors
- ‘A pipe leaked in the bathroom, there was **a lot of water** on the ground. I wiped and it turned out 4 buckets of water. Did I flood the neighbors?’⁷

- (21) *Pervyi opyt na Golos: razbor poletov i*
 first experience on golos sort flights and
- kucha udovolstvia***
heap pleasure-GEN

‘First experience at Golos: the sorting out of flights and **a lot of pleasure**.’⁸

- (22) *Perezjau v nov-uiu komnat-u i rechil povessit*
 Move to new-ACC room-ACC and decided hang up
- postery film-ov! Lia maman nie otsienila iz-za*
 posters Movies-GEN la maman NEG appreciate Because
- t-ogo, tchto tam vsio krasnoie agressivnoie i*
 That-GEN that there all red aggressive and
- kucha krov-i.*** *Chto dielat’?*
heap blood-GEN What do

‘I’m moving to a new room and I decided to hang up movie posters! La maman did not appreciate because everything is red, aggressive and there is **a lot of blood**. What do I do?’⁹

The examples (20), (21) and (22) were extracted from forums and a blog, genres reflecting a kind of relaxed everyday language, in which it was possible to find instances of such kinds of combination, differently from journalistic writing, where no such an example was found. Considering this as well as the insights provided by Rakhilina (2009),

⁷ <https://otvet.mail.ru/question/168564015>.

⁸ <https://golos.io/ru--golos/@atvaleyevsky/pervyi-opyt-na-golos-razbor-poletov-i-kucha-udovolstviya>

⁹ <https://thequestion.ru/questions/240901/perezzhayu-v-novuyu-komnatu-i-reshil-povesit-postery-filmov-lya-maman-ne-ocenila-iz-za-togo-chto-tam-vsyo-krasnoe-agressivnoe-i-kucha-krovi-chto-delat>

we would venture to say that this lexical combination is still associated to an informal discourse reinforcing the assumption that grammar is dynamic and spreads over language use.

Considering the overview presented here for Russian genitive quantifying binominal construction *NP NPgen*, the specificities related to the *kucha NPgen* micro-construction, and also the patterns related to Brazilian Portuguese, some research questions may be raised. The first one concerns the productivity of this micro-construction. Since it is considered the more entrenched genitive quantifying binominal construction in Russian, what is in fact the extensibility of this micro-construction in journalistic writing? Is it extensible to a large group of semantic domains regarding the *NPgen*?

Considering this question, another one may be raised: what are the collocational preferences of *kucha NPgen* in journalistic writing? Is it possible to map the semantic coverage of *kucha NPgen* micro-construction in this discourse domain?

In order to answer such questions, we have proceeded to a collexeme analysis¹⁰ of *kucha NPgen* in the same vein as the collexeme analysis processed to Brazilian Portuguese above. Table 2 below provides the top ten lexemes attracted to the *kucha NPgen* micro-construction:

Table 2 – Collexeme analysis of *kucha NPgen* construction

Word	Freq. in Corpus	Freq. in Cx	Expect. freq.	Collostructional strength
<i>diengi</i> (money)	1149	38	0.0134	116.9795
<i>mussor</i> (garbage)	44	15	0.0005	62.7590
<i>probiema</i> (problem)	5394	16	0.0745	31.5364
<i>vriemia</i> (time)	17551	15	0.2046	22.6550
<i>dokument</i> (document)	1426	9	0.0166	21.6316
<i>narod</i> (folk)	1237	8	0.0144	19.3810
<i>bumaga</i> (paper)	416	6	0.0048	16.7779
<i>rebionok</i> (child)	2952	7	0.0344	13.9854
<i>spravka</i> (certificate)	299	4	0.0035	11.2281
<i>pretenzia</i> (claim)	343	4	0.0040	10.9886

Coll.strength>3 => p<0.001; coll.strength>2 => p<0.01; coll.strength>1.30103 => p<0.05.

In terms of semantic coverage, it is possible to observe through the top ten lexemes in Table 2 that in journalistic writing the lexemes associated to the micro-construction *kucha NPgen* range from concrete referents that may be physically grouped together (documents, money, garbage and so on) to abstract things that not necessarily may be grouped

¹⁰ As well as in the collexeme analysis of Brazilian Portuguese we considered the scrip for collexeme analysis as proposed by Gries (2014), available at <http://www.linguistics.ucsb.edu/faculty/stgries/teaching/groningen/readme.txt>

in space and time (*problema* ‘problem’, *kompliment* ‘compliment’), from mass nouns as *diengi* (‘money’) or *mussor* (‘garbage’) to count nouns as *dokument* (‘document’) or *bu-maga* (‘paper’).

Although *kucha* can be combined with a great variety of lexical items in Russian as well as *monte* in Brazilian Portuguese, the collexeme analysis made for Russian reveals some kinds of collocational preferences for the *kucha NPgen* construction, differing from the *um monte de NP* construction. In Russian journalistic writing it was possible to see that *diengi* ‘money’ displays the greatest collostructional strength, followed by *mussor* ‘garbage’, *problema* ‘problem’, *vriemnia* ‘time’, *document* ‘document’ and so on. Although the nouns associated to the micro-construction designate some kind of generic entity, the degree of specificity is much greater than in Portuguese. A brief look at the top five lexemes instantiating the *kucha NPgen* micro-construction reveals at least two different semantic domains: concrete things that can be stacked, as *diengi*, *mussor* and *document*, despite their degrees of specificity, and abstracts things that may be individuated as *problema* ‘problem’ or divided in distinct portions as *vriemnia* ‘time’.

Differently from Brazilian Portuguese *um monte de SN*, *kucha NPgen* seems to have more semantic coherence among its members. One may ask what do *problema* and *vriemnia* have in common with *diengi*, *document* and other concrete referents in order to be good candidates to instantiate *kucha NPgen* micro-construction. Let’s look a few examples below:

- (23) *Rastavaisia po-dobromu dazhe s tiemi liudmi, kotoryie*
Part kindly even with those people who
sozdava-l-i tiebie kuchu problem na starom miestie.
create-PST-PL 2S.DAT heap-ACC problems-GEN in old place
‘Part kindly even with those people who created **a lot of problems** to you in the old place.’

- (24) *Nie khochu perechisliat kuchu problem sviazannykh*
NEG want list heap-ACC problem-GEN related
c nizkoi zarplaty uchiteliia
to low salaries teacher
‘I do not want to list **a lot of problems** associated with poor material base, with low teacher salaries.’

- (25) *Priezzhaiu domoi zloi i razdrzhionnyi: ia ubil*
come.1S home angry and upset 1SG killed
kuch-u vriemni.
heap-ACC time.GEN
‘I come home angry and upset: I’ve killed **a lot of time**.’

- (26) *malo* *togo* *chto* *on* *otnial* *u* *nas*
few that-GEN that 3S take in 1PL
- kuchu*** ***vriemieni***
heap-ACC **time.GEN**
- ‘not only did he take a lot of time from us’

The instances presented above show that *problema* ‘problem’ and *vriemia* ‘time’ may be metaphorized and conceptualized as entities that may be individuated, displaying some kind of materiality. According to the cognitive operation of reification as proposed by Talmy (2000, p. 43), *problema* and *vriemia* may be reified as objects that can be grouped together in large quantities, which may be transferred or taken away. In (23) the *old place* (*staroie miesto*) represents a kind of container where problems may be located in a large amount and in (24) problems are imagined as things cluttered together that can be presented and listed. As for *vriemia*, in (25) time is thought as a concrete thing that can be killed in some way and in (26) it is treated as something that in a way or another may be taken away from people.

In any case, analysis provided by Levontina (2004) of *kucha* grouping together things in a messy way applies to entities as time and problem. So, according to Li Su Hoyoun (2005) hypothesis discussed by Rakhilina (2009), *problema* and *vriemia* as well as *diengi*, *document*, *mussor* and so on designate entities that may be reunited somehow in a messy way, as empirically verified in corpus analysis. Thus, *kucha NPgen* semantic features can be distinguished from the other micro-constructions in a large amount sub-schema of quantifying binominal constructions in Russian.

4. Contrasting *um monte de NP* and *kucha NPgen*: a distinctive collexeme analysis

This section provides a comparison between Brazilian Portuguese *um monte de NP* construction and Russian *kucha NPgen* construction. Section 4.1. presents a distinctive collexeme analysis for these constructions indicating the preferences for lexemes that emerges when these constructions are compared. Section 4.2. offer a preliminary interpretation for the results obtained through the collostructional analysis performed in this research.

4.1 A distinctive collexeme analysis for Portuguese and Russian Quantifying binominal constructions

The collexeme analysis presented in the previous section were made separately for Brazilian Portuguese and Russian. Now, following Gilquin (2015), we walk toward a direct analysis between Brazilian Portuguese *um monte de NP* micro-construction and Russian *kucha NPgen* micro-construction by means of a distinctive collexeme analysis. As mentioned in section 2, the aforementioned type of analysis allows one to study one slot in two or more similar constructions. The analysis accomplished here considers the occurrences of lexemes within the *um monte de NP* micro-construction in Brazilian Portuguese and the translation, from Russian to Portuguese, of the lexemes occurring within the *kucha NP* micro-construction.

An interesting result is that out of the 364 different lexemes, 26 collexemes are shared by both constructions corresponding to a proportion of 7.14%, what may represent a small overlap in terms of the semantic fields covered by the respective constructions in Brazilian Portuguese and Russian. In addition, this overlap covers concrete inanimate referents (like money, paper, letter), concrete animate referents (like friend, tourist), as well as inanimate, abstract referents (like question, action, mistake).

However, these similarities between *um monte de NP* and *kucha NPgen* constructions seem not to be as significant as their specificities. Table 3 presents the statistically relevant collexemes occurring in *um monte de SN*, shown in the upper part of the table, and *kucha NPgen* constructions, shown in the lower part of it. It is possible to observe that despite Russian *kucha NPgen* micro-construction displayed a greater token frequency, the number of statistically relevant lexemes co-occurring in each construction is relatively the same (6 for Brazilian Portuguese and 7 for Russian). Besides, the force of attraction of lexemes to Brazilian Portuguese *um monte de NP* micro-construction is relatively greater than in Russian (the difference among thing, people in Brazilian Portuguese and the preferred collexemes for Russian *kucha NPgen* micro-construction is considerably greater), revealing that some lexemes in Portuguese have higher force of attraction in comparison with any preferable collexeme in Russian.

Table 3 – Distinctive collexeme analysis: *Um monte de NP* and *kucha NPgen*

Lexeme	Obs. freq. Portuguese	Obs. freq. Russian	Coll. strength	Pref. occur.
Coisa (thing)	38	0	14.5205	Portuguese
Gente (people)	46	6	11.7034	Portuguese
Cara (guy)	4	0	1.4677	Portuguese
Filme (movie)	4	0	1.4677	Portuguese
Homem (man)	4	0	1.4677	Portuguese
Mulher (woman)	4	0	1.4677	Portuguese
Dinheiro (money)	7	38	4.3008	Russian
Problema (problem)	0	16	3.9751	Russian
Tempo (time)	0	15	3.7229	Russian
Lixo (garbage)	1	15	2.8417	Russian
Documento (document)	0	9	2.2204	Russian
Povo (folk)	0	8	1.9718	Russian
Criança (child)	0	7	1.7236	Russian

Coll.strength>3 => p<0.001; coll.strength>2 => p<0.01; coll.strength>1.30103 => p<0.05.

The results presented in Table 3 point to differences between Portuguese and Russian binominal quantifying constructions in terms of the preferable collexemes occurring in *um monte de NP* and *kucha NPgen* respectively. It is interesting to note that Portuguese and Russian go in different directions in terms of preferable lexemes to occur in the micro-construction. While Portuguese *um monte de NP* covers more generic referents as thing or people, and mostly animate referents as woman, guy or man, Russian genitive quantifying binominal construction *kucha NPgen* relates to mostly inanimate and less generic referents, as money, problem, and time.

4.2. Interpreting Brazilian Portuguese *um monte de NP* and Russian *kucha NPgen* constructions

Collostructional analysis is certainly one of the best choices if you want to capture the meaning of a construction in a corpus. In this work we could observe Russian *kutcha NPgen* and Portuguese *um monte de NP* preferable collocates, i.e. piled stuff for Russian and generic referents for Portuguese. Although those results offer a good insight about constructional meaning of each construction, it seems to be necessary to take a further step in the interpretation of them, considering their implications to the study of grammar. So, a question which arises from the results of distinctive analysis, especially when two different languages are being compared, concerns to what differences on preferable col-

locates for Russian and Portuguese tell us about differences across languages in general.

As already demonstrated in this paper, both Russian and Portuguese quantifying binominal constructions can be associated to the cognitive process called by Talmy (2006) as *multiplexing*, which is a process by which different constructions affect the conceptualization of a referent, in terms of having multiple copies of it. Even though this is true, there is no guarantee that the same semantic cluster of lexemes will be preferably attracted by the quantifying binominal constructions in both languages. The reason for that relies on the dynamic development of such a language and tell us much about the close relationship between past and present on language.

Brazilian Portuguese *um monte de NP* construction was probably coined in the 18th century as a result of a long span of time of change, since 16th (at least). Fumaux (2018) observed that the spread of usages of *um monte de* being combined with lexemes referring to natural elements (*NP*) as *sand*, *straw*, etc., resulted from the iconic association between the shape of a heap (*monte*) and the shape of a *sand heap*, for example. So, contexts yielding quantity inferences were frequently associated to the form *um monte de NP* throughout the history of Portuguese and, then, a new form-meaning pairing, i.e. the quantifying binominal construction, arose. Initially, preferable collocates for this construction were piled items and then gathered items, not necessarily piled but contiguous in space and time. In 19th and 20th centuries, Fumaux (2018) observed an increasing of the spread of uses of the construction and consequently a decreasing of the semantic specification of the lexemes it recruits.

Although many different clusters can be recruited by the construction *um monte de NP* in contemporary Brazilian Portuguese, its preferable collocates correspond to more generic referents such as *gente* (people) and *coisas* (things). Assuming that it can be combined with any lexeme probably unrestrictedly, we still have to explain why generic referents instead of any other semantic cluster of lexemes are preferred by the construction.

The answer seems to be twofold: first, the increasing of instances of *um monte de gente* e *um monte de coisa(s)*, which affects the mental representation of the construction and, because of that, stimulates the recruitment of new types of generic referents (eg. *dinheiro*/ money; *caras*/ guys), increasing the type frequency but lower variability. Second, the competition between constructions on language system, by which new quantifying binominal constructions were coined in Brazilian Portuguese in the 20th century, such as *uma enxurrada de NP* ('a flash flood of NP', a lot of NP), *uma chuva de NP* ('a rain of NP', a lot of NP), *uma montanha de NP* ('a mountain of NP', a lot of NP). Those most recently

coined competitor constructions preferably recruit, respectively, lexemes that are being transferred from one place to another, lexemes that are interpreted as 'bad' and lexeme referring to piled stuff, which indicates that they carry a specific meaning (by relative semantic coherence between *enxurrada*, *chuva e montanha* and the respective *NP*) that *um monte de NP* cannot express anymore. Therefore, *um monte de NP* would be preferably used on unmarked contexts while the others would be preferably used on marked contexts, which present a higher level of expressivity.

Concerning to Russian, since there are no studies (from the best of our knowledge) which describe the history of *kutchá NPgen*, we assume that as a construction recently coined on language, on the basis of the fact that its performance is closer to the newest constructions of Portuguese quantifying binominal constructions mentioned above. *Kutchá NPgen* preserves a relative semantic coherence between *kutchá*, the name taken as the quantifier, and the lexemes recruited to the *NPgen* slot, taken as quantified. Thus, we believe the difference between Russian *kutchá NPgen* and Portuguese *um monte de NP* preferable collocates can bring out information about time and spreading and, consequently, extensibility of constructions.

This phenomenon illustrates the dynamics of language system, revealing the different forces in action. Oldest constructions undergoes changes in their behavior as well as changes in the language system as a whole, such as the influence of how oldest and newest competitor constructions (by similarities in form or function) behave. Since time affects grammar system and thus the performance of constructions, we should consider the following issues: the newest and the oldest binominal constructions are at different stages in terms of entrenchment (the oldest constructions tend to be more entrenched than the newest ones); productivity (the oldest constructions tend to be more productive in terms of extensibility than the newest ones); semantic coherence (the oldest constructions tend to present less semantic coherence than the newest ones); spreading (the oldest constructions tend to be allowed in more contexts than the newest ones). Therefore, we should consider that, if constructions have its own history on language, the comparison between them can not put aside crucial aspects that involves the dynamic system of grammar.

In addition, we believe that *um monte de NP* emerged as an alternative to the existent quantifiers and could give a kind of emphasis to the quantification, but, throughout the time, it (partially) lost that capacity since it increased in frequency and decreased in terms of its semantic coherence. Thus, the newest binominal constructions are used

today in competition with *um monte de NP* in contexts requiring a higher level of expressivity (pragmatically and semantically), following the Principle of Maximized Expressive Power, according to which the inventory of constructions is maximized for communicative purposes (GOLDBERG, 1995, p. 67).

5. Conclusion

Quantity is a basic concept for human cognition being expressed in language in various ways, and, worldwide, speakers of different languages recruit non-quantity nouns to quantify referents. In this paper, we investigated two quantifying binominal constructions – *um monte de NP*, in Brazilian Portuguese, and *kucha NPgen*, in Russian – under a usage-based constructionist approach, from a contrastive perspective.

The methodology applied to this research involves a family of methods named as collocation analysis, which in a usage-based perspective of language has been proven to be a useful tool. In this work, we used two of those methods in order to investigate the collocational preferences of each construction in Portuguese as well as Russian: a simple collexeme analysis and a distinctive collexeme analysis.

The results obtained through those analyses reveal that Brazilian Portuguese *um monte de NP* micro-construction attracts preferably lexemes denoting more generic referents. As far as Russian *kucha NPgen* is concerned, the preferably co-occurring lexemes tends to denote not so generic entities as in Portuguese. The semantic coverage of this construction in Russian seems to be more diverse, ranging from concrete massive nouns as *mussor* (garbage) to abstract nouns such as *problema* (problem). In addition, there seems to be a semantic difference between the micro-constructions analyzed in this paper. While Brazilian Portuguese *um monte de NP* seems to have lost the semantics related to heap, being associated to generic/general referents, Russian *kucha NPgen* seems to have preserved some features associated to the idea of a heap, designating entities that may be reunited somehow in a messy way.

Another question that can be raised concerns to how much a specific corpus is representative of language as a whole. Since we used a sample of journalistic texts, the results obtained in this work are valid to the extent it concerns to this specific discourse domain. Considering the benefits and the limitations of this work, our results show that, although Portuguese and Russian make use of metaphor and reification (see section 3) as cognitive operations directly associated to quantifying binominal constructions, collocational preferences demonstrate that each construction in each language exhibits semantic spec-

ificities which highlights how use impacts construction representation, corroborating to the assumption that grammar is language-specific.

References

- ALONSO, K. S. B. *Construções binominais quantitativas e construção de modificação de grau: uma abordagem baseada no uso*. 2010. Tese (Doutorado em Linguística) – PPG em Linguística, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.
- BARDDAL, J. *Productivity: evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins, 2008.
- BREMS, L. The establishment of quantifier constructions for size nouns: a diachronic study of heap(s) and lot(s). *Journal of Historical Pragmatics*, v. 13, p. 202– 231, 2012.
- 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*. 2010. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal de Juiz de Fora, 2010.
- BYBEE, J. *Language, usage and cognition*. New York: Cambridge University Press, 2010.
- DIESSEL, H. Usage-based construction grammar. In: DABROWSKA, Ewa; DIVJAK, Dagmar (eds.). *Handbook of Cognitive Linguistics*. Berlin: Mouton de Gruyter, 2015. p. 295-321.
- DIESSEL, H. *The grammar network: how language structure is shaped by language use*. Cambridge: University Press, 2019.
- FRANCIS, E. J., MICHAELIS, L. A. (Org.) *Mismatch: form-function incongruity and the architecture of grammar*. Stanford, CA: CSLI Publications, 2003.
- FUMAUX, N. C. A. *Construcionalização de ‘um monte de SN’: uma abordagem centrada no uso*. 2018. Dissertação (Mestrado em Linguística) – PPG em Linguística, Universidade Federal do Rio de Janeiro, 2018.
- GILQUIN, G. Contrastive colostruational analysis: causative constructions in English and French. *Zeitschrift für Anglistik und Amerikanistik*, v. 63, n. 3, p. 253-272, 2015.
- GOLDBERG, A. E. *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.
- GOLDBER, A. E. *Explain me this: creativity, competition and the partial production of constructions*. Princeton: Princeton University Press, 2019.
- GRIES, S. T. *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, 2004.

LEVONTINA, I. B. Slovarnyie stati kucha, gora, gruda, vorokh i kipa. In: *Novyi obiasnitelnyi slovar sinonimov russkogo iazyka*. Moskva i Viena: Iazyki slavianskoi kultury, Vienskii slavisticheski almanakh, 2004.

LEVSHINA, N. *How to do Linguistics with R*. Amsterdam: John Benjamins. 2015.

RAKHILINA, E. *Lingvistika konstruksii*. Moskva: Abukovnik, 2009.

SANTOS, C. *Gramática e cognição: um estudo de construções binominais*. 2014. Dissertação (Mestrado em Linguística) – PPG em Linguística, Universidade Federal do Rio de Janeiro, 2014.

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

TALMY, L. *Toward a Cognitive Semantics*. Volume I: Concept structuring systems. Cambridge, Massachusetts, London: The MIT Press, 2000.

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.

TAVARES, T. As construções binominais de quantificação indefinida: uma análise construcionista. In: *Litteris*, n. 4, set. 2014.

TRAUGOTT, E. C. Grammaticalization, constructions and the incremental development of language: suggestions from the development of degree modifiers in English. In: ECKARDT, R.; JÄGER, G.; VEENSTRA, T. (eds.). *Variation, selection, development: probing the evolutionary model of language change*. Berlin: Mouton de Gruyter, 2008. p. 219–250.

TRAUGOTT, E. C., TROUSDALE, G. G. *Constructionalization and constructional change*. Oxford University Press: Oxford, 2013.

VERVECKKEN, K. D. *Binominal quantifiers in Spanish: conceptually-driven analogy in diachrony and synchrony*. De Gruyter, 2015.



Data de submissão: 15/12/2019

Data de aceite: 05/05/2020