ROCZNIKI HUMANISTYCZNE
Tom LXXI, zeszyt 6 – 2023
DOI: https://doi.org/10.18290/rh23716.20

JAROSŁAW WILIŃSKI

A SEA OF-NOUN VS. *A MOUNTAIN OF*-NOUN: A QUANTITATIVE CORPUS-BASED STUDY OF TWO METAPHORICAL CONSTRUCTIONS

INTRODUCTION

To date, no research has attempted to compare and contrast the use of nouns in the *sea of*-NOUN-construction and the *mountain of*-NOUN-construction, to say nothing of their quantification and evaluation in terms of statistical significance. To the author's knowledge, previous work has focused solely on identifying several noun collocates of the pattern *sea of*-NP. Analyzing a few usage examples extracted from the BNC corpus, Hanks noticed that *a sea of* collocates with *mud*, *blood*, *people*, *faces*, *heads*, *hands*, and *hats*. In addition, he mentioned that there are 301 metaphorical uses of this construction in BNC.

The constructions under study have also received scant treatment in lexicographic research. Although lexicographers have provided similar definitions for their meanings and have given some usage examples, they have been unsuccessful in explaining the subtle differences in their semantics and collocability. For example, the lexicographers who created the online version of the *Macmillian English Dictionary* (MacmillanDictionary.com) in 2009 solely provided the definitions of these partitives (*a sea of* 'a large amount of something' and *a mountain of* 'a large pile or amount of something'), gave one or two illustrative examples of their usage, and grouped them semantically (along with some other partitives) on the dictionary's sister site *Macmillan Thesaurus* under the same topic 'large

JAROSŁAW WILIŃSKI, PhD, Siedlce University of Natural Sciences and Humanities, Faculty of Humanities, Institute of Linguistics and Literary Studies; e-mail: jaroslaw.wilinski@uph.edu.pl; ORCID: https://orcid.org/0000-0002-3136-6529.

Dr JAROSŁAW WILIŃSKI, Uniwersytet Przyrodniczo-Humanistyczny w Siedlcach, Wydział Nauk Humanistycznych, Instytut Językoznawstwa i Literaturoznawstwa; e-mail: jaroslaw.wilinski@uph.edu.pl; ORCID: https://orcid.org/0000-0002-3136-6529.

quantities and amounts', thereby signaling that both constructions are synonymous and related thematically. Thus far, however, the compilers of this dictionary and other reference works have failed to identify the nuances and subtleties of their meanings and indicate their potential noun collocates.

Since no study has investigated these two constructions in their respective collocational preferences in much detail and little attention has been paid to both constructions in dictionaries, there is still a need for determining the distribution of nouns and their frequency of occurrence in these patterns, because of the possible existence of distributional variations. This paper, therefore, aims to ascertain consequential, if slight, differences between the two patterns and to support the hypothesis that nearly synonymous though these two constructions might appear to be at first glance, they do display subtle differences in use and show definite preferences for specific categories of nouns.

The rest of this paper is organized into three sections. Section 2 discusses the theoretical and methodological frameworks underpinning the quantitative investigation: the usage-based model of construction grammar, the method of distinctive-collexeme analysis, the corpus, the data, the tools, and the statistical procedure followed in this study. Section 3 combines the findings of the quantitative analysis with a semantic description of nouns and elucidates subtle distributional differences between the two semantically near-equivalent constructions. Section 4 evaluates the results and puts forward some proposals for future research.

1. THEORETICAL AND METHODOLOGICAL FRAMEWORKS

The study rests on the usage-based approach to Construction Grammar (Hoffmann; Goldberg, *Constructions at Work* and "Constructionist Approaches"). This model of grammatical knowledge assumes that grammar is a structured inventory of constructions (pairings of form and meaning) at various levels of specificity and schematicity. Constructions encompass all linguistic levels, from morphemes through words to syntactic patterns and associated rules of their semantic, pragmatic, and discourse-functional interpretation (Hoffman). Thus, for example, *a sea of*-NOUN and *a mountain of*-NOUN are constructions since they have a form (a nominal partitive that is followed by the head noun) and a meaning (the sense 'a large amount of something') that are conventionally associated with each other. All linguistic units can be stored and represented as constructions as far as they occur with sufficient frequency, which in turn determines their entrenchment in a speaker/hearer's grammatical system (Croft and Cruse 292-293).

The method referred to as *Distinctive-Collexeme Analysis* (Gries and Stefanowitsch; Stefanowitsch; Hilpert) was used to determine those nouns that are distinctively associated with the *sea of*-NOUN-construction as compared to the *mountain of*-NOUN-construction. The method involved taking four steps. These steps can be clearly illustrated by reference to the noun *evidence* in the noun slot of both constructions in question. Table 1 shows the distribution of this noun in both patterns and other frequencies required for a distinctive-collexeme analysis.

Table 1. Contingency table cross-tabulating frequency scores of the noun *evidence* and the constructions under study

Constructions	**Nouns (*evidence*)**	**All other nouns**	**Total**
***A/the sea of*-NOUN**	Frequency of noun (*evidence*) in '*a/the sea of*-NOUN-construction' a = 0 (130.79)	Frequency of all other nouns in '*a/the sea of*-NOUN-construction' b = 4122	Total frequency of '*a/the sea of*-NOUN-construction' x = 4122
***A/the mountain of*-NOUN**	Frequency of noun (*evidence*) in '*a/the mountain of*-NOUN-construction' c = 198 (67.21)	Frequency of all other nouns in '*a/the mountain of*-NOUN-construction' d = 1920	Total frequency of '*a/the mountain of*-NOUN-construction' y = 2118
Total	Total frequency of noun (*evidence*) e = 198	Total frequency of all other nouns f = 6042	Total frequency of both constructions z = 6240

The first step entailed searching for these patterns in the Corpus of Contemporary American English (COCA), extracting their occurrences in concordance lines, identifying the metaphorical senses, and calculating the observed frequencies. The choice of this data source was dictated by the enormous size of the corpus and its representativeness. The corpus is well-balanced, is representative of the language variety, and covers over one billion words of data, including 20 million words each year from 1990-2020, derived from eight genres: spoken, fiction, popular magazines, newspapers, academic texts, TV and movies subtitles, blogs, and other web pages. The search engine in this corpus allowed for the retrieval of noun collocates of the two patterns. The retrieval was restricted to four places to the right of the combinations under scrutiny (i.e., *a/the sea of* and *a/the mountain of*). The corpus search initially uncovered 2058 different noun collocates of *a/the sea of* and 957 collocates of *a/the mountain of*. Next, all potential collocates of the patterns were manually examined to identify genuine combinations. All false combinations (the occurrences which did not correspond

with the metaphorical sense 'a large quantity of something': e.g., *the territorial sea of a coastal state* or *the mountain of the Lord's house*) were discarded from further analysis. The metaphorical nature of the constructions was established based on a lexical semantic test that considers distinct but comparable senses of the words *sea* and *mountain*: *sea* and *mountain* are used metaphorically in the patterns under study when their most basic literal senses ('a large area of salt water' and 'a very large hill') stand in contrast to their current contextual meanings (e.g. *a mountain of evidence* or *a sea of people*), and there is a cross-domain correspondence between the two senses (a large quantity of something is understood as a sea or a mountain) (cf. Steen et al.). Finally, the observed frequencies of the patterns with the metaphorical senses and the remaining instances of nouns were calculated manually by reading concordance lines. The figures (a, x, c, y) in Table 1 were obtained from the corpus directly, while the remaining ones result from addition and subtraction.

In the second step, these observed frequencies were used to calculate the expected frequencies of the noun (*evidence*) in both constructions. This calculation was performed in Microsoft Excel in the following order: for the lemma *evidence* in each pattern, its column total was multiplied by its row total, and this final score was divided by the overall table total. For illustrative purposes, the expected frequencies for the lemma *evidence* in each construction are provided in parentheses (see Table 1). If the observed frequency of the noun (*evidence*) in the *sea of*-NOUN-construction is significantly higher or lower than expected, the mutual association between the noun *evidence* and this construction is one of attraction or repulsion respectively (the noun is then deemed to be a significantly attracted or repelled *collexeme* of the *sea of*-NOUN-construction). Likewise, if the observed frequency of the noun (*evidence*) in the *mountain of*-NOUN-construction is significantly higher or lower than expected, then the noun is a significantly attracted or repelled *collexeme* of this construction.

In the third step, the strength of association (the so-called *collostruction strength*) between the noun (*evidence*) and the constructions under scrutiny was estimated. To this end, the figures (a, b, c, d) from Table 1 were entered in a 2-by-2 table and submitted to the Fisher exact test, which is considered by mathematicians as "the most appropriate significance test for contingency tables" (Evert 1235). In particular, it is strongly recommended if at least one expected value in the table is smaller than 5 (Levshina 29) or if data is very unevenly distributed and/or infrequent (cf. Gries and Stefanowitsch 10; Gries, "Frequencies"; Gries, "More (Old and New) Misunderstandings"). The *p*-value obtained from this test was used to gauge the collostruction strength of each noun, i.e. the

degree of attraction to the *sea of*-NOUN-construction or the *mountain of*-NOUN-construction: the smaller the *p*-value, the higher the probability that the observed distribution is not due to chance and the higher the strength of the attraction between a noun and one of the synonymous constructions in question. This computation of statistical significance was performed using an online Fisher's exact test calculator for two-by-two contingency tables. Then, each *p*-value (e.g. 2.1340291864802483e-96 for *evidence*) was transformed into the logarithm to the base 10 (e.g. 95.67), which provided a more readable score than the *p*-values, which are frequently expressed in powers of ten. The collostruction strength above 1.301 means that the noun is strongly associated with the construction, while the collostruction strength below 1.301 means that the noun is significantly repelled by the construction. This shows that the noun *evidence* is highly significant for one of the two constructions. However, it does not suggest which one. In order to determine this, the observed frequencies of the noun need to be compared with the expected ones. As this comparison reveals, the noun *evidence* occurs more frequently than expected in the *mountain of*-NOUN-construction and less frequently than expected in the *sea of*-NOUN-construction. In other words, *evidence* is a highly significant, very strongly distinctive collexeme of the *mountain of*-NOUN-construction as opposed to the *sea of*-NOUN-construction

Finally, the results of the quantitative investigation were sorted according to the strength of attraction and then evaluated qualitatively. More clearly, the quantitative findings were integrated with a semantic classification of nouns, and subtle semantic differences between these two near-equivalent constructions were explained.

2. RESULTS AND DISCUSSION

A careful examination of concordance lines revealed 4122 occurrences of the *sea of*-NOUN-construction and 2118 occurrences of the *mountain of*-NOUN-construction. In other words, the occurrence of the former appears to be approximately twice as frequent as the latter in COCA. The observed frequencies obtained from the calculation of the tokens of nouns in both constructions indicate that *a/the sea of* collocates with 1656 types of nouns, out of which 1124 types occurred only once in the construction in question. By contrast, *a/the mountain of* combines with 769 types of nouns, out of which 505 types were used only once with this construction. This in turn means that the bulk of nouns are rather

loosely associated with both patterns, and it also suggests that the remaining ones are more strongly attracted to one of these near-synonymous constructions.

Given that many nouns occur in both of these constructions, we should expect that the two constructions are to some extent synonymous. However, the findings of this study imply significant differences between both constructions concerning the semantic constraints they impose on the nouns that can occur in them. This section will solely report the results for the 60 most strongly attracted collexemes of the constructions, since it is impossible to present and evaluate the results for all these nouns in the space here allotted.

2.1 FINDINGS FOR THE *SEA OF*-NOUN-CONSTRUCTION

Table 2 below displays the thirty most distinctive collexemes of the *sea of*-NOUN-construction, the observed frequencies used to calculate the direction of association (attracted or repelled) and the strength of association (the distinctiveness of nouns), the expected frequencies for each noun: (a) and (c), as well as the findings of the distinctive-collexeme analysis (*p*-values and collostructional strength).

Table 2. The thirty most distinctive collexemes of the *sea of*-NOUN-construction

a = Observed frequency of nouns (e.g. *faces*) in the *sea of*-NOUN-construction; **b** = Frequency of all other nouns in the *sea of*-NOUN-construction; **c** = Observed frequency of noun (e.g. *faces*) in the *mountain of*-NOUN-construction; **d** = Frequency of all other nouns in the *mountain of*-NOUN-construction; **e** = Total frequency of noun (e.g. *faces*); **f** = Total frequency of all other nouns; **x** = Total frequency of the *sea of*-NOUN-construction; **y** = Total frequency of the *mountain of*-NOUN-construction; **z** = Total frequency of both constructions; **(a)** = Expected frequency of noun (e.g. *faces*) in the *sea of*-NOUN-construction; **(c)** = Expected frequency of noun (e.g. *evidence*) in the *mountain of*-NOUN-construction; ***p*-values** and **collostructional strength** = indexes of statistical significance

rank	noun	a	c	e	f	x	y	z	b	d	(a)	(c)	*p*-values	coll. strength
1.	**faces**	138	0	138	6102	4122	2118	6240	3984	2118	91.16	46.84	9.847066828235347e-26	25.01
2.	**people**	137	3	140	6100	4122	2118	6240	3985	2115	92.48	47.52	4.0986305824023634e-21	20.39
3.	**red**	56	0	56	6184	4122	2118	6240	4066	2118	36.99	19.01	1.2732030907405394e-10	9.90
4.	**troubles**	50	0	50	6190	4122	2118	6240	4072	2118	33.03	16.97	1.0864944958347145e-9	8.96
5.	**green**	45	0	45	6195	4122	2118	6240	4077	2118	29.73	15.27	9.188015293087668e-9	8.04
6.	**grass**	44	0	44	6196	4122	2118	6240	4078	2118	29.07	14.93	1.5509346760815443e-8	7.81

rank	noun	a	c	e	f	x	y	z	b	d	(a)	(c)	*p*-values	coll. strength
7.	**blood**	41	1	42	6198	4122	2118	6240	4081	2117	27.74	14.26	7.194332331631862e-7	6.14
8.	**heads**	26	0	26	6214	4122	2118	6240	4096	2118	17.18	8.83	0.00002904639453136692	5.64
9.	**humanity**	46	3	49	6191	4122	2118	6240	4076	2115	32.37	16.63	0.000005697158768851033	5.24
10.	**stars**	27	0	27	6213	4122	2118	6240	4095	2118	17.84	9.16	0.00001713804727101634	4.77
11.	**red ink**	26	0	26	6214	4122	2118	6240	4096	2118	17.18	8.83	0.00002904639453136692	4.54
12.	**darkness**	22	0	22	6218	4122	2118	6240	4100	2118	14.53	7.47	0.00014638634302460693	3.83
13.	**change**	21	0	21	6219	4122	2118	6240	4101	2118	13.87	7.13	0.00025299156483017194	3.60
14.	**mud**	35	3	38	6202	4122	2118	6240	4087	2115	25.10	12.90	0.00039825372874367185	3.40
15.	**fans**	19	0	19	6221	4122	2118	6240	4103	2118	12.55	6.45	0.00044872144071614014	3.35
16.	**particles**	18	0	18	6222	4122	2118	6240	4104	2118	11.89	6.11	0.0007433632204970012	3.13
17.	**lights**	17	0	17	6223	4122	2118	6240	4105	2118	11.23	5.77	0.0012697149474754903	2.90
18.	**oil**	17	0	17	6223	4122	2118	6240	4105	2118	11.23	5.77	0.0012697149474754903	2.90
19.	**time**	34	4	38	6202	4122	2118	6240	4088	2114	25.10	12.90	0.001597840753759005	2.80
20.	**glass**	32	4	36	6204	4122	2118	6240	4090	2114	23.78	12.22	0.0024389173255296402	2.61
21.	**men**	24	2	26	6214	4122	2118	6240	4098	2116	17.18	8.83	0.003069884806141511	2.51
22.	**bodies**	24	2	26	6214	4122	2118	6240	4098	2116	17.18	8.83	0.003069884806141511	2.51
23.	**orange**	14	0	14	6226	4122	2118	6240	4108	2118	9.25	4.75	0.0038027995555504008	2.42
24.	**trees**	19	1	20	6220	4122	2118	6240	4103	2117	13.21	6.79	0.0038499646977987498	2.41
25.	**hats**	15	0	15	6225	4122	2118	6240	4107	2118	9.91	5.09	0.00408482590941959	2.39
26.	**flags**	13	0	13	6227	4122	2118	6240	4109	2118	8.59	4.41	0.006440314527638787	2.19
27.	**clouds**	13	0	13	6227	4122	2118	6240	4109	2118	8.59	4.41	0.006440314527638787	2.19
28.	**sand**	37	7	44	6196	4122	2118	6240	4085	2111	29.07	14.93	0.010168879396360408	1.99
29.	**tears**	12	0	12	6228	4122	2118	6240	4110	2118	7.93	4.07	0.011265339653102089	1.95
30.	**suits**	12	0	12	6228	4122	2118	6240	4110	2118	7.93	4.07	0.011265339653102089	1.95

In the case of the *sea of*-NOUN-construction, the findings imply that the five most distinctive nouns are *faces*, *people*, *red*, *troubles*, and *green.* The log transformations taken to be indicators of their distinctiveness are very high: 25.01, 20.39, 9.90, 8.96, and 8.04, respectively. A comparison of the observed and the expected frequencies of each of these nouns and each of the two constructions shows us that the nouns occur more frequently than expected in the *sea of*-NOUN-construction and less frequently than expected in the *mountain of*-NOUN-construction. In other words, they are very strongly distinctive collexemes of the former as compared to the latter. Note also that *faces* is the strongest collexeme for the *sea of*-NOUN-construction, since its collostructional strength resulting from the calculation of the Fisher exact test is exceptionally high (25.01), and the expected frequency is lower than the observed frequency in the pattern in question.

Even a cursory examination of the results in Table 2 reveals that the most significant group of collexemes is constituted by nouns associated with people or persons. *Faces*, *people*, *heads*, *humanity*, *fans*, *men*, *bodies*, *hats*, and *suits* in ranks 1, 2, 8, 9, 15, 21, 25, and 30 fall into this category of distinctive collexemes of the *sea of*-NOUN-construction. *Humanity* denotes 'all the people who live in the world', while *faces*, *heads*, *bodies*, *hats*, and *suits* are used metonymically to stand for people. As can be seen in Table 2, solely a few occurrences of *people*, *humanity*, and *bodies* were observed in the *mountain of*-NOUN-construction. Hence, the nouns are most distinctive for the *sea of*-NOUN-construction in direct comparison with the *mountain of*-NOUN-construction

Another set of the more distinctive collexemes of the *sea of*-NOUN-construction consists of nouns denoting colors. *Red*, its leading collexeme in rank 3, is accompanied by *green* and *orange* in ranks 5 and 23. All these nouns seem to be used metonymically to refer to some physical entities possessing these specific colors. For example, *green* may stand for green grass, trees, meadows, or hills. The word *green* itself precedes *grass* in rank 6, which may suggest that the combination *a sea of green* is commonly used to designate 'a large quantity of grass'. The ranking list also includes the word *trees*, in rank 24, which metonymically stands for 'a large area of land covered by trees'.

The next category comprises a range of countable nouns. *Troubles*, ranked 4, is the most significant lexeme of this group. It is frequently used with *a sea of* to pertain to 'a very large number of problems'. It is followed by nouns designating various material objects, physical entities, or substances that possess boundaries and therefore are observable as wholes in the world. These are *stars*, *particles*, *lights*, *flags*, *clouds*, and *tears.* The semantic property of *sea* that is systematically exploited in the combinations with these nouns is its vastness or immensity. All

these metaphorical combinations share the feature of being perceived as a vast expanse of something that is not salty water. Thus, *stars*, *particles*, and *clouds* are perceived as physical entities that are scattered in large numbers around a huge area. For example, *a sea of stars* can be understood as a large number of small bright lights covering a wide area of the sky at night.

The last group of the most significant lexemes is constituted by a variety of uncountable nouns such as *blood*, *red ink*, *darkness*, *change*, *mud*, *oil*, *time*, *sand*, and *glass*. *Blood*, ranked 7, holds the highest position among the most distinctive nouns belonging to this set. *Blood* is used here metonymically to refer to 'violence and death'. *Red ink*, ranked 11, metonymically stands for 'a financial deficit, loss or debt'. The phrase derives from the practice of using red ink to denote 'debt or losses on financial balance sheets'. *A sea of darkness*, ranked 12, means 'a complete lack of light, especially because it is night'. Occasionally, it can also have the metaphorical sense 'much evil', thus reflecting the metaphor EVIL IS DARK. The phrase *a sea of change* denotes 'a series of many actions or events by which things become completely different'. *A sea of mud* designates 'a large quantity of very soft wet earth'. *A sea of oil* pertains to 'a large amount of a thick dark smooth liquid used for making petrol and other fuels'. It frequently collocates with the verbs *float on* or *sit on/atop*. *A sea of time* denotes 'a plenty of time', while *a sea of sand* means 'a large amount of a loose pale brown substance at a beach or in the desert'. *A sea of glass* is used in COCA in two senses. The first sense refers to 'a large quantity of a hard clear substance used for making objects'. The second sense originates from biblical cosmology, and it is used in Revelation 4:6 in the following context: *Also before the throne*, *there was what looked like a sea of glass*, *clear as crystal* (COCA, WEB: bible.cc/revelation/5-6.htm). In this passage, *the sea of glass* pertains to a part of the surroundings of God's throne in heaven. This space reflects the qualities of the throne and functions as a barrier between the holy God and the corrupt sin-filled universe. *The sea of glass* points to a throne room that is holy, pure, glorious, and eternal, thus symbolizing the holiness required of those who draw near the throne.

2.2 FINDINGS FOR THE *MOUNTAIN OF*-NOUN-CONSTRUCTION

Concerning the *mountain of*-NOUN-construction, the results confirm the hypothesis predicting that there are nouns strongly attracted to this construction. The results of the distinctive collexeme analysis for the thirty most strongly attracted collexemes of the *mountain of*-NOUN-construction in direct comparison with the *sea of*-NOUN-construction are rendered in Table 3. As can be observed,

evidence is the most distinctive collexeme of the *mountain of*-NOUN-construction, as the log transformation of the *p*-value (2.1340291864802483e-96) resulting from the calculation of the Fisher exact test for this noun is high: 95.67. In addition, a comparison of the observed values with the expected ones shows that *evidence* occurs more frequently than expected by chance in the *mountain of*-NOUN-construction as compared to the pattern with *a/the sea of*. This and three other nouns, such as *data*, *knowledge*, and *research* in ranks 4, 29, and 10, are semantically related to facts, information, or the study of data in order to discover new facts. In the *mountain of-NOUN*-construction, *evidence* is used in the context of a prosecutor saying that a suspect left many traces behind, or in the context of a scientist stating that there are many facts supporting his/her hypotheses.

The next group in the ranking is constituted by a range of nouns denoting money or an amount of money that somebody owes. *Debt* in rank 2, the most significant collexeme of this group, is accompanied by *money*, *cash*, and *bills* in ranks 7, 11, and 17, respectively. In this context, the word *mountain* is used figuratively to suggest that *debt*, *money*, *cash*, or *bills* are present in a large and growing pile – a pile roughly resembling a mountain of rocks heaped upon rocks.

In addition to the collexemes associated with money and debt, the top of the table contains nouns connected with some work to do. *Work*, *laundry*, and *paper-work*, ranked third, fifth, and sixth, fall into this semantic category. These nouns are highly distinctive for the pattern in question in direct comparison with the pattern with *a/the sea of*, in which they are used extremely infrequently.

Another group in the ranking is constituted by nouns pertaining to the things that can be eaten by people or animals. *Food*, its central collexeme in rank 8, is followed by *potatoes*, *meat*, and *flesh* in ranks 12, 19, and 25. *A mountain of flesh* has two meanings in COCA. It may refer to 'a large quantity of the soft part of the body of a person or animal', or it may denote 'a fat man', as in "But I wasn't prepared for how massively fat he was − a flat-topped mountain of flesh" (COCA, FIC: Analog Science Fiction & Fact).

Table 3. The thirty most distinctive collexemes of the *mountain of*-NOUN-construction

rank	noun	a	c	e	f	x	y	z	b	d	(a)	(c)	*p*-values	coll. strength
1.	**evidence**	0	198	198	6042	4122	2118	6240	4122	1920	130.79	67.21	2.1340291864802483e-96	95.67
2.	**debt**	19	164	183	6057	4122	2118	6240	4103	1954	120.89	62.11	2.911793179099751e-57	56.54
3.	**work**	1	27	28	6212	4122	2118	6240	4121	2091	18.50	9.50	3.6221682150229795e-12	11.44
4.	**data**	21	49	70	6170	4122	2118	6240	4101	2069	46.24	23.76	7.954582554962355e-10	9.10

rank	noun	a	c	e	f	x	y	z	b	d	(a)	(c)	*p*-values	coll. strength
5.	**laundry**	0	16	16	6224	4122	2118	6240	4122	2102	10.57	5.43	2.9891962881596876e-8	7.52
6.	**paperwork**	3	22	25	6215	4122	2118	6240	4119	2096	16.51	8.49	3.161845506604641 6e-8	7.50
7.	**money**	9	26	35	6205	4122	2118	6240	4113	2092	23.12	11.88	0.0000016667370589759613	5.78
8.	**food**	0	12	12	6228	4122	2118	6240	4122	2106	7.93	4.07	0.0000022905242857723503	5.64
9.	**garbage**	5	20	25	6215	4122	2118	6240	4117	2098	16.51	8.49	0.0000029847335642339054	5.53
10.	**research**	1	13	14	6226	4122	2118	6240	4121	2105	9.25	4.75	0.00000743980210436 0489	5.13
11.	**cash**	4	17	21	6219	4122	2118	6240	4118	2101	13.87	7.13	0.000013054514926709481	4.88
12.	**potatoes**	0	10	10	6230	4122	2118	6240	4122	2108	6.61	3.39	0.000020012720786483177	4.70
13.	**material**	1	12	13	6227	4122	2118	6240	4121	2106	8.59	4.41	0.000020482318324265633	4.69
14.	**rubble**	3	15	18	6222	4122	2118	6240	4119	2103	11.89	6.11	0.00002308727675524678	4.64
15.	**trash**	2	13	15	6225	4122	2118	6240	4120	2105	9.91	5.09	0.000038322033732615866	4.42
16.	**dirt**	3	14	17	6223	4122	2118	6240	4119	2104	11.23	5.77	0.00005727802174361751	4.24
17.	**bills**	4	15	19	6221	4122	2118	6240	4118	2103	12.55	6.45	0.00007511523778837332	4.12
18.	**stuff**	4	15	19	6221	4122	2118	6240	4118	2103	12.55	6.45	0.00007511523778837332	4.12
19.	**meat**	1	10	11	6229	4122	2118	6240	4121	2108	7.27	3.73	0.00015242433408032034	3.82
20.	**documents**	0	8	8	6232	4122	2118	6240	4122	2110	5.28	2.72	0.00017463540556019385	3.76
21.	**gold**	11	21	32	6208	4122	2118	6240	4111	2097	21.14	10.86	0.00027066197295613 12	3.57
22.	**corpses**	0	7	7	6233	4122	2118	6240	4122	2111	4.62	2.38	0.00051563357786264 09	3.29
23.	**complaints**	0	7	7	6233	4122	2118	6240	4122	2111	4.62	2.38	0.00051563357786264 09	3.29
24.	**wood**	0	7	7	6233	4122	2118	6240	4122	2111	4.62	2.38	0.00051563357786264 09	3.29
25.	**flesh**	6	15	21	6219	4122	2118	6240	4116	2103	13.87	7.13	0.000657235773825704	3.18
26.	**papers**	3	11	14	6226	4122	2118	6240	4119	2107	9.25	4.75	0.00081282601135575 08	3,09
27.	**snow**	7	15	22	6218	4122	2118	6240	4115	2103	14.53	7.47	0.00117520540532018 83	2.93
28.	**files**	0	6	6	6234	4122	2118	6240	4122	2112	3.96	2.04	0.00152199797556833 8	2.82
29.	**knowledge**	1	7	8	6232	4122	2118	6240	4121	2111	5.28	2.72	0.00290262078393677	2.54
30.	**letters**	2	8	10	6230	4122	2118	6240	4120	2110	6.61	3.39	0.00384887525845233 4	2.41

The ranking list of distinctive collexemes for the construction under discussion also contains a range of nouns denoting some material, mass substances, or physical objects. *Garbage*, ranked 9, is followed by *material*, *rubble*, and *trash* in ranks 13, 14, and 15, respectively. In addition to these lexemes, other nouns occupying the subsequent positions, such as *dirt*, *stuff*, *gold*, *corpses*, and *wood*, can be assigned to this category.

Finally, apart from the nouns mentioned above, the bottom of the ranking list in Table 3 includes nouns pertaining to documents, papers, or pieces of writing. *Documents*, ranked 20, occupies the highest position among the lexemes belonging to this semantic category. It is followed by *complaints*, *papers*, *files*, and *letters* in ranks 23, 26, 28, and 30, respectively. These nouns are used metaphorically with the quantifier *a mountain of* to refer to 'documents or pieces of writing placed in a huge and growing pile that resembles a mountain'. For example, *a mountain of complaints* denotes 'a pile of written statements in which someone says they are not satisfied with something', while *a mountain of files* designates 'a pile of papers, documents or records'.

CONCLUSION

In conclusion, the findings indicate that some nouns combine with both partitive constructions (e.g. *debt*, *data*, *money*, *gold*, and *snow*), while others collocate either only with *a sea of* (e.g. *faces*, *red*, *troubles*, *green*, *grass*, *heads*, *stars*, or *darkness*) or solely with *a mountain of* (e.g. *evidence*, *laundry*, *food*, *potatoes*, *documents*, *corpses*, and *complaints*). A possible explanation for this discrepancy may lie in the semantics of both nominal partitives. The meaning of *a sea of* is a metaphorical extension of the basic sense 'a large area of salt water'; hence, it is motivated by the metaphor A LARGE QUANTITY OF SOMETHING IS A VAST EXPANSE OF THE SEA. In this construction, the word *sea* is used metaphorically to highlight that some entities, such as people (*faces*, *heads*), material objects (*stars*), plants (*grass*), colors (*red*, *green*), and abstract concepts (*troubles*, *darkness*) occur in a large quantity around a particular area, a quantity that resembles the vastness of the sea or a seascape with an infinite horizon. Thus, the construction in question exhibits a noticeable preference for entities that can be perceived as objects occurring in a wide area, a space roughly resembling the immensity of the sea. By contrast, the meaning of *a mountain of* seems to be a metaphorical extension of the concrete sense: 'a very large hill of rocks'. Thus, a very large pile of something (e.g., *evidence*, *laundry*, *food*, *potatoes*, *documents*) is conceptualized

in terms of a vertical orientation, i.e. a mountain of rocks heaped upon other rocks. In other words, this sense is motivated by the conceptual metaphor QUANTITY IS VERTICAL ELEVATION, a metaphor that is grounded directly in our embodied experience: physical objects that are perceived as being high are also large. In this metaphor, a quantity scale is understood as a verticality scale, while a measure of quantity is comprehended as a scalar position (Lakoff and Johnson).

Regarding the *sea of*-NOUN-construction, the findings indicate that the most distinctive collexemes are nouns connected with people (*faces*, *people*, *heads*, *humanity*, *fans*, *men*, *bodies*, *hats*, *suits*) and colors (*red*, *green*, *orange*). The other distinctive collexemes for this construction are countable nouns (*troubles*, *stars*, *particles*, *lights*, *flags*, *clouds*, *tears*) and uncountable nouns (*blood*, *grass*, *red ink*, *darkness*, *change*, *mud*, *oil*, *time*, *sand*, *glass*). All these metaphorical uses share the semantic property of *sea* which is understood as a vast expanse of something that is not salt water. In contrast to the pattern with *a sea of*, the *mountain of*-NOUN-construction exhibits a strong preference for entities that can be arranged into large and growing piles. Thus, it collocates with nouns connected with facts, information, or the study of data (*evidence*, *data*, *knowledge*, *research*), money or an amount of money that somebody owes (*debt*, *money*, *cash*, *bills*), some work to do (*work*, *laundry*, *paperwork*), the things that can be eaten by people or animals (*food*, *potatoes*, *meat*, *flesh*), some material, mass substances, or physical objects (*garbage*, *material*, *rubble*, *trash*, *dirt*, *stuff*, *gold*, *corpses*, *wood*), and documents, papers, or pieces of writing (*documents complaints*, *papers*, *files*, *letters*).

The results also seem to imply that the uses of both partitives and their noun collocates display varying degrees of metaphoricity (cf. Hanks). If a noun collocating with the partitives denotes an abstract entity, the metaphoricity seems to be greater, since there are no shared properties other than a vast expanse or a vertical elevation, respectively. For example, the metaphoricity of *a sea of troubles* and *a mountain of research* is greater than in *a sea of faces* and *a mountain of snow*, because of the greater abstractness of *troubles* and *research* than *faces* and *snow*. However, when a noun collocate refers to a physical entity, the metaphoricity appears to be weaker, because they share physical properties. For example, *faces* and *snow* share the feature of physical entities with *sea* and *mountain*, respectively, so that the semantic connection between the two combinations can be more readily established than in the case of *a mountain of research* and *a sea of troubles*, where noun collocates are abstract qualities.

The distinctive-collexeme analysis applied in this study has proved to be an effective technique for the identification of the most distinctive lexemes of both

constructions, and hence could be adopted elsewhere for the determination of the most significant lexemes occurring in different types of metaphorical constructions. Future research might, for example, focus on determining the nouns distinctive for other metaphorical partitives, such as *a torrent of*, *an ocean of*, or *a tsunami of*. In particular, it would be interesting to explore the distribution of nouns in these constructions across different types of both written and spoken registers, given the possible existence of slight variations in their occurrence.

BIBLIOGRAPHY

Croft, William, and D. Alan Cruse. *Cognitive Linguistics*. Cambridge UP, 2004.

Evert, Stefan. "Corpora and Collocations." Vol. 2: *An International Handbook*, edited by Anke Lüdeling and Merja Kytö, De Gruyter Mouton, 2009, pp. 1212-1248.

Goldberg, Adele. *Constructions at Work: The Nature of Generalization in Language*. Oxford UP, 2006.

Goldberg, Adele. "Constructionist Approaches to Language." *Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale, Oxford UP, 2013, pp. 15-48.

Gries, Stefan Th. "Frequencies, Probabilities, and Association Measures in Usage-/Exemplar-based Linguistics: Some Necessary Clarifications." *Studies in Language*, vol. 36, 2012, pp. 477-510.

Gries, Stefan Th. "More (Old and New) Misunderstandings of Collostructional Analysis: On Schmid and Küchenhoff (2013)." *Cognitive Linguistics*, vol. 26, no. 3, 2015, pp. 505-536.

Gries, Stefan Th., and Anatol Stefanowitsch. "Extending Collostructional Analysis: A Corpus-based Perspective on Alternations." *International Journal of Corpus Linguistics*, vol. 9, 2004, pp. 97-129.

Hanks, Patrick. "Metaphoricity is Gradable." *Corpus-Based Approaches to Metaphor and Metonymy*, edited by Anatol Stefanowitsch and Stefan Th. Gries, De Gruyter Mouton, 2008, pp. 17-35.

Hoffmann, Thomas. "Construction Grammars." *The Cambridge Handbook of Cognitive Linguistics*, edited by Barbara Dancygier, Cambridge UP, 2017, pp. 310-329.

Levshina, Natalia. *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. John Benjamins, 2015.

Lakoff, George, and Mark Johnson. *Metaphors We Live by*. U of Chicago P, 1980.

Steen, Gerard J, Aletta G. Dorst, J. Berenike Herrmann, Anna A. Kaal, Tina Krennmayr, and Tryntje Pasma. *A Method for Linguistic Metaphor Identification. From MIP to MIPVU*. John Benjamins, 2010.

Stefanowitsch, Anatol. "Collostructional Analysis." *Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale, Oxford UP, 2013, pp. 290-306.

DATA SOURCES AND TOOLS

The Corpus of Contemporary American English (COCA). Available from www.corpusdata.org/purchase.asp.

Fisher's Exact Test. Available from www.langsrud.com/fisher.htm.

The Macmillian English Dictionary and *Macmillian Thesaurus*. Available from www.macmillan dictionary.com.

A SEA OF-NOUN VS. *A MOUNTAIN OF*-NOUN: A QUANTITATIVE CORPUS-BASED STUDY OF TWO METAPHORICAL CONSTRUCTIONS

Summary

Applying the perspective of Construction Grammar and the corpus-based method known as *Distinctive-Collexeme Analysis*, this study seeks to identify nouns that indicate a strong preference either for the *sea of*-NOUN-construction or the *mountain of*-NOUN-construction. Based on the data extracted from the corpus, this study reveals that the constructions convey subtle nuances of meaning, reflect different degrees of metaphoricity, and have a strong tendency to occur with specific nouns.

Keywords: COCA; construction; distinctive collexeme analysis; Fisher's exact test

A SEA OF-NOUN VS. *A MOUNTAIN OF*-NOUN: ILOŚCIOWE BADANIE KORPUSOWE DWÓCH KONSTRUKCJI METAFORYCZNYCH

Streszczenie

Stosując podejście gramatyki konstrukcyjnej oraz metodę korpusową znaną jako analiza dystynktywno-koleksemowa, autor ma na celu zidentyfikowanie rzeczowników, które wykazują silne preferencje do łączliwości z konstrukcjami *a sea of*-NOUN lub *a mountain of*-NOUN. Na podstawie danych uzyskanych z korpusu studium ujawnia, że obie konstrukcje przenoszą subtelne niuanse znaczeniowe, odzwierciedlają różne stopnie metaforyczności i mają wyraźną tendencję do występowania z określonymi rzeczownikami.

Słowa kluczowe: COCA; konstrukcja; analiza dystynktywnych koleksemów; dokładny test Fishera