

ON THE GRADIENCE OF ENGLISH SIZE NOUNS: FREQUENCY, PRODUCTIVITY, AND EXPANSION

DAMIAN HERDA

Jagiellonian University, Poland

Abstract. The synchronic degrees of grammaticalization of size nouns are traditionally measured based on proportionate frequencies of their quantificational attestations in corpus samples. However, grammaticalization, in general, is associated not only with an increased frequency of grammaticalized uses but also with a rise in productivity and distributional expansion. Thus, drawing on corpus data encompassing selected English size nouns which originally individuate concrete inanimate nominals, this paper investigates the relationship between the three aforementioned parameters. Productivity is operationalized as the arithmetic mean of two measures, namely type-token ratio (TTR) and hapax-token ratio (HTR), i.e. the number of, respectively, types of quantified collocates and hapax legomena N2s divided by the number of all quantifier tokens of each expression, while host-class expansion is construed as the proportion of animate and abstract collocates among the respective items' quantifier uses. Contrary to expectations, the results reveal only a weak positive correlation between the elements' frequency values and their levels of productivity, and the same holds for the relation between frequency and distributional extension. Also surprising is the moderate negative correlation observed between productivity and expansion, which can nevertheless be elucidated in terms of a high type frequency of semantically general animate and abstract N2-collocates of the most distributionally extended expressions.

Key words: size nouns, vague quantifiers, grammaticalization, synchrony, type-token ratio, hapax-token ratio, corpus-based study

INTRODUCTION

Apart from standard vague quantifiers such as *a little* and *much/many*, English possesses an open class of so-called size nouns (cf. Brems, 2003, 2007, 2011), which may be employed to indicate non-specific, subjectively assessed small (cf. (1)–(2)) or large (cf. (3)–(4)) quantities of what the concomitant nominal refers to, as

illustrated below with examples taken from the 1.9 billion-token Corpus of Global Web-Based English (GloWbE).

- (1) Small pumpkins are easy to cook in a slow cooker with *a bit of water* or cut and roast them in an oven. (GloWbE)
- (2) People who go around accusing others of dishonesty without *a shred of evidence* disgust me. (GloWbE)
- (3) His acceptance speech was even surprisingly devoid of real forward momentum, despite *a lot of platitudes*. (GloWbE)
- (4) I've got a new job that is puuuuurty exciting for me and will be *heaps of work* so I've had to make some adjustments to my life. (GloWbE)

It has nonetheless been found that size nouns, most of which originally function partitively, i.e. serve to bound or unitize the concomitant nominals' reference (cf. Verreckken, 2015: 48), as exemplified by *scrap* in *some scraps of bread* or *heap* in *three heaps of stones*, exhibit synchronic gradience, i.e. differ internally in the extent to which they have grammaticalized in the quantifier function. In general, the synchronic degrees of grammaticalization of size nouns have been measured by determining the proportionate frequencies of their quantifier attestations, vis-à-vis basic partitive ones, in corpus samples (cf., among others, Brems, 2003, 2007, 2011; Delbecque and Verreckken, 2014; Verreckken, 2015).

However, grammaticalization, especially in functionalist approaches, tends to be associated not only with increased frequency, but also with distributional expansion and a rise in the grammaticalizing expression's productivity (cf. Himmelmann, 2004), operating on tokens and types of collocates. Thus, based on random samples of attestations of nine English size nouns, namely *bit(s)*, *scrap(s)*, *shred(s)*, *heap*, *heaps*, *load*, *loads*, *lot*, and *lots* (N = 2250), extracted from the GloWbE corpus, this paper examines the relationship between the three aforementioned parameters. Productivity will be represented by the arithmetic mean of two measures traditionally employed in morphological research, namely type-token ratio (TTR) and hapax-token ratio (HTR), i.e. the number of, respectively, types of quantified collocates and hapax legomena N2s divided by the number of all quantifier tokens of each expression, while host-class expansion is understood as the percentage of animate and abstract N2-collocates among the quantifier attestations of a particular size noun. The elements for analysis have been selected based on the existing literature, including an etymological dictionary (cf. Klein, 1966) and a preliminary corpus investigation. More specifically, the items, while having been partially dealt with in previous studies, originally function partitively in relation to concrete inanimate nouns and have all developed a purely quantificational meaning. As can be noted, there are three items suggesting small quantities and another three elements indicating large quantities. Furthermore, in the latter case, the plural forms are treated as separate items in view of the observation that pluralization may lead to substantial differences in the degree of grammaticalization of the singular forms and plural variants of 'large size' nouns (cf. Brems, 2003).

In accordance with the above, the present study aims to provide answers to three primary research questions, all of which revolve around the relationship between frequency, extension, and productivity. First, what is the correlation between the proportionate frequency of the quantifier uses of the scrutinized items and the extent of their distributional extension? Second, what is the correlation between the frequency of the expressions' quantifier attestations and their level of productivity? And third, what is the correlation between the elements' productivity and expansion values? Since all of the relevant phenomena may be expected to intensify with the progress of grammaticalization, a strong positive correlation coefficient is expected in each case.

The remainder of the paper is organized in the following way. Section 2 provides an account of the grammaticalization of size nouns into quantifiers. Section 3 describes the empirical material and the adopted methodology. Section 4 offers an analysis of naturally-occurring English data. Finally, Section 5 summarizes the main observations arrived at in the study as well as outlines prospects for further research on the topic.

THE GRAMMATICALIZATION OF SIZE NOUNS

As stated before, most size nouns originally function as partitives, i.e. the leftmost nominal elements in binominal (N1 *of* N2) syntagms whose function consists in 'bounding or unitizing the entities expressed by the second constituent' (Verveckken, 2015: 48). Particularly noteworthy as regards the semantics of partitive nouns is the fact that they exhibit more or less specific lexical requirements pertaining to the types of nouns with which they may co-occur (cf. Doetjes, 1997: 183–184; Brems, 2011: 133). For instance, all of the items under analysis here originally combine with concrete inanimate nominals.

Easily inviting scalar inferences (cf. Langacker, 1991: 88), partitives tend to evolve into vague quantifiers, which, in contrast to numerals, are 'imprecise in their specification of number or amount' (Jackson, 2013: 119). Huddleston and Pullum (2002: 365–366) further differentiate between *multal* and *paucal* vague quantifiers, which indicate, respectively, a non-specific high or low quantity. The very transition of nouns into quantifiers exemplifies grammaticalization (cf., among others, Brems, 2011; Verveckken, 2015; Giacalone Ramat, 2019), i.e. a process whereby lexical, more contentive expressions, in specific syntagmatic environments, develop more abstract, grammatical meanings (Hopper and Traugott, 2003: 1), whose initial phase manifests itself in semantic generalization (cf. Lehmann, 1985), or, more specifically, in 'the semanticization of quantifier meaning through repeated pragmatic inferencing of size or scalar implications that are part of the lexical semantics of the [size noun]' (Brems, 2011: 108). At the same time, this instance of grammaticalization involves subjectification since the newly emerged quantifiers convey 'meaning that indexes speaker-relatedness, in that quantifier meaning involves a speaker assessing size relative to a scale' (Brems, 2011: 231).

As can be expected, the semantic generalization of partitive nouns bears on their distributional patterns in a number of ways. First of all, partitives affected by grammaticalization lose compatibility with other quantifiers, including numerals (cf. Keizer, 2007: 136), the only exception here being paucal quantifiers functioning as negative polarity items, capable of co-occurring with the numeral *one*, which, in this case, performs an emphatic function (cf. Brems, 2007), as in (5).

(5) You have not provided *one shred of evidence* that any of those groups of people I have mentioned who rose up against a foreign occupying power, a foreign occupying power that consistently behaved with genocidal savagery against the people of the lands they had invaded became “new bosses just as bad as the old” you haven’t done it because you can’t. (GloWbE)

Moreover, grammaticalizing partitives undergo host-class expansion to novel N2-classes (cf. Himmelmann, 2004: 31–34), as vague quantifiers, in contrast to partitive nouns, typically do not exhibit any specific restrictions pertaining to the kinds of nouns they combine with. And finally, such nominal quantifiers differ from partitives in that the former can only be pre-modified by intensifying elements, such as *whole* (Brems, 2011: 201), e.g. *a whole heap of time* vs. *a neatly arranged heap of books*. Again, paucal quantifiers which exhibit negative polarity patterns are special in that it is not infrequent for them to combine with the superlative forms of adjectives invoking small size, as in (6), as well as with the adjective *single*, as in (7), both of which, just like the numeral *one* used in this context, serve to further reinforce the lack of what the N2 refers to (cf. Brems, 2007).

(6) I’d be disgusted if any journalist labelled Alex Salmond a serial child abuser on Twitter without *the slightest shred of evidence*. (GloWbE)

(7) And before you claim they were: Show me *a single scrap of evidence* that anyone depicted in the pictures above was ever arrested or questioned by the Secret Service. (GloWbE)

Rather than occurring all at once, the above-described changes are obviously extended in time, as is the case with all instances of grammatical evolution. Nevertheless, given the synchronic availability of ‘alternate strategies which enjoy different levels of grammatical autonomy’ (Lehmann, 1985: 309), grammaticalization, including the development of partitives into vague quantifiers, even though diachronic in nature, can in fact also be examined synchronically. As Haspelmath (2001: 16539) explains, ‘[s]ince grammaticalization is generally regarded as a gradual diachronic process, it is expected that the resulting function words form a gradient from full content words to clear function words’ (cf. also Hopper, 1991). Another fundamental observation here is that the grammaticalization process normally leads to an increased frequency of the grammaticalizing expression (cf. Hopper and Traugott, 2003), which means that it generally appears increasingly more often in language use (absolute frequency) and, more importantly, that the percentage of its

grammaticalized uses in samples of naturally-occurring data keeps increasing over time (proportionate frequency). Thus, the synchronic extent to which a partitive noun has grammaticalized is typically measured by determining the proportion of its grammaticalized attestations in corpus material (cf. Brems, 2011; Verwekken, 2015). For instance, it has been noted that in English, the plural forms of nominal quantifiers tend to display higher frequencies of grammaticalized uses than is the case with the singular ones (cf. Brems, 2003, 2011), which can be attributed to the intensifying effect of pluralization.

However, grammaticalization, especially in functionalist frameworks, is taken to manifest itself not only in increased (both absolute and proportionate) frequency, but also in distributional expansion and a rise in the grammaticalizing expression's productivity (cf. Himmelmann, 2004), conceived of as an enhanced collocational openness attendant upon semantic schematization. In particular, distributional extension and strengthened productivity are the main factors distinguishing between grammaticalization and lexicalization, the latter phenomenon instead involving freezing, i.e. a decrease in the pertinent expression's collocational freedom (cf. Brinton and Traugott, 2005). It is, therefore, worth investigating the relationship between the three grammaticalization-related phenomena, i.e. frequency, expansion, and productivity.

METHOD

As already mentioned in the introduction, the overarching aim of the present paper is to investigate the relationship between the grammaticalization parameters of nine English size nouns, namely *bit(s)*, *scrap(s)*, *shred(s)*, *heap*, *heaps*, *load*, *loads*, *lot*, and *lots*. The empirical material for the study was derived from the 1.9 billion-token Corpus of Global Web-based English (GloWbE), which, apart from representing different varieties of English, incorporates a significant amount of data reflective of informal registers, where language change is most likely to take place.

The first stage of the analysis consisted of the extraction of random samples of 250 adnominal attestations of each of the analyzed items (N = 2250) by means of the corpus search engine. As stressed before, the singular forms of multal quantifiers were treated separately from the plural ones. As far as small size nouns are concerned, the proportion of occurrences of the plural variants in the sample was left to chance, the assumption being that each attestation of this kind belongs to the partitive category. Thus, while the corpus queries in the former case were *heap of*, *heaps of*, etc., the commands used in the latter situation were *BIT of*, *SCRAP of*, and *SHRED of*, capitalization indicating that the pertinent item may be either in the singular or in the plural form. The tokens were then classified into (i) partitive, (ii) quantifier, and (iii) indeterminate uses. Partitive attestations were distinguished from quantifier ones on the basis of their distributional and semantic properties: in the former case, the N1s can be substituted with standard partitive nouns such as *piece(s)*, *fragment(s)*, *trace(s)*, etc., while in the latter, the pertinent items can be

replaced with canonical quantifiers such as *a little, some, many/much*, or, in the case of paucal quantifiers occurring in negative polarity settings, *any*. Indeterminate uses, in turn, include instances which, due to a shortage of co-textual clues, allow the partitive as well as the quantifier reading. Notably, attestations labelled by Brems (2003) as ‘valuing quantifier uses,’ e.g. *a vacuous and meaningless load of pap*, were treated here as partitive rather than quantifier ones on account of their distributional similarity to the basic uses of partitive nouns, i.e. practically unrestricted pre-modification patterns coupled with selectional requirements pertaining to N2-elements. More precisely, size nouns in their evaluative uses only take negatively charged animate or, more frequently, abstract N2s.

To further examine the distribution of the scrutinized elements, the N2-collocates in each subclass underwent additional labelling into (a) concrete inanimate (or simply concrete), (b) concrete animate (or simply animate), and (c) abstract. Notably, the percentage of animate and abstract collocates among the quantifier attestations of a given expression will be taken to reflect its level of host-class expansion. Next, two productivity measures were calculated, namely TTR (type-token ratio) and HTR (hapax-token ratio), i.e. the number of, respectively, types of quantified collocates and hapax legomena N2s divided by the number of all quantifier tokens of a particular form, and the arithmetic mean of the two values was established for each item. When determining the number of N2-types, only the head elements, without pre- and/or post-modifiers, were taken into account, which means that phrases such as *a shred of credible evidence* and *a shred of evidence that Iran has been weaponizing* were both analyzed as instantiating the N2-type *evidence*. Also, the abbreviated and the full form of a particular word, e.g., *info* and *information*, were both analyzed as exemplifying the same N2-type. Likewise, when the N2-slot was occupied by a series consisting of at least two nouns, only the linearly first, i.e. leftmost, one was taken into consideration. The only exception to the above-specified criteria were instances involving lexicalized phrases, such as *effing and blinding* (cf. *a bit of effing and blinding*) or *weak spot* (cf. *a bit of a weak spot*), which were obviously treated as inherent units. Finally, in the case of each expression, correlation coefficients were established for the relation between frequency and productivity, frequency and expansion, as well as productivity and expansion.

RESULTS

Table 1 presents the numerical results pertaining to the frequency (F), productivity (P), and extension (E) of the quantifier uses of the analyzed English size nouns. The values have been rounded up to two decimal places.

Further quantitative analysis of the results points to there being only a weak positive correlation between the frequency and the productivity of the scrutinized items’ quantifier uses ($r = .1171$), and the same holds for the relation between frequency and extension, although in the latter case, the correlation coefficient is slightly higher ($r = .3466$). This finding indicates that high-frequency values

need not, and often do not, go hand in hand with an advanced level of the other grammaticalization-related phenomena, which can be best illustrated with *bit* on the one hand and *shred* on the other: while the former exhibits comparably high degrees of frequency, productivity, and extension, the latter displays a moderate frequency of quantifier uses, a very low productivity in the quantifier function, and a very high distributional extension. Curiously, there is even a moderate negative correlation between productivity and extension ($r = -.4311$), which, as will be shown in the following parts of the text, can nonetheless be elucidated in terms of a high type frequency of some animate and abstract N2-collocates of the most distributionally extended items.

Table 1 Frequency, productivity, and extension values

Size noun	F	TTR	HTR	P	E
Bit(s)	0.81	0.85	0.77	0.81	0.92
Scrap(s)	0.10	0.65	0.46	0.56	0.65
Shred(s)	0.64	0.31	0.23	0.27	1
Heap	0.54	0.77	0.66	0.72	0.72
Heaps	0.87	0.67	0.54	0.61	0.71
Load	0.46	0.80	0.69	0.75	0.72
Loads	0.95	0.72	0.61	0.67	0.73
Lot	1	0.65	0.53	0.59	0.87
Lots	1	0.71	0.61	0.66	0.75

The following subsections offer a more detailed account of the empirical distribution of each of the analyzed expressions, and some qualitative comments are made as regards the items' distributional characteristics, with a focus on the most frequent N2-collocates in their quantifier attestations. What will serve as a point of departure for the qualitative discussion, which will be illustrated with examples derived from the investigated data, are the results shown in Table 2, revealing the collocability of the scrutinized elements with concrete inanimate (CI), concrete animate (CA), and abstract (A) N2s in each type of use.

Table 2 Empirical distribution of the analyzed items

Size noun	N2-type	Type of use		
		Partitive (#)	Quantifier (#)	Indefinite (#)
<i>Bit(s)</i>	CI	18	17	1
	CA	1	17	0
	A	25	168	3

Size noun	N2-type	Type of use		
		Partitive (#)	Quantifier (#)	Indefinite (#)
<i>Scrap(s)</i>	CI	153	9	5
	CA	2	0	0
	A	64	17	0
<i>Shred(s)</i>	CI	17	0	4
	CA	1	1	0
	A	68	159	0
<i>Heap</i>	CI	84	38	0
	CA	3	11	0
	A	24	87	3
<i>Heaps</i>	CI	28	63	2
	CA	0	25	0
	A	2	130	0
<i>Load</i>	CI	26	32	8
	CA	6	25	2
	A	91	59	1
<i>Loads</i>	CI	9	64	1
	CA	3	46	0
	A	0	127	0
<i>Lot</i>	CI	0	33	0
	CA	0	56	0
	A	0	161	0
<i>Lots</i>	CI	1	63	0
	CA	0	42	0
	A	0	144	0

1 BIT(S) OF N2

In its basic partitive uses, the item *bit* combines with concrete nouns denoting substances of solid consistency, such as *bread* (cf. Klein, 1966: 176; Traugott, 2008b). As shown in Table 1, *bit* is currently one of the partitives most frequently used in the quantifier function. In addition, the expression has undergone a considerable extent of distributional expansion to animate and abstract nominals (cf. Table 2).

With count animate and abstract collocates, *bit* typically takes a bounded complement, i.e. a singular noun form preceded by the indefinite article, carrying information about the relatively low degree of some gradable property implied by the relevant nominal, as in (10) and (11).

- (10) It's tough luck I'd say if people think you are *a bit of a religious nut* if you are one. (GloWbE)
- (11) The yellow tile scheme feels right to me, bringing *a bit of a big city terminal feel* without being annoying about it. (GloWbE)

Displaying a remarkably high level of productivity (cf. Table 1), the quantifier *bit* does not occur in many strong collocations. Nevertheless, its most frequent N2-collocates in the scrutinized data include the positively colored mass abstract nouns *fun* (6 occurrences), as in (12), and *luck* (also 6 occurrences), as in (13).

- (12) Well let's have *a bit of fun* shall we to examine this. (GloWbE)
- (13) Come back, creep down to the landing and, with *a bit of luck*, you should be able to pick off two of the monsters. (GloWbE)

Interestingly, *bit(s)* has likewise developed partitive uses involving abstract N2s, in which case *bit* is functionally equivalent to, and thus substitutable with, the general partitive noun *piece* (cf. (14)).

- (14) And he'd be taping *different bits of music* all the time. (GloWbE)

The above fact sometimes leads to interpretational ambiguities, as exemplified by (15), where the segment *a bit of* may be felicitously replaced with the partitive expression *a piece of* or the quantifier *some*:

- (15) But, I digress: there is *a bit of positive news* in the Apple slaps Samsung injunction. (GloWbE)

2 SCRAP(S) OF N2

Like *bit*, the item *scrap* in its basic partitive uses co-occurs with concrete nouns standing for solid substances, such as *meat* (cf. Klein, 1966: 1402; Brems, 2007). Yet, in contrast to *bit(s)*, *scrap(s)* generally prefers concrete over abstract nominals. In its scarce purely quantificational uses, however, *scrap* collocates with abstract nouns more frequently than it does with concrete ones (cf. Table 2).

Despite its low frequency of quantifier uses, *scrap* exhibits a moderate level of collocational openness in the quantifier function. Its recurrent collocates in the dataset include the abstract mass nouns *difference* (3 occurrences), as in (16), *evidence* (6 occurrences), as in (17), and *honor* (2 occurrences), as in (18), as well as the concrete mass noun *make-up* (2 occurrences), as in (19).

- (16) It's never good finding out someone has died, the medium of transferring that knowledge however makes not *a scrap of difference*. (GloWbE)
- (17) But there isn't *a scrap of evidence* for this, not even from the disingenuous spin he puts on the 12 year-old comments of a newspaper

columnist (nor, farcically, from the observation that the British team received their medals to the British national anthem). (GloWbE)

(18) Then I'd like to see the Lockerbie families return the millions that was shamefully extorted from Libya, as they surely must do if they have *a scrap of honour*. (GloWbE)

(19) I know for a fact that all the models that day looked fab without *a scrap of make-up* and how unfair is that! (GloWbE)

As can be seen above, when employed quantificationally, *scrap* typically functions as a negative polarity item, analogous to *any* rather than *a little* or *some*. In other words, when occurring in non-assertive contexts, the discussed expression emphasizes the non-attestation of what the N2 refers to (cf. Brems, 2007).

As can be seen in Table 2, *scrap(s)* is typically used partitively in relation to concrete inanimate mass nouns, as in (20). Like *bit(s)*, the scrutinized element has additionally developed quite frequent partitive uses involving mass abstract nominals, in which case it is functionally akin to expressions such as *piece(s)*, *trace(s)*, or *remnant(s)* (cf. (21)).

(20) Clutching a scrap of paper scrawled with the call number, I searched the shelves. (GloWbE)

(21) Those who only see scraps of misinformation should educate themselves first, before shooting their mouths off ... (GloWbE)

Marginally, *scrap* may be also used partitively in relation to count-to-mass coerced animate nouns, as in (22):

(22) Any pressure on a premmey's skin is painful, but I was assured that **the scrap of boy** before me would be comforted if you held a hand near, but not touching, his face. (GloWbE)

3 SHRED(S) OF N2

Like both *bit(s)* and *scrap(s)*, the partitive *shred(s)* originally combines with concrete collocates referring to solid substances (cf. Klein, 1966: 1439; Traugott, 2008b). Yet, similarly to *bit(s)*, and in contrast to *scrap(s)*, *shred(s)* reveals a general predilection for abstract NPs, and the same applies to its quantifier attestations (cf. Table 2).

What is quite striking about *shred*, however, is that despite being relatively often used as a quantifier, it displays a conspicuously low productivity level (cf. Table 1), as almost half of its N2-collocates instantiate one N2-type, namely *evidence* (72 occurrences), as illustrated by (2), (5), and (6). Among the other habitual N2-collocates of the quantifier *shred of N2* are the mass abstract nouns *decency* (7 occurrences), as in (23), *proof* (also 7 occurrences), as in (24), *sense* (5 occurrences), as in (25), and *truth* (6 occurrences), as in (26).

- (23) Disgusting... any professional musician with a *shred of decency* would
1) not ask musicians to play for free, and 2) NOT do this show. (GloWbE)
- (24) And Vice President Dick Cheney continues to say without
a *shred of proof* that there is “overwhelming evidence” justifying
the administration’s pre-war charges. (GloWbE)
- (25) Doesn’t make a *shred of sense*, might as well be a lifetime ban...
(GloWbE)
- (26) Is there a *shred of truth* there? (GloWbE)

As can be seen above, analogously to *scrap*, and as opposed to *bit*, the quantifier *shred* normally functions as a negative polarity item (cf. *any*). Another commonality between the two expressions manifests itself in the fact that both typically quantify over abstract nouns encoding epistemic notions, even though the tendency is much more pronounced in the case of *shred*. In the investigated data, the item under discussion has additionally been found to combine with one proper name with an animate human referent, in which case the negative segment *not a shred of* can be paraphrased as *nothing of* (cf. (27)).

- (27) Charles Johnson on Andrew Sullivan, 5 Apr 2008: “Not a *shred of the post-9/11 Sullivan* remains; all that’s left is a rhetoric-spewing empty shill for “progressive” causes.” (GloWbE)

Moreover, like both *bit(s)* and *scrap(s)*, *shred(s)* has developed partitive uses involving abstract nominals, where it bears semantic resemblance to the general partitive noun *piece* (cf. (28)).

- (28) Not to condone for one second what Brown has done in the past (and to reiterate: *every shred of evidence* points to the fact that he’s a wang of near-galactic proportions), but nobody apart from Louis Walsh wants to be entertained by a bunch of fucking Blue Peter presenters, do they? (GloWbE)

4 HEAP OF N2

The item *heap* originally combines with nominals denoting stackable entities or masses (cf. Klein, 1966: 710–711; Brems, 2012). Indeed, in the analyzed data, *heap* most frequently combines with concrete inanimate nominals, although in its quantifier uses, it displays a preference for abstract NPs (Table 2).

The most frequent N2-collocates of *heap* in its quantifier attestations include the emotively loaded mass abstract nouns *fun* (5 occurrences), as in (29), and *trouble* (6 occurrences), as in (30), as well as the non-count concrete noun *money* (5 occurrences), as in (31).

- (29) The first full day in Kyoto I spent cruising around on my unladen bike, which was a *whole heap of fun*! (GloWbE)

(30) Kevin Wu (aka Youtuber, KevJumba) and Dante Basco seem to be getting into *a whole heap of trouble* in the trailer for their latest flick, *Hang Loose*. (GloWbE)

(31) The NBN will make *a whole heap of money* and that profit will be used to pay back the debt. (GloWbE)

Additionally, *heap* occasionally occurs in metaphorical partitive uses involving abstract NPs, such as (33). Attestations of this kind crucially rely on the expression's basic meaning coupled with its tendency to indicate portions of waste substances in its standard partitive occurrences, as exemplified by (32).

(32) He could find something fascinating anywhere, from a stunning high alpine lake to a neighborhood garden to *a heap of junk* along the side of the road, and would always take time to check it out. (GloWbE)

(33) It remains to be seen, however, whether the public-health community will give this landmark work due credit or continue to rubber stamp an outdated policy that, like bloodletting and trepanation, properly belongs on *the scrap heap of sham medical interventions*. (GloWbE)

5 HEAPS OF N2

Heaps generally reveals a slight preference for abstract over concrete (inanimate and animate) nominals. Likewise, the expression displays a propensity to quantify over abstract NPs (Table 2).

Interestingly, while the extension parameter is almost equal in the case of both *heap* and *heaps*, the productivity level of the latter is slightly lower than that of the former (cf. Table 1), which indicates that the quantifier *heaps* participates in stronger collocations. The most frequent N2-collocates of the expression at issue are the mass abstract noun *fun* (16 occurrences), as in (34), *room* (5 occurrences), as in (35), *stuff* (5 occurrences), as in (36), the count animate noun *people* (6 occurrences), as in (37), and the mass concrete nouns *food* (6 occurrences), as in (38), and *money* (11 occurrences), as in (39).

(34) The workshop was *heaps of fun*, with everyone making mini comics and hanging out. (GloWbE)

(35) I love it. *Heaps of room*, nice to drive. (GloWbE)

(36) You can count on us for tickets to the sweetest gigs and *heaps of FREE stuff*. (GloWbE)

(37) "It doesn't get any better than this, I get to do what I love and meet *heaps of new people* along the way." (GloWbE)

(38) On top of the sporting events -- there was a big cheerleading competition, a marching band and *heaps of free food*. (GloWbE)

(39) They're contracted by the government to do these things and it costs *heap of money* for contractors to go out to communities. (GloWbE)

6 LOAD OF N2

In its basic partitive uses, *load* refers to a collection of goods or a portion of a substance carried at a time (cf. Klein, 1966: 899). In the dataset under scrutiny, however, *load* reveals an overall preference for abstract N2-complements, even though its purely quantificational attestations are almost equally distributed between concrete (inanimate and animate) and abstract nominals (cf. Table 2).

With its relatively high level of productivity, the quantifier *load* does not occur in many strong collocations. The most frequent N2-elements quantified by *load* include the nouns *money* (5 occurrences), as in (40), *stuff* (5 occurrences), as in (41), and *things* (4 occurrences), as in (42):

(40) If you have Microsoft shares, sell them now, or you will lose a *bucket load of money*. (GloWbE)

(41) Now it is worth mentioning that there is a *whole load of stuff* I want to say about my experience with this particular dealership but none of it is relevant to online communications. (GloWbE)

(42) "He was fantastic to discuss about batting one to one. But on a *load of things* I would not follow him," he said, to peels of laughter from the gathering. (GloWbE)

As already suggested by (40), *load* displays a high degree of conceptual persistence, reflected in its pre-modification patterns. In other words, the size noun *load* tends to be modified by elements which, while intensifying its inherent scalar implications, reinvoke its lexical meaning, implying what can be generally labelled as containers, where literal loads can be kept or carried. Among such items observed in the data, apart from *bucket*, are *barrel*, as in (43), and *shed*, as in (44).

(43) And it's a *barrel load of fun*. (GloWbE)

(44) A great driver who never quite lived up to his potential? Or the greatest of the modern era who won a *shed load of championships*? Perhaps the greatest ever? (GloWbE)

Another important finding as regards *load* is that the expression very frequently occurs in evaluative uses involving abstract nominals, such as (45), which, in this paper, are analyzed as a special kind of partitive attestations on account of their limited distribution:

(45) The realisation that some people have two jobs is not an astounding new find published by esteemed sociologists -- but a *vacuous and meaningless load of pap* that's made the front page because this keen editor is proud at having coined a new term. (GloWbE)

7 LOADS OF N2

Loads of N2 displays a slight general preference for abstract over concrete (inanimate and animate) N2-collocates, and the same tendency applies to its quantifier attestations (cf. Table 2). That *loads* exhibits an overall lower percentage of uses with abstract N2-complements, and at the same time a markedly higher proportion of purely quantificational occurrences, than is the case with *load* stems from the latter's frequent appearance in evaluative attestations involving nouns such as *nonsense*, and an apparent lack thereof as far as the former is concerned.

Nevertheless, the quantifier *loads* shares two of the most common N2-collocates with *load*, namely the nouns *money* (10 occurrences), as in (46), and *things* (8 occurrences), as in (47). In addition, *loads* habitually quantifies over the animate noun *people* (16 occurrences), as in (48).

(46) They don't want to invest their own money and they want to make *shed loads of money* by selling our best players -- its fact pure and simple. (GloWbE)

(47) *Loads of other things*, it's actually an interesting question once you consider all the factors... (GloWbE)

(48) Although many folks try and eat a healthy diet nowadays you are going to discover that *loads of people* still don't get the nutrition they require. (GloWbE)

8 LOT OF N2

In its basic partitive uses, the item *lot* refers to a parcel of land (cf. Klein, 1966: 907). However, in the material under analysis, *lot*, like *load*, generally exhibits a clear preference for abstract N2-collocates over concrete (inanimate and animate) ones (cf. Table 2), yet in contrast to the latter, the former has not developed evaluative uses.

The collocability of the quantifier *lot* nonetheless resembles that of both *load* and *loads*, as in the data under analysis, *lot* most frequently quantifies over the nouns *money* (10 occurrences), as in (49), *people* (22 occurrences), as in (50), *things* (9 occurrences), as in (51), and *time* (5 occurrences), as in (52).

(49) Hello rich people, we normally hijack *a lot of your money* but now we'll hijack a little less of your money. (GloWbE)

(50) There are *a lot of much smarter people* doing a lot more successful ventures than me, but for a hobby, I really do love it, and it completely reinvigorated my interests in programming. (GloWbE)

(51) Older people are always impressed when they meet a young man or girl that works and create cool things and they are also willing to teach you *a lot of things* as well as to help you. (GloWbE)

(52) In my opinion teachers are the second mothers for the students because students spend *a lot of time* with their teachers.

What may be surprising is the fact that *lot*, despite its remarkable frequency of quantifier attestations, has been shown to display a lower degree of productivity than *loads*, which is nevertheless counterbalanced by the former's greater extent of host-class expansion. This finding can be accounted for in view of the observation that a higher expansion value implies a higher type frequency of semantically general animate and abstract N2-collocates such as *people* and *things*, respectively, which, in turn, translates into a lower productivity level.

9 LOTS OF N2

Like all *heaps*, *load*, *loads*, and *lot*, *lots* more frequently combines with abstract complements than it does with concrete (inanimate and animate) ones (cf. Table 2). Interestingly, as in the case of both *loads* and *lot*, the most frequent N2-collocates of *lots* include the nouns *money* (5 occurrences), as in (55), *people* (23 occurrences), as in (56), and *things* (10 occurrences), as in (57).

(53) People assume that there's *lots of money* in the game because it's a professional sport but every cent somehow finds a home quickly enough. (GloWbE)

(54) Through that online community I have "met" *lots of people* who I keep in touch with via facebook mainly. (GloWbE)

(55) I'd say after that things fragmented and now there are *lots of interesting little things* going on but seem slightly underdeveloped, less ambitious or meaningful than those giants of '05 either that or I turned thirty (GloWbE)

Notably, *lots* displays a higher productivity level than does *lot*, which is nonetheless coupled with the former's lower value of host-class expansion. This finding further substantiates the claim that at least in the case of plural nominal quantifiers, wide collocability with animate and abstract nouns tends to entail a high type frequency of items with general meanings.

CONCLUSION

Drawing on empirical data derived from the Corpus of Global Web-based English (GloWbE), this paper examined the grammaticalization patterns of nine size nouns, namely *bit(s)*, *scrap(s)*, *shred(s)*, *heap*, *heaps*, *load*, *loads*, *lot*, and *lots*. The research objectives included determining the relationship between the expressions' frequencies of quantifier attestations, degrees of productivity, as well as extents of host class-expansion.

Contrary to the assumption that the three grammaticalization parameters should be strongly positively correlated, the obtained results reveal only a weak positive correlation between the frequency and the productivity of the items'

quantifier uses ($r = .1171$), and the same holds for the relation between frequency and extension ($r = .3466$), which indicates that, at least as far as size nouns are concerned, high-frequency values typically do not go hand in hand with an advanced level of the other grammaticalization-related phenomena. Perhaps even more surprising at first glance is the moderate negative correlation observed between productivity and extension ($r = -.4311$), a finding which can nonetheless be explained in terms of a high type frequency of semantically general animate and abstract N2-collocates of the most distributionally extended constructions. Particularly noteworthy in this context is the paucal quantifier *shred*, which, despite exhibiting the highest extension value of all the expressions under scrutiny, displays the lowest productivity in the quantifier function, as almost half of its N2-collocates in the quantifier uses instantiate one N2-type, namely *evidence*.

Since the scrutinized quantifiers, especially the multal ones, share a number of their most frequent N2-collocates, such as *money*, *people*, *fun*, and *things*, a next step in the research on their grammaticalization should involve a fine-grained diachronic study aimed at elucidating the extent to which the expressions' collocability has been shaped by the forces of paradigmatic analogy (cf. Fischer 2011). The same applies to the paucal quantifiers *scrap* and *shred*, which, in addition to typically appearing in negative polarity contexts, tend to collocate with epistemic nominals. Future studies on the topic should likewise include an investigation into the adverbialization of the analyzed size nouns, a development which has been observed to constitute the subsequent phase of their grammatical evolution (cf. Traugott, 2008a,b), and which has not yet been systematically examined based on extensive corpus data (cf. De Clerck and Brems, 2016).

ACKNOWLEDGEMENT

This research was funded by the National Science Center of Poland (Grant No. 2019/33/N/HS2/01695).

REFERENCES

- Brems, L. (2003) Measure noun constructions: an instance of semantically-driven grammaticalization. *International Journal of Corpus Linguistics*, 8 (2): 283–312.
- Brems, L. (2007) The grammaticalization of small size nouns: Reconsidering frequency and analogy. *Journal of English Linguistics*, 35: 235–293.
- Brems, L. (2011) *The Layering of Size Noun and Type Noun Constructions in English*. Boston, Berlin: Mouton de Gruyter.
- Brems, L. (2012). The establishment of quantifier constructions for size nouns: A diachronic case study of *heap(s)* and *lot(s)*. *Journal of Historical Pragmatics*, 13 (2): 202–231.
- Brinton, L. J. and Traugott, E. C. (2005) *Lexicalization and Language Change*. Cambridge: Cambridge University Press.

- De Clerck, B. and Brems, L. (2016) Size nouns matter: a closer look at *mass(es)* of and extended uses of SNs. *Language Sciences*, 53: 160–176.
- Delbecque, N. and Verwecken, K. D. (2014) Conceptually-driven analogy in the grammaticalization of Spanish binominal quantifiers. *Linguistics*, 52 (3): 637–684.
- Doetjes, J. (1997) *Quantifiers and Selection: on the distribution of quantifying expressions in French, Dutch and English*. Leiden: University of Leiden.
- Fischer, O. (2011) Grammaticalization as analogically driven change? In H. Narrog and B. Heine (eds.) *The Oxford Handbook of Grammaticalization* (pp. 31–42). Oxford: Oxford University Press.
- Giacalone Ramat, A. (2019) Degrees of grammaticalization and measure constructions in Italian. *Revue Romane*, 54 (2): 257–277.
- Haspelmath, M. (2001). Word classes and parts of speech. In P. B. Baltes and N. J. Smelser (eds.). *International Encyclopedia of the Social and Behavioral Sciences* (pp. 16538–16545). Amsterdam: Pergamon.
- Himmelman, N. P. (2004) Lexicalization and grammaticization: Opposite or orthogonal? In W. Bisang, N. P. Himmelman, and B. Wiemer (eds.). *What Makes Grammaticalization? A Look from its Fringes and Its Components* (pp. 21–42). Berlin, New York: Mouton de Gruyter.
- Hopper, P. J. (1991) On some principles of grammaticization. In E. C. Traugott and B. Heine (eds.) *Approaches to Grammaticalization: Volume 1* (pp. 17–35). Amsterdam, Philadelphia: John Benjamins.
- Hopper, P. J. and Traugott, E. C. (2003) *Grammaticalization*. New York: Cambridge University Press.
- Huddleston, R. and Pullum, G. (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Jackson, H. (2013) *Grammar and Meaning: a semantic approach to English grammar*. London, New York: Routledge.
- Keizer, E. (2007) *The English Noun Phrase: The Nature of Linguistic Categorization*. Cambridge: Cambridge University Press.
- Klein, E. (1966) *A Comprehensive Etymological Dictionary of the English Language*. Amsterdam, London, New York: Elsevier.
- Langacker, R. (1991) *Foundations of Cognitive Grammar. Volume 2: descriptive application*. Stanford: Stanford University Press.
- Lehmann, C. (1985) Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile*, 20: 303–318.
- Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985) *A Comprehensive Grammar of the English Language*. London, New York: Longman.
- Traugott, E. C. (2008a) Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In R. Eckardt, G. Jäger and T. Veenstra (eds.) *Variation, Selection, Development – Probing the Evolutionary Model Language Change* (pp. 219–250). Berlin, New York: Mouton de Gruyter.
- Traugott, E. C. (2008b). The grammaticalization of NP of NP patterns. In A. Bergs and G. Diewald (eds.) *Constructions and Language Change* (pp. 23–45). Berlin, New York: Mouton de Gruyter.
- Verwecken, K. D. (2015) *Binominal Quantifiers in Spanish: conceptually-driven analogy in diachrony and synchrony*. Boston, Berlin: Mouton de Gruyter.

TOOL

[GloWbE] Davies, M. (2013) Corpus of Global Web-based English: 1.9 Billion Words from Speakers in 20 Countries. Available from <https://www.english-corpora.org/glowbe/> [Last accessed on 25 July 2021].

Damian Herda (MA in English Studies and Swedish Philology) is a PhD candidate in Linguistics at the Jagiellonian University in Kraków, Poland. His main research interests include grammaticalization theory and lexical semantics. ORCID ID: <https://orcid.org/0000-0002-0925-1805>.

E-mail: damian.herda@uj.edu.pl