
A Corpus-Based Study of the Grammaticalization of Partitive *a body of*

Huiqing Huang, Qingshun He*

School of Foreign Languages, Sun Yat-Sen University, Guangzhou, China

Email address

markman1998@163.com (Qingshun He)

*Corresponding author

Citation

Huiqing Huang, Qingshun He. A Corpus-Based Study of the Grammaticalization of Partitive *a body of*. *Language, Literature and Culture*. Vol. 2, No. 1, 2019, pp. 16-22.

Received: January 20, 2019; Accepted: March 25, 2019; Published: April 9, 2019

Abstract: During the past decades, grammaticalization has been an important topic for linguistic studies, and size noun constructions, or binominal quantifiers have been taken by researchers as an ideal entry point for investigating the phenomenon. The result is the emergence of many studies focusing on the grammaticalization of size noun constructions in different languages, including English, Spanish, and Polish. Most of these studies attempt to explain the mechanism or motivation behind grammaticalization and many are corpus-based. Being equally aware of the advantages of a corpus-based approach to the study of a grammaticalization process, in this research, we intend to investigate the grammaticalization of the English partitive *a body of*, a construction belonging to the category of size noun constructions and not yet touched upon so far, and its stylistic preference based on evidences derived from two corpora, the COHA and the BNC. The findings of our research include: (1) Partitive *a body of* has the potential to be fully grammaticalized, but currently this process has shown no sign of completion; (2) Partitive *a body of* tends to be used mostly in formal academic texts rather than in informal spoken texts. These findings can be explained by the association and dissociation of the semantic focus and logical focus within the target construction *a body of*.

Keywords: Grammaticalization, Corpus-Based, *a body of*, Stylistic Preference

1. Introduction

How is the common noun *body* in example (1a) used as a partitive noun in example (1c), and what difference can we find from the two *body of* expressions in examples (1b) and (1c)?

(1) a. The stamp is pushed, rather than kicked, into the *body of* the attacker. (BNC_MISC)

b. Formerly known as managing clerks, this *body of* employees has been known as legal executives since 1963. (BNC_AC)

c. This social context is not one where a *body of* intellectuals think up resolutions to an objectively defined urban crisis and then make this knowledge available to the wide church of policy makers and practitioners in the best enlightenment spirit. (BNC_AC)

By what grammatical steps does this shift of usage take place? This may be explained by grammaticalization. Grammaticalization is described as the evolution of grammatical elements from lexical sources [1]. In example (1a), *body* is a head noun, and *the attacker* functions as its

post-modifier. In example (1c), *a body of* is a partitive construction modifying *intellectuals* and hence can be taken as an instance of grammaticalization. However, *body* in example (1b) lies in between; it may have either explanation.

We can hereby hypothesize that the partitive construction *a body of* has experienced a process of grammaticalization and it is sensitive with regard to text style. To test this hypothesis, we will conduct a corpus-based study on the diachronic and synchronic distributions of the partitive construction *a body of*.

We will review grammaticalization in literature in Section 2, and introduce the corpora used in this research and data collection in Section 3. The findings of the research on the diachronic and synchronic distributions of the *a body of* construction will be shown in Sections 4 and 5 respectively. A discussion on the findings is presented in Section 6.

2. Literature Review

The term *grammaticalization* is believed to be coined by Meillet [2] in the early 20th century although phenomena

concerning grammaticalization had been noticed and observed with a longer history than that. The study of grammaticalization is therefore considered a young sub-field in the modern studies of language. Ever since the creation of the term, the study of grammaticalization has experienced three critical stages of development. The early 1970s marks the beginning of the study of grammaticalization when Givón [3] emphasized the importance of knowledge about the historical development of languages for understanding the language structure. The status of grammaticalization as an independent field within the study of language was cemented by Heine et al. [4] and Hopper & Traugott [5] in the 1990s.

Since the beginning of this century, the study of grammaticalization has entered into a new phase to associate with other fields in linguistic studies. Although seriously criticized, grammaticalization has been increasingly accepted by linguists from both the functional and the formal schools and has been extended to regions outside the traditional centers of linguistics [6]. Today, grammaticalization is a widely discussed topic especially when language change is concerned. Theories and studies of grammaticalization have been a major source of inspiration for recent studies of language change [7].

Grammaticalization is a process leading from lexemes to grammatical formatives [8]. In accordance with this definition, the primary goal of grammaticalization theory is to describe how grammatical forms and constructions arise and develop through space and time, and to explain why they are structured the way they are [9]. Since grammaticalization is a diachronic process, evidences derived from records of the earlier stages of development of a language are helpful to verify or falsify any grammaticalization hypothesis made about that language. This is where corpus linguistics comes to play an important role in grammaticalization studies.

Today, quantitative studies based on corpora have become an important approach to the investigation of grammaticalization. The combination of these two sub-fields within linguistics is recognized and approved by more and more linguists. Mair [10] sees that corpus linguistics and grammaticalization theory share considerable common grounds and that corpus linguistics can bring at least three other benefits to the study of grammaticalization besides hypotheses testing:

First, they make it possible to study incipient or ongoing processes of grammaticalisation. Secondly, a quantitative-cum-qualitative analysis of corpus data makes it possible to shed light on important theoretical issues. Thirdly, corpora allow a better take on the text linguistic, genre and discourse factors relevant to grammaticalisation. [10]

Many well-done corpus-based studies have made contributions to the development of grammaticalization as an independent field in the modern studies of language by discovering aspects that would not have been noticed through traditional qualitative approaches and by verifying currently available theories and findings in the field. Wu et al. [11] conducted a corpus-based quantitative study on the grammaticalization of *be going to* to investigate its realization

time of full grammaticalization. Tagliamonte [12] studied the grammaticalization of forms used for the expression of necessity and/or obligation such as *have to* and their variations with the York English Corpus. Based on the data collected from the English-Swedish Parallel Corpus, Aijmer [13] described the meaning and uses of the English verb phrases *be able to* and *manage to*. With evidences drawn from at least nine computerized corpora to illustrate his discussion, Rissanen [14] described the development and grammaticalization of the preposition and conjunct *beside(s)* in English.

To investigate the process of grammaticalization, many researchers have taken size noun constructions, or binominal quantifiers as their starting point as these constructions are often viewed as the locus of grammaticalization processes [15-17]. Size noun constructions are “structures that incorporate a size noun expression, i.e. a nominal expression that describes size or shape, implying a measure” [18]. They can be schematized as (D) (M) N1 *of* (D) (M) N2¹. Basically, the concept of binominal quantifiers overlaps with that of size noun constructions. One the one hand, it is based on the definition of quantifying nouns, which are “nouns that refer to containers (*a barrel, a mouthful*, etc.), configurations of masses (*a heap, a pile, a bunch*, etc.) or collectives (*a flock, a swarm*, etc.) when used as a lexical head” [15]. A binominal quantifier is formed “when a prepositional phrase is added to specify the constituents” [15]. On the other hand, the structure suggested above for size noun constructions can be applied to binominal quantifiers perfectly.

In recent years, most of the published works on the grammaticalization of size noun constructions or binominal quantifiers in different languages have attempted to explain the mechanism behind the phenomenon and adopted a corpus-based approach. For example, Herda [19] conducted a corpus investigation of the grammaticalization, specifically the delexicalization, of quantifying nouns *pile* and *stack* in English and their Polish counterparts *streta* and *stos*. She argues that the conceptual mechanism behind the delexicalization of the quantifying nouns is metonymy. Verveckken [20], however, considered analogy as a mechanism or a motivation for diachronic change of language. Based on corpus data, he studied the role of analogy in the grammaticalization process of seven Spanish binominal quantifiers. De Clerck & Brems [21] investigated the grammaticalization of size nouns *mass of* and *masses of* and their various uses caused by the process with data derived from corpora. At the same time, they looked at the grammaticalization of size nouns in general from a broader perspective.

Being equally aware of the advantages of a corpus-based approach and considering that, to our knowledge, the grammaticalization process of English partitive *a body of* has not yet been touched upon, we intend to investigate the grammaticalization of this particular construction and its

1 (D), (M) and N1 are optional determiner, optional modifier and size noun respectively.

stylistic preference.

Grammaticalization is always linked to the process of semantic bleaching, phonetic erosion or morphological reduction. Semantic bleaching can be described as the loss of semantic content [22]. The grammaticalization of *a body of* can thus be described as the process of semantic bleaching of the noun *body* in this expression.

3. Methodology

3.1. Corpora

A corpus is a collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written texts or tape recordings, which have been collected for linguistic study [23]. Recently, this concept has been further explained as being used as collections of texts (or parts of text) that are stored and accessed electronically [23]. Corpora are designed for different linguistic purposes and these purposes in turn decide what a corpus will be like, including its size, the types of texts included and etc. Corpora offer us information about frequency of a specific word or construction, the co-text and context where it is used, its collocations and the historical development of its frequency as well as usage.

In this study, we will use the Corpus of Historical American English (COHA) and the British National Corpus (BNC) to examine the historical development of the partitive construction *a body of* in English and its stylistic preference respectively. We choose the COHA because it is a corpus for the study of historical English, containing more than 400 million words and covering a time span of 200 years from the 1810s to the 2000s. Each of the decades forms a sub-corpus and they are all well-balanced by genre. With such features, the COHA allows researchers to examine a wide range of changes in English with much more accuracy and detail than with any other available corpus [24].

The BNC was originally created by Oxford University

Press in the 1980s to early 1990s, containing 100 million words of text. We choose this corpus to examine the stylistic preference of our target construction *a body of* because it covers a wide range of genres, such as spoken, fiction, magazine, newspaper, non-academic, academic and miscellaneous.

3.2. Data Collection

In order to collect data for our study, an exhaustive concordance of the construction “*a body of* NOUN” was performed in the COHA and we get the raw frequencies and the normalized frequencies of per million words of the target construction in each of the 20 decades as shown in Table 1.

Table 1. Diachronic Distribution of Partitive *a body of* in the COHA.

	Raw Frequency	Normalized Frequency
1810s	9	7.50
1820s	51	7.39
1830s	69	5.00
1840s	101	6.31
1850s	93	5.64
1860s	84	4.91
1870s	109	5.86
1880s	121	5.79
1890s	129	6.08
1900s	102	4.53
1910s	74	3.26
1920s	66	2.58
1930s	43	1.76
1940s	28	1.16
1950s	25	1.02
1960s	38	1.59
1970s	28	1.18
1980s	31	1.23
1990s	32	1.15
2000s	30	1.02

The types of the nouns that immediately follow *a body of* and their frequencies were also obtained as shown in Table 2.

Table 2. Diachronic distributions of typical nouns following partitive *a body of* in the COHA (normalized frequency).

	man	water	troop	knowledge	Indian	law	cavalry
1810s	5.08	0	0.85	0	0	0	0
1820s	2.45	0.58	0.43	0	0	0.29	0.29
1830s	1.52	0.07	0.22	0	0.22	0.07	0
1840s	1.43	0	0.37	0	0.50	0.12	0.12
1850s	1.76	0.30	0.36	0	0.24	0.18	0.24
1860s	0.94	0.18	0.29	0	0.12	0.18	0.12
1870s	1.51	0.43	0.16	0	0	0.16	0.16
1880s	1.72	0.44	0.15	0	0.10	0	0.10
1890s	1.84	0.24	0.44	0	0.05	0	0.19
1900s	0.86	0.09	0.05	0.05	0	0.05	0
1910s	0.66	0.04	0.09	0.09	0	0.04	0.04
1920s	0.23	0.16	0	0.08	0	0.04	0
1930s	0.33	0.24	0.04	0.04	0.04	0.04	0.04
1940s	0.12	0	0.04	0.12	0	0	0.04
1950s	0.04	0.20	0	0.12	0	0	0
1960s	0.21	0.29	0	0.21	0.04	0	0
1970s	0.17	0.38	0	0	0.04	0.04	0
1980s	0.04	0.28	0	0.16	0	0.04	0
1990s	0.04	0.07	0	0.11	0	0.04	0
2000s	0	0.20	0.03	0.03	0	0.07	0

Table 3. Genre distributions of *a body of* and *a lot of* in the BNC.

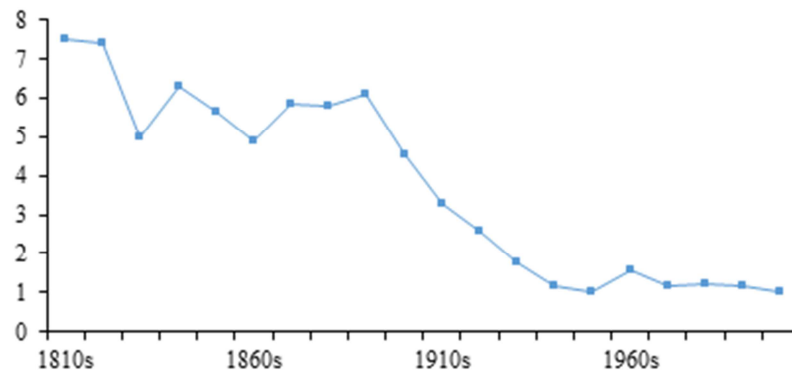
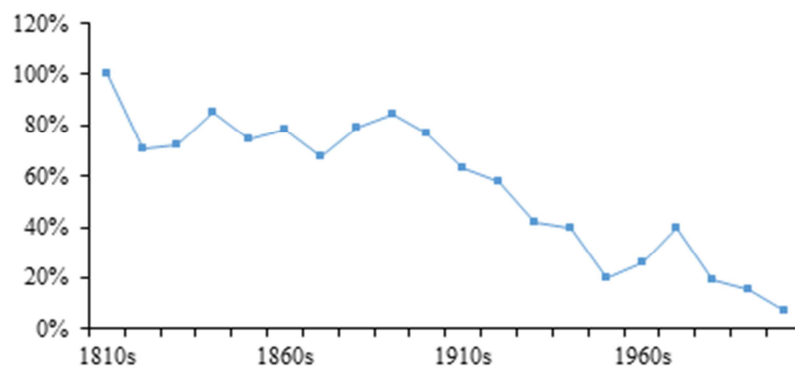
	<i>a body of</i> NOUN	<i>a lot of</i> NOUN
Spoken	15	3,528
Fiction	5	1,527
Magazine	16	934
Newspaper	2	1,174
Non-academic	53	680
Academic	77	270
MISC	29	1,332

An exhaustive concordance of the construction “*a body of* NOUN” was also performed in the BNC in each of its seven sections (genres). Frequencies of the target construction in

each section were acquired and we also searched in these sections for frequencies of the construction “*a lot of* NOUN” for comparison. Table 3 shows the results of these two concordances conducted.

4. Diachronic Distribution of Partitive *a body of*

In this section, we will investigate the grammaticalization of partitive *a body of* from the diachronic perspective and illustrate this construction with evidences derived from the COHA. See Figure 1:

**Figure 1.** Diachronic distribution of partitive *a body of* in the COHA.**Figure 2.** Proportion of nouns denoting human beings to all the nouns immediately following partitive *a body of* in the COHA.

In the past 20 decades, the frequency of this construction has been decreasing as shown in Figure 1, and the 1890s is a critical period. Before the 1890s, most nouns immediately following *a body of* denote human beings, such as *men, troops, Indians, cavalry, horsemen, soldiers, citizens, electors, persons, experts* and so on. We posit the reason is that the conventional meaning of *body* is inseparable from the concept of human being. Figure 2 shows the proportion of nouns denoting human beings among all the nouns that immediately follow the partitive construction *a body of* in the past 200 years.

Before the 1910s, nouns that denote human beings take up more than 60 percent of all the nouns that follow *a body of*. After that, the proportion has been decreasing to 20 to 60 percent until the 1950s. Later, the proportion fluctuates and finally goes lower than 10 percent in the 2000s.

As the proportion of nouns denoting human beings decreases, the diachronic frequency of the construction reduces dramatically, this construction, however, is still far away from being abandoned completely because in some cases, it cannot be replaced by other semantically and structurally similar expressions. For example, a countable noun following *a lot of* is more acceptably singular, while that following *a body of* may be either singular or plural. Take the countable noun *law* as an example, the data retrieved from the COHA are shown in Table 4:

Table 4. Occurrences of *law* and *laws* following *a lot of* and *a body of* in the COHA.

	Total	<i>law</i>	<i>laws</i>
<i>a lot of</i>	4	1	3
<i>a body of</i>	23	12	11

5. Genre Distribution of Partitive *a body of*

An important feature of academic writing is its high lexical density. Compared with the *a lot of* NOUN construction which contains only one notional word at the end, the *a body of* NOUN construction contains two notional words since part of

the conventional meaning of the word *body* still remains. Therefore, the construction has a greater contribution to the lexical density of text. We can hereby hypothesize that *a body of* tends to appear more in academic texts than in other types of texts with regard to genre distribution. Compare Figures 3 and 4:

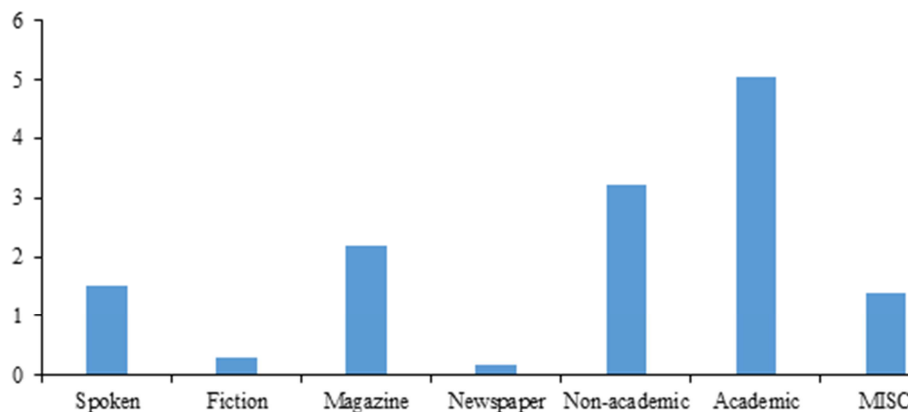


Figure 3. Genre distribution of partitive *a body of* in the BNC (normalized frequency).

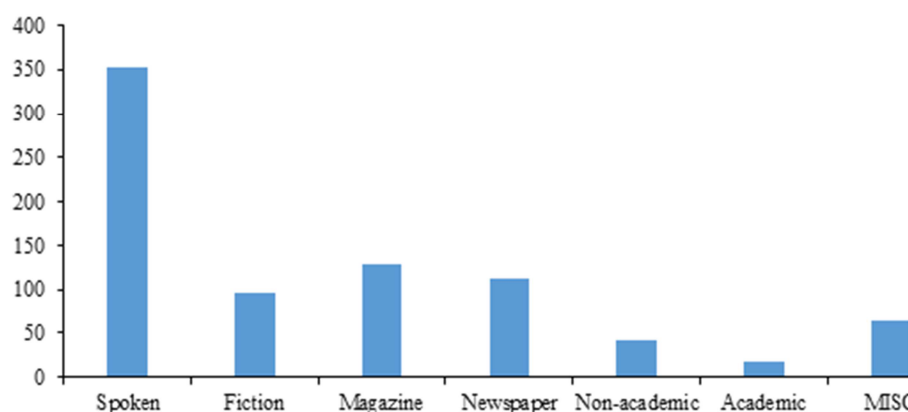


Figure 4. Style distribution of partitive *a lot of* in the BNC (normalized frequency).

While *a body of* tends to appear in formal academic texts, *a lot of* is most popularly used in spoken texts. This indicates that although structurally the same, the two partitive constructions are significantly different in genre distribution.

6. Discussion

Halliday [25] describes the structure of nominal groups in English from the logical perspective and the semantic perspective within the framework of systemic functional linguistics. The logical structure of a nominal group consists of a head noun and its modifiers, and the semantic structure, a thing noun and one or more characterizing elements preceding and a quality following the thing noun. The structure of the nominal group *a body of students* can be illustrated in Table 5:

Table 5. Functional structure of *a body of*.

	a	body	of	students
Logical	Premodifier	Head	Postmodifier	
Semantic	Numerative			Thing

An important indicator of full grammaticalization for constructions having the same structure and similar meaning with *a body of* is the move of both their semantic and logical focuses to the noun that follows. In the past 200 years, *a body of* has been under grammaticalization, but since only its semantic focus has moved to the slot after it, the construction has not yet fully grammaticalized as *a lot of*, whose semantic and logical focuses have fused into one slot, has done. See Table 6:

Table 6. Functional structure of *a lot of*.

	a lot of	students
Logical	Premodifier	Head
Semantic	Numerative	Thing

With the frequency decrease of nouns denoting human being, nouns which are less closely related to human beings account for a larger proportion of all the nouns that immediately follow *a body of*, indicating a process of grammaticalization of this construction. The lexical word

body in the construction further loses its conventional meaning, and the connection between the partitive and nouns that denote human beings has in turn become weaker. The construction thus goes beyond this particular type of nouns and attracts more nouns that have a less direct connection with the semantic meaning of the word *body*. In this way, the construction *a body of* as a whole has been gradually grammaticalized. Along this process of grammaticalization, the semantic focus of *a body of* has been moving from the lexical noun *body* to the noun that immediately follows. Its logical focus, however, remains with the word *body*, accounting for the reason why *a body of* has not yet been fully grammaticalized.

The dissociation of the semantic focus and logical focus of the construction can be further illustrated with a comparison between *a body of* and the fully grammaticalized *a lot of*. In the latter construction, both the semantic and logical focuses have moved from *lot* to the noun that follows. In *a body of*, however, the conventional meaning of the word *body* still has some logical contribution to the meaning of the construction as a whole. Although both of the constructions mean “a large number or amount”, the meaning of *a lot of* is more indefinite and that of *a body of* is more definite. Think of the two expressions *a body of students* and *a lot of students*.

Like *a lot of*, *a body of* has the potential to be fully grammaticalized, but as its use has sharply decreased, the grammaticalization process can be considered as having stopped at the current stage because a grammaticalization process can stop at any point of development [9].

The dissociation of the semantic focus from the logical focus can also explain the stylistic preference of the partitive *a body of*. This is because, in the *a body of* construction, the semantic focus and the logical focus falling in different slots results in three characteristics that are consistent with the features of academic texts. Compared with fully grammaticalized constructions, such as *a lot of*, which has become an indefinite determiner, *a body of* first evokes a more concrete and tangible sense of amount and second contributes better to the lexical density of text. Last, because the semantic focus and the logical focus of *a body of* do not lie in the same slot, it requires more effort to process its syntactic and semantic structure and thus explains why *a lot of* is more frequently used in spoken texts while *a body of* in academic texts.

7. Conclusion

The COHA-based research shows that the partitive *a body of* has been experiencing grammaticalization, but there is no sign of realizing full grammaticalization. The reason is that part of the conventional meaning of the word *body* is still attaching to the construction and influencing its meaning. This can also be explained by the dissociation of the semantic focus and logical focus of the construction. The BNC-based research shows that *a body of* has a strong tendency to occur in academic texts than in other types of texts, the reason being that the characteristics of the construction are more consistent

with the features of academic texts.

Acknowledgements

This research is supported by China National Social Science Fund (17BYY185).

References

- [1] Langacker, R. W. (1999). *Grammar and Conceptualization*. Berlin: Mouton de Gruyter.
- [2] Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia (Rivista di scienza)* 12 (6), 384–400.
- [3] Givón, T. (1971). Historical syntax and synchronic morphology. *Chicago Linguistic Society Proceedings* 7, 394–415.
- [4] Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.
- [5] Hopper, P. J., & Traugott, E. C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- [6] Heine, B., & Narrog, H. (2011). *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press.
- [7] Hancil, S., & König, E. (2014). *Grammaticalization – Theory and Data*. Amsterdam: John Benjamins.
- [8] Ramat, P. (2001). Grammaticalization. *International Encyclopedia of the Social & Behavioral Sciences*, 6353–6357.
- [9] Heine, B., & Kuteva, T. (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- [10] Mair, C. (2004). Corpus linguistics and grammaticalisation theory: Statistics, frequencies and beyond. In H. Lindquist and C. Mair (Eds), *Corpus Approaches to Grammaticalization in English* (pp. 121–150). Amsterdam: John Benjamins.
- [11] Wu, J., He, Q., & Feng, G. (2016). Rethinking the grammaticalization of future *be going to*: A corpus-based approach. *Journal of Quantitative Linguistics*, 23 (4), 317–341.
- [12] Tagliamonte, S. (2004). *Have to, Gotta, Must* Grammaticalisation, variation and specialization in English deontic modality. In H. Lindquist & C. Mair (Eds), *Corpus Approaches to Grammaticalization in English* (pp. 33–55). Amsterdam: John Benjamins.
- [13] Aijmer, K. (2004). The Semantic path from modality to aspect: *Be able to* in a cross-linguistic perspective. In H. Lindquist & C. Mair (Eds), *Corpus Approaches to Grammaticalization in English* (pp. 57–78). Amsterdam: John Benjamins.
- [14] Rissanen, M. (2004). Grammaticalisation from side to side: On the development of *beside(s)*. In H. Lindquist & C. Mair (Eds), *Corpus Approaches to Grammaticalization in English* (pp. 152–170). Amsterdam: John Benjamins.
- [15] Verveckken, K. (2012). Towards a constructional account of high and low frequency binominal quantifiers in Spanish. *Cognitive Linguistics* 23 (2). 421–478.

- [16] Brems, L. (2007). The grammaticalization of small size nouns. Reconsidering frequency and analogy. *Journal of English Linguistics* 35 (4). 293-324.
- [17] Traugott, E. C. (2007). The concepts of constructional mismatch and type-shifting from the perspective of grammaticalization. *Cognitive Linguistics* 18 (4). 523-557.
- [18] Brems, L. (2011). *Layering of size and type noun constructions in English*. Berlin: De Gruyter Mouton.
- [19] Herda, D. (2017). On Metonymically Motivated Delexicalization of Quantifying Nouns in English and Polish: A corpus investigation. *Studies in Polish Linguistics*, 12 (4), 199-219.
- [20] Verveckken, K. (2016). Binominal quantifiers in Spanish: syntagmatic and paradigmatic analogy in interaction. *Language Sciences*, 53 (Part B), 114-135.
- [21] De Clerck, B., & Brems, L. (2016). Size nouns matter: a closer look at mass(es) of and extended uses of SNs. *Language Sciences*, 53, 160–176.
- [22] Heine, Bernd. *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford: Oxford University Press, 1993.
- [23] Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- [24] Davies, M. (2010). *Corpus of Historical American English (COHA)*. <http://davies-linguistics.byu.edu/personal/>
- [25] Halliday, M. A. K. (1994). *An introduction to Functional Grammar (2nd edition)*. London: Edward Arnold.