

## A Corpus-based Study of the Grammaticalization Paths of the *a NOUN of* Construction

Huiqing Huang

Sun Yat-Sen University, China

Qingshun He \*

Sun Yat-Sen University, China

### Abstract

In this article a corpus-based study was conducted to investigate the grammaticalization paths of complex quantifiers with the *a NOUN of* construction in English. The nouns within the *a NOUN of* constructions in English were classified into two broad types, quantifiers and non-quantifiers, and complex quantifiers were recategorized into two types: definite complex quantifiers and indefinite complex quantifiers. According to the characteristics of nouns within complex quantifiers, two different grammaticalization paths were identified, one being from definite to indefinite, the other being from less indefinite to more indefinite. The research also indicated that the types of nouns following the complex quantifiers change with the process of grammaticalization.

**Keywords:** corpus-based, grammaticalization, complex quantifier, *a NOUN of* construction

### INTRODUCTION

How is the construction *a bunch of* in example (1a) different from the one in example (1b), and in what way do the *a bunch of* construction in example (1b) and the *a body of* construction in example (1c) relate to each other?

- (1) a. He said he had *a bunch of* flowers — and he really had. (BNC\_FIC)  
b. I hate to listen to guitarists rattle off rubbish about “the only thing that matters is the story of the song” because that’s just *a bunch of* crap! (BNC\_MAG)  
c. Once frozen *a body of* water remains frozen for the rest of the game. (BNC\_MISC)

The relationship between these three *a NOUN of* expressions may be explained by grammaticalization, a process defined as “the development from lexical to grammatical forms and from grammatical to even more grammatical forms” (Heine & Kuteva, 2002, p. 2).

Expressions like *a bunch of* are referred to as size noun constructions (SNCs) or binominal quantifier constructions (BQCs). SNCs are “structures that incorporate a size noun

expression, i.e. a nominal expression that describes size or shape, implying a measure” (Brem, 2011, p. 2). They can be schematized as (D) (M) N<sub>1</sub> of (D) (M) N<sub>2</sub>, (D) and (M) being an optional determiner and modifier respectively, and N<sub>1</sub> a size noun. Many linguists (e.g., Brems, 2011; Delbecque & Verveckken, 2014) have noticed the various synchronic uses of such constructions and consider them as the result of grammaticalization.

The scope of our target constructions in this research is much constrained because they are confined to the *a NOUN of* construction, consisting of an indefinite article, a size noun and a preposition *of*. We refer to this type of construction as complex quantifier. The concept of BQC is based on the definition of quantifying noun. Quantifying nouns 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” and a BQC is formed “when a prepositional phrase is added to specify the constituents” (Verveckken, 2012, p. 422).

The constructions discussed in this study differ from BQCs in that while abstract notions of quantity such as *kilo, number* and *part* are excluded from BQCs because they are not expected to undergo the semantic changes involved in the grammaticalization process of the remaining QNs, complex quantifiers with the *a NOUN of* construction do not single out the binominal constructions made up of such nouns. Our aim is not to distinguish constructions having undergone grammaticalization from those that are not likely or less likely to undergo grammaticalization; rather it is to examine by what paths complex quantifiers with the *a NOUN of* construction have been grammaticalized and in what way we can determine whether our target constructions have been grammaticalized or not.

The layout of this research is as follows. Section 2 reviews quantifying constructions and their grammaticalization. Section 3 introduces the two corpora that we will use in this study and the method of data collection. The findings and the discussion will be presented in Sections 4 and 5 respectively.

## LITERATURE REVIEW

Traditionally, quantifiers refer to determiners or pronouns indicating quantity, such as *all, some, both, much, many* and *few*, etc. In studies of constructions like *a bunch of, a cup of* and *a pile of*, however, the nouns within are also considered as quantifiers, or quantifying nouns and size nouns. Langacker (1991) explains how such lexical nouns denoting physical objects or the shape of certain objects can be considered as quantifiers and notices that the first noun appearing in expressions like *a bunch of carrots, a bucket of water* and *a lot of sharks* as head often has more than one possible interpretation.

These nouns still have an interpretation in which they designate a physical, spatially-continuous entity that either serves as the container for some portion of a mass (*bucket, cup, barrel, crate, jar, tub, vat, keg, box*) or else is constituted of some such portion (*bunch, pile, heap, loaf, sprig, head, stack, flock, herd*). Thus, *a bucket of water* may in fact be a physical bucket that is filled with water, and *a bunch of carrots* may in fact be a lump-like object formed by tying a number of carrots together

into a bunch... Most such nouns have taken on a different sense in which size becomes the most salient specification. (Langacker, 1991, p. 88)

Taking *a bucket of water* as an example, Langacker (1991) explains that the expression is used to refer to a bucket actually filled with water as well as to the water in a bathtub whose amount is equal to the volume of a bucket. In this way, a notional noun incorporated in a binomial quantifying construction can be regarded as a quantifier itself. Langacker (1991) also suggests that the second interpretation comes from the first one:

The notion of a discrete physical object has faded, leaving behind the conception of a schematically characterized mass (the mass that, in the original sense, either fills or constitutes the object) whose projection on the scale of magnitude then provides its primary semantic content. (p. 88-89)

Such a process whereby a new explanation or function of a noun derives from an old one is actually a process of grammaticalization. The synchronically coexisting uses of SNCs are “the result of (diachronic) processes of grammaticalization” (Brems, 2011, p. 79). The factors at work in the grammaticalization of SNCs include reanalysis and analogy, semantic change, layering and persistence, (inter)subjectification, decategorialization, paradigmaticization, syntactic extension, coalescence and phonetic erosion. Reanalysis is “the typically covert reinterpretation of a given construction, both on the semantic and formal level” (Brems, 2011, p. 98). Analogy refers to “the similarity a speaker perceives between two patterns” (Delbecque & Verveckken, 2014, p. 640). Brems (2011) agrees with Hopper & Traugott (2003) and considers reanalysis as the primary mechanism for grammaticalization and analogy the secondary. Based on their study of the grammaticalization of Spanish binominal quantifiers, Delbecque and Verveckken (2014) claim that analogy is more essential as it motivates (morpho)syntactic reanalysis.

Other cross-linguistic studies (e.g., Herda, 2017) have explored the mechanisms behind the grammaticalization of complex quantifiers or SNCs. Agreeing on the claim that the synchronic uses of SNCs are products of grammaticalization, this article looks on this matter from a different perspective. It captures an important feature of the target constructions, whether they are definite or indefinite, and proposes grammaticalization paths. It also attempts to show how we can conduct corpus-based research to determine whether a complex quantifier is grammaticalized and, if it is, to track the timeline of the process.

## METHOD

### Corpora

Corpora are recognized as “collections of texts (or parts of text) that are stored and accessed electronically” (Huston, 2002, p. 2). They are often designed different in text types, text size and context for different research purposes and are rather instrumental in hypotheses testing and insight discovery for language studies. We will use the Corpus of Historical American English (COHA) and the British National Corpus (BNC) in this research to illustrate the grammaticalization process of our example constructions and to exemplify our arguments with example sentences respectively. The COHA is designed

for the study of historical English, containing 406 million words with a time span of 20 decades, each of which is a well-balanced sub-corpus by themselves. These characteristics make the COHA a corpus that “allows researchers to examine a wide range of changes in English with much more accuracy and detail than any other available corpus” (Davies, 2010). The BNC is a corpus developed by Oxford University Press in the 1980s to early 1990s. Compared to the COHA, it has a relatively small text size, containing 100 million words. However, it has a wide range of well-balanced text genres and is convenient to retrieve data for genre analysis.

### Data Collection

In order to distinguish our target constructions from non-complex quantifiers with the same structure and to observe the characteristics of the nouns that can be included in a complex quantifier, we conducted an exhaustive concordance of the *a NOUN of* construction in the BNC. Some typical concordance lines retrieved from the BNC are shown in Table 1:

**Table 1.** “*a NOUN of*” constructions retrieved from the BNC (per million words)

Expression	Frequency	Expression	Frequency
<i>a number of</i>	149.66	<i>a range of</i>	33.28
<i>a lot of</i>	144.40	<i>a piece of</i>	25.20
<i>a couple of</i>	69.02	<i>a kind of</i>	24.89
<i>a series of</i>	58.88	<i>a sense of</i>	24.80
<i>a result of</i>	55.49	<i>a period of</i>	23.78
<i>a bit of</i>	49.49	<i>a pair of</i>	23.05
<i>a variety of</i>	42.56	<i>a set of</i>	23.04
<i>a matter of</i>	41.60	<i>a sort of</i>	21.38
<i>a member of</i>	38.74	<i>a way of</i>	19.51
<i>a group of</i>	34.86	<i>a total of</i>	19.50

To illustrate the grammaticalization paths, exhaustive concordances of complex quantifiers *a bunch of* and *a body of* were also performed in the COHA. Some typical nouns that follow the two complex quantifiers are shown in Tables 2 and 3:

**Table 2.** Diachronic distributions of typical nouns following complex quantifier *a bunch of* in the COHA (per million words)

	<i>key</i>	<i>flower</i>	<i>grape</i>	<i>people</i>	<i>man</i>	<i>cattle</i>	<i>crap</i>	<i>time</i>	<i>nonsense</i>
1810s	0	0	0	0	0	0	0	0	0
1820s	0.14	0.29	0	0	0	0	0	0	0
1830s	0.36	0.36	0.36	0	0	0	0	0	0
1840s	0.31	0.37	0.06	0	0	0	0	0	0
1850s	0.49	0.49	0.18	0	0	0	0	0	0
1860s	0.23	0.29	0.23	0	0	0	0	0	0
1870s	0.16	0.43	0.22	0	0	0	0	0	0
1880s	0.74	0.30	0.15	0	0	0.05	0	0	0
1890s	0.19	0.29	0.19	0.04	0	0.15	0	0	0
1900s	0.32	0.23	0.09	0.08	0.05	0.09	0	0	0
1910s	0.48	0.22	0.04	0.04	0.09	0.13	0	0	0
1920s	0.31	0.20	0.04	0.04	0.20	0.08	0	0	0
1930s	0.20	0.16	0.12	0.21	0.16	0.08	0.04	0	0
1940s	0.33	0.12	0.21	0.50	0.12	0.04	0.04	0	0
1950s	0.16	0.12	0.16	0.12	0.04	0	0.04	0	0.04
1960s	0.13	0.04	0.25	0.43	0.17	0	0	0	0.08
1970s	0.29	0.04	0.13	0.27	0.04	0.04	0.42	0	0
1980s	0.08	0.16	0.20	0	0.08	0	0	0.04	0.04
1990s	0	0.32	0.25	0	0.21	0	0.04	0.18	0.14
2000s	0.03	0.03	0.34	0	0.07	0	0	0.20	0.03

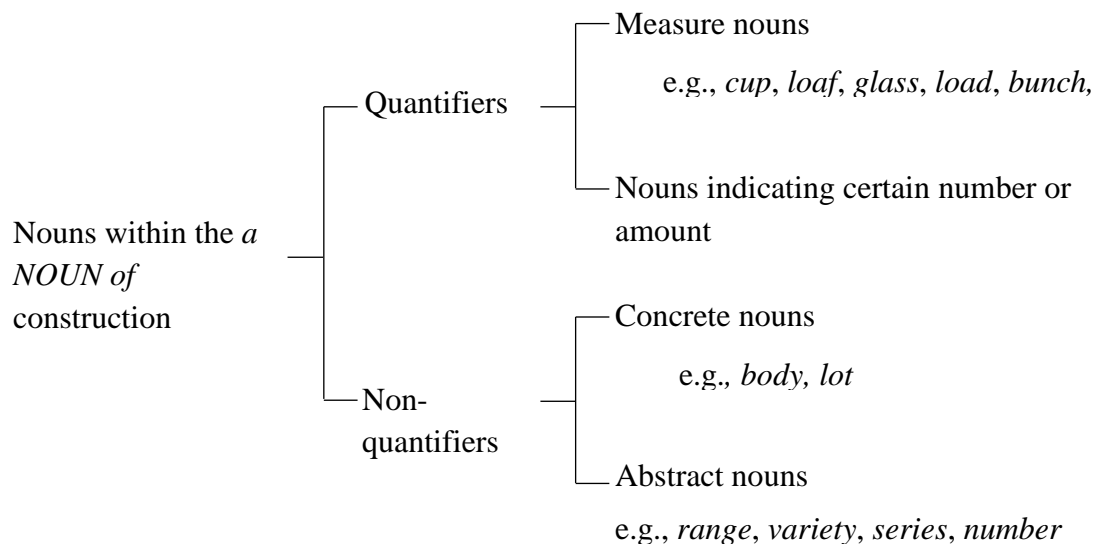
**Table 3.** Diachronic distributions of typical nouns following complex quantifier *a body of* in the COHA (per million words)

	<i>man</i>	<i>water</i>	<i>troop</i>	<i>knowledge</i>	<i>Indian</i>	<i>law</i>	<i>cavalry</i>
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

## RESEARCH FINDINGS

### Classification of Nouns within the *a NOUN of* Construction

Corpus data show that all the nouns within the *a NOUN of* constructions in English can be classified into two broad types, quantifiers and non-quantifiers, each of which can be further divided into two sub-types respectively as shown in Figure 1:



**Figure 1.** Classification of nouns appearing in the *a NOUN of* construction

These types of complex quantifiers can be recategorized into two types: definite complex quantifiers and indefinite complex quantifiers. The noun within a definite *a NOUN of* construction must be a quantifier, i.e. either a measure noun or a noun that indicates a certain number. For example:

- (2) a. On the bedside table was *a glass of water* and a bible. (BNC\_FIC)  
 b. One morning a carter and his son were delivering *a load of firewood*. (BNC\_MISC)  
 c. I started skating with *a couple of guys* called Sam and Luke. (BNC\_MAG)

The noun within an indefinite *a NOUN of* construction can be a quantifier or a non-quantifier. A quantifier, however, does not function the same as it does when the construction as a whole is a definite modifier. All the indefinite complex quantifiers with the *a NOUN of* construction have experienced a grammaticalization process.

### Types of Nouns Following the *a bunch of* Construction

It can be seen in Table 2 that nouns following the complex quantifier *a bunch of* can be generally divided into three types according to the relationship between their meanings and the meaning of the word *bunch*. The first type are nouns denoting flowers, vegetables and small objects fastened together, whose shape can be properly described by the measure noun *bunch*. The second type are nouns denoting human beings, animals and buildings, which have the characteristic of gathering together and thus have a much weaker but still observable relation with the meaning of the measure noun *bunch*. The

last type are nouns having no obvious relation to the word *bunch*, which are abstract nouns such as *crap* and *nonsense*. The second type of nouns generally came to collocate with the construction *a bunch of* a few decades later than the first type while the third type appeared even later as can be seen in Table 2.

The proportions of the three types of nouns following the construction *a bunch of* remain steady during the 200 years from the 1810s to the 2000s as is shown in Figure 2. Those nouns that are directly related to the meaning of the measure noun *bunch* takes up a proportion from 40 to 60 percent, nouns that are indirectly related to the measure noun between 20 to 40 percent and abstract nouns that have no obvious relation to the measure noun below 10 percent.

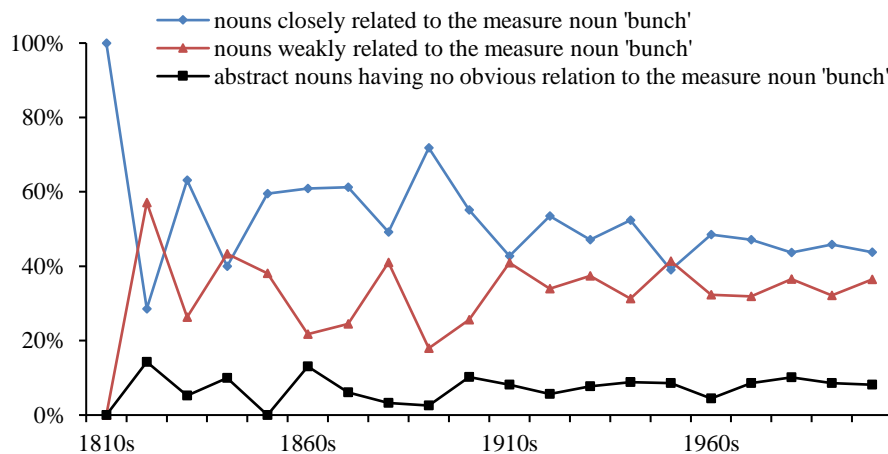


Figure 2. Proportions of the three types of nouns following *a bunch of* in the COHA

### Diachronic Distribution of Complex Quantifier *a body of*

It can be seen from Table 3 that the proportions of the types of the nouns following the complex quantifier *a body of* change dramatically. Before the 1910s, most of the nouns following *a body of* denote human beings, such as *men, troops, Indians, horsemen, soldiers* and *citizens, etc.*, all of which are closely related to the meaning of the word *body* and take up over 60 percent of all the nouns following the construction. During the 1920s to the 1980s, however, the proportion goes down and fluctuates within the range from 20 to 60 percent. After that, the proportion decreases even further and finally goes to lower than 10 percent in the 2000s.

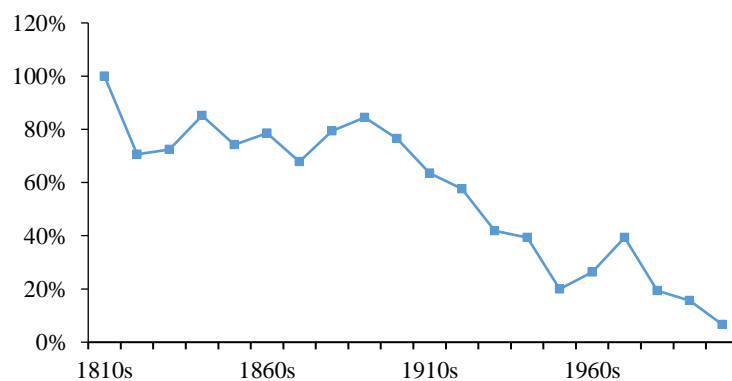


Figure 3. Proportion of nouns denoting human beings to all the nouns immediately following partitive *a body of* in the COHA

The decrease of the proportion of the nouns denoting human being is compensated for the increase of the proportion of nouns having no direct relationship with the meaning of the noun *body*, many of which are abstract nouns such as *knowledge, work, literature, law(s), evidence, information, facts, rules* and so on.

## DISCUSSION

This section is dedicated to the discussion of the two grammaticalization paths of the complex quantifiers *a NOUN of* respectively according to the types of the noun within.

### The First Grammaticalization Path: From Definite to Indefinite

Indefinite complex quantifiers containing a quantifier noun developed from definite complex quantifiers containing the same quantifier noun. For example, the *a bunch of* construction in example (3b) is derived from that in (3a), and the *a couple of* construction in example (3c), from that in (2c).

- (3) a. Marek snatched up *a bunch of* grapes from the bedside table.  
(BNC\_FIC)
- b. Graham started chatting away to *a bunch of* people he took to be the support group. (BNC\_MAG)
- c. If you remember *a couple of* years ago. Er three or four years ago.  
(BNC\_SPOK)

Along this process of grammaticalization, the quantifier noun within the construction has experienced a loss of meaning and the relationship between the three elements in the *a NOUN of* construction has become more intimate. Such a process can also be seen as the change of the function of the quantifier within the construction from head to modifier (Huang & He, 2019). When a SNC is grammaticalized, the original structure of the [SN] + [of + N2] construction is “rebracketed” as [SN + of] + [N2], “which entails an inversion of the hierarchic dependency relations, i.e. the head/ modifier-relations” (Brems, 2011, p. 98). Such a generalization applies to this first grammaticalization path proposed in terms of definiteness and indefiniteness.

When the noun incorporated in a complex quantifier is a measure noun, a useful way to determine whether the construction is definite (ungrammaticalized) or indefinite (grammaticalized) is to see whether the construction has the plural form or not. A definite complex quantifier has its plural form. For example:

- (4) *Two bunches of* roses from the dead girl’s parents were placed at the scene after the jury left. (BNC\_NEWS)

The grammaticalization of the *a NOUN of* construction can also be tested by the type of noun that follows. The noun following an ungrammaticalized complex quantifier has a strong relationship with the noun in the construction. Take definite complex quantifiers with measure noun as examples. The measure noun within a definite complex quantifier could be the shape or the container of the noun following the construction. For example:

- (5) a. A vapour, *a drop of* water suffices to kill him. (BNC\_MISC)
- b. Out of his pocket he pulled *a bunch of* keys. (BNC\_FIC)



c. Anyway, so they gave us *a bottle of wine*. (BNC\_SPOK)

d. Take him *a load of coal*, and he took him. (BNC\_SPOK)

Along this grammaticalization process, the noun within a complex quantifier gradually loses its conventional meaning and thus the meaning of the quantifier changes as a whole. The construction, consequently, is used in new contexts and attracts nouns having a less direct relationship with the noun within it. The more grammaticalized a complex quantifier is, the more abstract the noun that follows it is. Therefore, the nouns that follow a complex quantifier determines whether or not the construction is grammaticalized.

The *a bunch of* construction, for example, can be either definite or indefinite. The nouns following this quantifier can be classified into three types. When it functions as head followed by a prepositional phrase, the quantifier noun *bunch* describes the physical shape of the object that the noun following the construction denotes and the noun would be from the first type we have categorized. As the word *bunch* has experienced a loss of conventional meaning, the construction as a whole was grammaticalized and no longer was limited to the first type of nouns. Rather, it began to attract nouns denoting human beings, animals, buildings and even more abstract nouns. The observable gaps between the times when the three different types of nouns started to collocate with *a bunch of* respectively provide evidence to the idea that nouns not closely related to the conventional meaning of the measure noun appeared after the construction only when it was grammaticalized and became indefinite.

On the other hand, we can see that when new types of nouns arose, the existing ones do not just disappear. Different usages of a construction, definite or indefinite, ungrammaticalized or grammaticalized, often co-exist. The proportions of different types of nouns following the complex quantifier, however, do change over time. We then suggest that with the increase of the degree of grammaticalization of a construction, the proportion of the nouns that are more abstract and less directly related to the noun within the construction will increase and finally exceed that of nouns having a more direct relation to the noun within.

However, since grammaticalization is a long-term process that may span centuries and because the COHA only covers a time span of 200 years, data collected from the corpus can only reflect a very short period of the grammaticalization of our target constructions. From the data we collected for the construction *a bunch of*, we have not found noteworthy changes of the proportions of the three types of nouns. Nevertheless, in accordance with our hypothesis, the frequencies of the second and third types of nouns increase with the grammaticalization of the complex quantifier *a bunch of*.

### **The Second Grammaticalization Path: From Less Indefinite to More Indefinite**

A complex quantifier with a non-quantifier noun within is always indefinite; it does not have a definite equivalent. This is because such a complex quantifier follows a different grammaticalization path. Although the origin of such complex quantifiers is still subject to further research and confirmation, we can be sure that since such constructions came into being and began to function as quantifiers, they all have undergone a

grammaticalization process, which is very possibly to continue. In this sense, within the *a NOUN of* construction the grammaticalization of complex quantifiers such as *a body of* and *a lot of* can be characterized as something from less indefinite to more indefinite.

From the nouns following a complex quantifier with a non-quantifier noun, we can find indications of the grammaticalization of quantifiers. Take the partitive *a body of* as an example. During the 200 years, *a body of* has experienced the process of grammaticalization, noticeably marked by the decrease of the proportion of nouns denoting human beings among all the nouns following the quantifier and the consequent increase of other nouns that are much less closely related to the conventional meaning of the noun *body*. In other word, as the result of grammaticalization, the complex quantifier *a body of* is able to collocate with a wider variety of nouns.

## CONCLUSION

In complex quantifiers with the *a NOUN of* construction, we can identify two grammaticalization paths. One is from definite to indefinite, meaning that a definite complex quantifier carrying a quantifier noun develops into an indefinite complex quantifier having the same form. The other is from less indefinite complex quantifier to more indefinite complex quantifier, meaning that complex quantifiers with a non-quantifier noun are usually grammaticalized constructions and they are being further grammaticalized. An important way to identify which complex quantifiers have been grammaticalized or are undergoing a grammaticalization process is to observe the type of noun that follows.

## ACKNOWLEDGEMENTS

This research is supported by China National Social Science Fund (17BY185).

## REFERENCES

- Brems, L. (2011). *Layering of Size and Type Noun Constructions in English*. Berlin: De Gruyter Mouton.
- Davies, M. (2010). *Corpus of Historical American English (COHA)*. <http://davies-linguistics.byu.edu/personal/>
- Delbecque, N., & Verveckken, K. (2014). Conceptually-driven analogy in the grammaticalization of Spanish binominal quantifiers. *Linguistics*, 52(3), 637–684.
- Heine, B. & Kuteva, T (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- 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.
- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.
- Huang, H. & He, Q. (2019). A corpus-based study of the grammaticalization of partitive *a body of*. *Language, Literature and Culture*, 2(2), 1-7.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Langacker, R. W. (1991). *Foundations of Cognitive Grammar (Volume II): Descriptive application*. Stanford: Stanford University Press.
- Verveckken, K. (2012). Towards a constructional account of high and low frequency binominal quantifiers in Spanish. *Cognitive Linguistics*, 2, 421-478.