

English Language and Linguistics, 25.1: 1–33. © The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

doi:[10.1017/S136067431900039X](https://doi.org/10.1017/S136067431900039X)

Constructional competition and network reconfiguration: investigating *sum(e)* in Old, Middle and Early Modern English

LOTTE SOMMERER and KLAUS HOFMANN

University of Vienna

(Received 10 May 2019; revised 3 October 2019)

This article investigates some functions of the determinative *sum(e)* in Old, Middle and Early Modern English. It traces, quantifies and models the diachronic development of *sum(e)* as a pre-head element from a usage-based, cognitive Construction Grammar perspective by postulating several semi-specified but also abstract constructional OE and ME NP-schemas and sketching the observable (changing) network (re)configurations. By analyzing texts from the *Penn–Helsinki Parsed Corpus of Middle English* (PPCME) and the *Penn–Helsinki Parsed Corpus of Early Modern English* (PPCEME), the article especially focuses on the demise of the so-called ‘individualizing’ usage with singular nouns and traces the incipient stages of *sum(e)* as an indefinite near-article with plural and mass nouns. R was used to calculate correlation coefficients and measures of statistical significance in univariate analyses, and for multivariate regression models to address questions involving more than one predictor variable. It is shown that the usage of *sum(e)* with singular nouns became marginalized because of constructional competition with the numeral *ān*. In Old English, the two forms were both occasionally used to mark indefiniteness before singular nouns, but ultimately *ān* became the default marker of indefiniteness ousting *sum(e)*. We also show that the usage of *sum(e)* as an indefiniteness marker for plural nouns increased drastically from the later ME period onwards, particularly in informal text genres. Moreover, from the earliest periods onwards, there is a strong preference for this function to occur with complex NPs with pre- and post-head modification, which seem to have acted as bridging contexts.

Keywords: indefinite article, nominal determination, Diachronic Construction Grammar, grammaticalization, constructionalization, constructional competition

1 Introduction

This article investigates some functions of the determinative *sum(e)* in Old, Middle and Early Modern English. It traces, quantifies and models the diachronic development of *sum(e)* as a prehead element from a constructional perspective. Our contribution especially focuses on the demise of the so-called ‘individualizing’ usage with singular nouns (1) and traces the incipient stages of *sum(e)* as an indefinite near-article with plural and mass nouns (2).

(1) & **sum wif** hatte Uenus, seo wæs Ioues dohtor ...

‘and there was a woman called Venus, who was Jupiter’s daughter’ (*Dictionary of Old English Web Corpus* [DOEC] 2009, WHom 12 [0026 (77)])

- (2) ... þat he schuld sende **summe prestes** to þis lond ...
 ‘... that he should send some priests to this land ...’ (PPCME, CMCAPCHR, 54.655)

We will argue that the use of *sum(e)* with singular nouns became marginalized because of constructional competition with the numeral *ān*. In Old English (OE), the two forms were both occasionally used to mark indefiniteness before singular nouns, but ultimately *ān* became the default marker of indefiniteness ousting *sum(e)*. We will also show that Middle English (ME) *sum(e)* started to grammaticalize into an indefinite near-article for plural and mass nouns in late Middle English, following the previously established indefinite singular article *a/an*.

Originally lacking an indefinite article category, English developed one in the wake of the complete systemic reorganization of definiteness and indefiniteness marking during late Old English and Middle English (Christophersen 1939; Rissanen 1967; Traugott 1992; Sommerer 2018). Overt (in)definiteness marking became obligatory for referential NPs during the late OE period, which corresponds to the emergence of an abstract NP-construction with a determination slot requiring to be filled. In the singular, OE demonstrative *se* (‘that’) and the numeral *ān* (‘one’) were recruited as default markers of (in)definiteness, grammaticalizing into today’s *the* and *a/an*. In the plural, *the* also established itself to mark definiteness. Indefinite *ān*, on the other hand, remained restricted to singular nouns and did not branch out into the plural. It is our view that *sum(e)* started to grammaticalize before plural and mass nouns to fill the gaps left by *ān*, while at the same time relinquishing its OE individualizing function with singular count nouns to *ān*.

In terms of its theoretical approach, this article subscribes to a usage-based, cognitive, Construction Grammar model (e.g. Goldberg 2006, 2019; Hilpert 2014; Diessel 2015, 2019; Ellis, Römer & O’Donnell 2016), which postulates that linguistic knowledge is stored in the form of constructions (i.e. form–function pairings) and in which ‘grammaticalization’ is reconceptualized as ‘grammatical constructionalization’ (e.g. Traugott & Trousdale 2013; Trousdale 2014; Sommerer 2018). In such a model, language change is conceived as the reconfiguration of the ‘construction’ (i.e. network reconfiguration via node emergence and node-external reconfiguration of links between constructions). It is primarily cognitive factors (frequency, processing efficiency, analogical thinking, ability to schematize and abstract ...) and the influence of related constructions which are responsible for the constructionalization of [*sum(e)*]_{ART}. To our knowledge, this article is the first to model the grammaticalization of *sum(e)* within this framework.

Empirically, this article relies on the quantitative and qualitative analysis of texts from the *Penn–Helsinki Parsed Corpus of Middle English* (PPCME 2000) and the *Penn–Helsinki Parsed Corpus of Early Modern English* (PPCEME 2004). The open statistical software environment *R* (R Core Team 2018) was used to extract the data and to calculate correlation coefficients and measures of statistical significance in univariate analyses. To address questions involving more than one predictor variable, we constructed a multivariate logistic regression model (Baayen 2008; Gries 2009; Levshina 2015). One main finding from these analyses is that the usage of *sum(e)* as

an indefiniteness marker for plural nouns increased drastically from the later ME period onwards, particularly in informal text genres. Moreover, from the earliest periods onwards, there is a strong preference for this function to occur with complex NPs with pre- and post-head modification, which seem to have acted as bridging contexts.

The remainder of this article is structured as follows: [section 2](#) provides an overview of the various functions of *sum(e)*. First, we discuss its functions in Present-day English to introduce essential terminology used throughout the article. After that, we consider the etymological roots of *sum(e)* and its Old and Middle English uses. In [section 3](#), we present the results from the empirical corpus study. The findings are then interpreted through a constructional lens in [section 4](#). [Section 5](#) concludes the article.

2 Changing functions of *some*

The functional inventory of *some* has not remained constant in the history of English. Various OE uses (Mustanoja 1960; Mitchell 1985) have become rare or obsolete, while new ones have emerged over time. Before we present the historical functions of *sum(e)*, we briefly turn to functions of *some* in Present-day English.

2.1 Versatility of *some* in Present-day English

In the comprehensive grammars of English, *some* has been classified as an independent indefinite pronoun heading its own NP (3) or as determinative which occurs in the pre-head of the NP (4) (Quirk *et al.* 1985: 255; Huddleston & Pullum 2002: 92). *Some* is also found in the partitive [*some of* DET+CN] construction with the preposition *of* (5) and in compound forms like *somebody*, *someone* (Huddleston & Pullum 2002: 100).

- (3) If there is pizza left, I'd like to get **some**.
- (4) **Some people** like pineapple on their pizza.
- (5) **Some of my friends** have season tickets for the opera.

In its function as a dependent determiner, *some* is an 'existential determinative' (Payne & Huddleston 2002: 380), which is located in the determination zone before any modifiers.¹ *Some* often functions as a quantifier. In English, quantification can be expressed by a range of lexical items such as *always*, *rarely*, *numerous*. Mostly, however, it is expressed by a small set of quantifiers, e.g. *some*, *all*, *several*, *none*. One can distinguish between existential quantification and universal quantification. Universal quantification, which is typically expressed by *all* (e.g. *All students were tired*), refers to a complete set of entities. Existential quantification, which indicates a quantity or number greater than zero, is most often expressed by *some*. According to Jespersen (1949: 49), the quantifier *some* expresses an unspecified quantity, amount or number of persons and

¹ Based on Halliday's Systemic Functional Grammar, Downing & Locke (2006: 404) classify *some* as an 'indefinite inexact determinative'. Halliday & Matthiessen (2014: 365) also describe *some* as a determiner and classify it as a 'non-specific deictic'.

things. Biber *et al.* (1999: 275–8) are more specific, stating that *some* usually specifies a moderate quantity. It generally refers to a number not less than two. In examples (6) and (7), *some* has a quantifier function.

- (6) **Some candidates** misunderstood the question.
 (7) **Some people** don't know how to love.

The examples indicate the existence of a quantity or number of candidates and people. This number has a certain property or behaves in a certain way (e.g. misunderstanding the question). We are concerned with a subset of people belonging to an implied larger set, which is why in all cases one could easily substitute *some* by *not all*, in the sense of NOT ALL people misunderstood the question and NOT ALL people don't know how to love. This function of *some* has been called '(basic) proportional' (Payne & Huddleston 2002: 381) 'partitive' (Israel 2000: 173) and 'selective' (Halliday & Matthiessen 2014: 366–9). Sahlin calls this function 'an indefinite assertive limiting quantifier' (1979: 14–16).

Some is not always used to express moderate quantity but can also be used to express considerable quantity in the sense of *many* (8)–(9):

- (8) It was **some years** before she saw him again.
 (9) We discussed the problem at **some length**.

In addition, *some* can also have other functions that need to be distinguished from those above. Especially in combination with singular nouns, it can express non-specificity, vagueness or speaker attitude (10)–(13) (Biber *et al.* 1999: 278; Israel 2000: 170; Payne & Huddleston 2002: 380f.).

- (10) **Some day** I will win the lottery.
 (11) **Some idiot** must have left the fridge open!
 (12) When I arrived, **some student** was waiting outside the door.
 (13) **Some 18 percent of managing directors** secured pay increases of over 20 percent.

The usage in (11) has been called 'exclamatory use'² (Payne & Huddleston 2002: 381) due to the prosodic emphasis that *some* receives in such cases. In (13) *some* is used as an approximating adverb before numerals.

Finally, and most importantly for our purposes, *some* can take yet another function that is rather different. It may be used in a non-proportional, non-selective, non-partitive way (14)–(17).

- (14) I need to buy **some apples**.
 (15) We need **some milk**.
 (16) I see **some cars** driving along the road.
 (17) There were **some children** in the park.

² The exclamative function of the determiner *some* falls outside the distinction partitive versus near-article and will not be discussed in this article. For a detailed description see Israel (2000).

The above examples are different from the partitive/quantifier usage of *some* described earlier. Here, *some* lacks the proportional aspect: there is no ‘not all’ implicature. Several scholars have classified this kind of *some* as an indefinite ‘near-article’, selecting plural and mass heads (e.g. Sahlin 1979; Chesterman 1993; Jacobsson 2002). For example, Sahin calls this function ‘indefinite assertive article’ (Sahin 1979: 14; see also Israel 2000; Jacobsson 2002). In this function, *some* is usually reduced to /s(ə)m/. The reduced phonology reflects the general pattern of quantitative and qualitative vowel reduction (and eventual elision) in function words due to lack of stress (Selkirk 1996; Bybee 2001: 74; Kingston 2007: 419–22). Its non-specific, individualizing interpretation shows significant resemblance to the indefinite *a/an* (e.g. in clauses introduced by *there is/are*). In this function, *some* often serves to introduce a new entity in discourse and designates the existence of a given nominal type (Israel 2000: 172).³

Although the above classification of functions might not do justice to every semantic nuance associated with various uses of *some*, it suffices as a backdrop for our diachronic investigation. We agree with previous accounts that in cases like (14)–(17) *some* functions as an indefinite near-article which selects plural and non-count heads. In this article we are particularly interested in the diachronic development of this function. It should be noted, however, that an analysis which grants *some* a status similar to the canonical articles *the* and *a/an* can be criticized. One of the reasons for skepticism is that *some* is much more restricted in its distribution than *the* and *a/an*. In Present-day English, indefinite plural NPs are not necessarily marked by any overt determinative (18)–(19). Frequently, the NP consists of only the nominal occurring as a ‘bare’ noun.

(18) I see **cars** driving along the road.

(19) There were **children** in the park.

Additionally, *some* cannot be used in ascriptive and predicative complements, as in (20c) and (21b).

(20) (a) As a doctor, she should know better.

(b) As doctors, they should know better.

(c) *As some doctors, they should know better.

(21) (a) This liquid is acid.

(b) *This liquid is some acid.

Also, *some* cannot be used in generic constructions without changing the meaning. Thus, insertion of *some* automatically triggers a quantifier reading (22).

(22) (a) Lions are ferocious beasts.

(b) Some lions are ferocious beasts.

However, we contend that the fact that the near-article use of *some* is barred in the cited examples is not a sufficient argument against its potential article status. We argue that an

³ We adopt the terminology from Israel (2000) in distinguishing partitive usage and non-partitive, article-like usage.

Table 1. *(In)definiteness marking in Present-day English*

	DEFINITE		INDEFINITE	
	Count	Non-count	Count	Non-count
SINGULAR	the book	the ink	a book	–/(some) ink
PLURAL	the books		–/(some) books	

element deserves to be classified as an article based on its behavior in referential NPs, not based on a complete distributional congruency with canonical representatives of the article class. Arguably, a similarly uneven situation obtained at historical stages of English, when *the* had already been firmly established as an article in definite contexts while the indefinite determiner *a/an* was still trailing behind (Sommerer 2018). One may speculate that although *some* may not be compatible with non-referential or generic uses yet (as this represents the last step in the grammaticalization cline), it may be assigned these functions in the future.⁴ All the cases (20)–(22) above are non-referential or generic.⁵

We admit, however, that it also true that *some* does not have to be used obligatorily in the same way as the definite article *the* or the indefinite article *a/an*. As can be seen in table 1, the marking of definiteness is obligatory in Present-day English with count and mass noun heads. In contrast, indefiniteness is not obligatorily marked with plural and mass noun heads. It only has to be marked with count heads in the singular. For this reason, we argue that the unstressed, non-partitive PDE *some* functions as an indefinite near-article for referential plural and non-count nouns. It does not deserve full article status (yet), as we define a full article as an element which ‘is a syntactically fixed default slotfiller used to exclusively and obligatorily mark (in)definiteness’ (Sommerer 2018: 72).

2.2 *Sum(e) in Old and Middle English*

Etymologically, PDE *some* goes back to PGmc. **sumas* from the PIE root **sem-* ‘one, as one’ (cf. Skt. *samah* ‘even, level, similar, identical’) (OED).⁶ OE and ME *sum(e)* could

⁴ It is worth noting that in French the so-called *article partitif* serves a function very similar to the indefinite near-article use of *some* argued for here. As the name suggests, the *article partitif* is diachronically related to a partitive quantifier (Carlier 2007), as is English *some*. Unlike in English, however, the overt (in)definiteness marking of plural and mass nouns is next to obligatory in French, to the extent that even generic NPs as in (22) would be marked by an overt article (albeit not necessarily with the same definiteness values as in English), e.g. *Les lions sont des bêtes féroces*.

⁵ We consider generic NPs to be referential.

⁶ There is considerable dialectal and diachronic variation regarding the orthographic representation of *sum(e)* in English texts, which may or may not be indicative of its phonological form in the respective dialects at the various points in time. The attested forms include *sum*, *zum*, *summ*, *sume*, *summe*, *soumme*, *soume*, *soom*, *som*, *zom*, *somp*, *somme*, *zome*. By Early Modern English, the form *some* had largely established itself (OED) for both singular and plural usage. In this article we use the form *sum(e)* as the formal shape of the diachronically

express a number of different meanings and had several syntactic functions (Wülfing 1901 [1894]; Heltveit 1977; Mitchell 1985: 152). Like in Present-day English, it could be used either as an independent pronoun or as a pre-head dependent. In the latter function, OE *sum(e)* was consistently declined as strong⁷ and used both in the singular and plural. Several of the OE functions have survived but some others have become obsolete.

2.2.1 Independent pronoun usage in Old and Middle English

For completeness, we briefly mention the use of *sum(e)* as an independent pronoun. OE and ME *sum(e)* is often used in contrasting constructions [*some...*, *some/other ...*] as in (23).

(23) **Sume** hi beoton, **sume** hi ofslogon.

‘Some they beat, some they killed.’ (DOEC 2009, Mk (WSCp) [0461 (12.5)])

Pronominal *sum(e)* clearly has a partitive function. It denotes a certain part of something, a portion: an indefinite or unspecified but not large number of people (24).

(24) & þær ofslogon manige Walas & **sume** on fleame bedrifon, on þone wudu

‘and (they) killed many Welsh there and drove some to flight into the wood’ (DOEC, ChronE, [0094 (477.1)])

In example (25), OE *sum(e)* is used in combination with a partitive genitive plural. *Sum(e)* in this construction can precede or follow the genitive construction.

(25) Para manna **sum** wæs, ..., bescoren preost, **sum** wes læwde, **sum** wæs wifmon.

‘One of these people (lit. men) was ... a tonsured priest, one was a layman, one was a woman.’ (DOEC, Bede 5 [0265 (13.428.1)])

This construction still existed in Middle English but was increasingly substituted by the *some of X* construction.⁸ As the independent usage of *sum(e)* is not the focus of this article, examples like (23)–(25) will not be investigated any further.

2.2.2 Dependent determiner usage in Old and Middle English

In Old and Middle English, *sum(e)* was also used dependently as a determiner with singular and plural noun heads and had several functions. One frequent usage was the partitive quantifier function (26) that has survived into Present-day English (see section 2.1).

(26) þæt ælc gesceaft bið healdon locen wið hire gecynde, ..., buton monnum & **sumum englum**.

‘that every creature is kept within bounds with its kind, [...] except human beings and some angels.’ (DOEC, Bo [0678 (25.57.5)])

evolving construction, an orthography which attempts to highlight the fact that the *-e* ending, which has become fossilized in the modern orthography, used to carry inflectional meaning, particularly concerning the singular/plural distinction.

⁷ This is due to the fact that in pre-head dependent uses *sum(e)* could not be combined with preceding determiners, which would have triggered a weak adjectival inflection.

⁸ and **somme of hem** ben to harde to thy tendir age. ‘and some of them are too hard to your tender age’ (PPCME, CMASTRO, 662.C1.11).

- (27) **Summe iuglurs** beoð þt ne cunne seruin of nan oðer gleo, buten makien cheres
 ‘There are some jesters who can provide no other entertainment but to make faces’
 (PPCME, CMANCRIW-1,II.157.2131)

Sum(e)’s general indefinite reading sometimes indicated an approximate amount or estimate, similar to an adverb with the sense ‘about, nearly, approximately’ (28), again a use that is still relatively common today (see [section 2.1](#)).⁹

- (28) þa hæfde he **sume hundred scipa**; þa wæron hi **sume ten gear** on þam gewinne
 ‘then he had about a hundred ships; then they were at war for about ten years.’
 (DOEC, Bo [1487 (38.115.17)])

In addition, *sum(e)* could combine with singular nouns in a way very similar to the ‘individualizing’ use of OE *ān* ‘one’, which would later go on to grammaticalize into the PDE indefinite article *a/an* (29)–(30).

- (29) þa wæs **sum consul** ... Boetius wæs gehaten.
 ‘There was some/one/a consul ... called Boethius.’ (DOEC, Bo [0006 (1.7.11)])
- (30) **Sum iungling** him fyligde mid anre scytan bewæfed ...
 ‘Some/one/a young man followed him clothed in a piece of cloth ...’ (DOEC, Mk (WSCp) [0588 (14.51)])

In the earlier literature on the topic, one finds an animated discussion concerning the question whether *sum(e)* can be regarded as an indefinite article for singular head nouns (Ropers 1918; Süßskand 1935; Rissanen 1967). For example, Mustanoja (1960: 211) remarks that ‘in OE and early ME the dependent use of *sum* is often practically the same as that of the indefinite article’.¹⁰ *Sum* and *ān* are regularly used to individualize and single out an individual among several of the same kind (31)–(32).

- (31) æt Finchamstæde **an mere** blod weoll
 ‘at Finchampstead a pool welled up blood’ (DOEC, ChronE [1550 (1098.5)])
- (32) Ðær wearð Alexander þurhscoten mid **anre flan** ...
 ‘There Alexander was pierced by an arrow ...’ (DOEC, Or 3 [0247 (9.73.18)])

However, both *ān* and *sum(e)* were extremely rare in this function. In the vast majority of cases, indefiniteness was not marked overtly by any determinative in Old English (see Sommerer 2018: 227 for frequency and distribution of *ān* in Old English). For example, indefiniteness marking was very rare in predicate NPs (33)–(34).

⁹ Like today, *sum(e)* was often combined with terms of time or space: ... *þær lagon sume hwile* ‘... (they) lay there for some time’ (DOEC, ChronC [0834 (1055.18)]).

¹⁰ ‘*An* and *sum* appear to have been used more or less side by side throughout the OE period in the individualizing function. Both are common in Ælfric. Süßskand [1935: 49–50] believes that this use of *an* was characteristic of spoken language, whereas *sum* was literary. But Rissanen [1967: 299–302] advances the possibility that the difference may originally have been dialectal’ (Mitchell 1985: 153).

- (33) He cuæð: ða ic hæfde ðone weall ðurhðyrelod, ða geseah ic **duru**
 ‘He said: When I had pierced through the wall, I saw a door’ (DOEC, CP [0723
 (21.155.3)])
- (34) Gif hwa slea his ðone nehstan mid **stane** oððe mid **fyste** ...
 ‘If anyone beats his neighbour with a stone or a fist’ (DOEC, LawAfeI [0027 (16)])

Similarly, overt marking of indefiniteness was the exception in NPs denoting the result of an action expressed by a transitive verb (35).

- (35) & þa gewrohte he **weall** mid turfum
 ‘and then he built a wall with turf’ (DOEC, ChronE [0051 (189.1)])

Following our definition, which only regards an item as an article if its use is obligatory (see section 2.1), and considering the sporadic use of *sum(e)* and *ān* in Old English, we conclude that the two cannot be classified as full articles at that stage. At best, they functioned as near-articles. This near-article function of *sum(e)* with singular nouns continued for some time into Middle English (36)–(37).

- (36) þenne hie beð ofþurst cumeð to **sum welle**
 ‘when she is thirsty, (she) comes to a well’ (PPCME, CMTRINIT, 199.2758)
- (37) vse wel this remedye, that whether thou slepe or wake, thy mynde be euere vpon sum sentence of holy writ or vpon **som seyntes lyf** ...
 ‘use this remedy well, so that whether you sleep or are awake, your mind is always focused on a sentence from the Bible or on a saint’s life’ (PPCME, CMAELR4, 12.328)

Ultimately, however, the singular near-article function was marginalized in Middle English compared to Old English due to the fact that *ān* supplanted *sum(e)* in this context. According to Mustanoja (1960: 262), *ān* became more and more frequent as a determinative (i.e. the ratio of *ān* versus *som(e)* being 10:1). It remains to be seen whether the empirical analysis in section 3 can corroborate this assessment.

Finally, in addition to the functions of *sum(e)* discussed so far, there is evidence for another function appearing in Middle English: near-article usage before plural noun heads (38)–(39).

- (38) ‘Lat us spyrrre **som tydynges**,’ seyde Percyvale ...
 ‘“Let us make some inquiries,” said Percival ...’ (PPCME, CMMALORY, 652.4362)
- (39) He receyued a letter fro þe kyng of Grete Britayn, ... þat he schuld sende **summe prestes** to þis lond to baptize him and his puple.
 ‘He received a letter from the king of Britain, ... that he should send some priests to this land to baptize him and his people.’ (PPCME, CMCAPCHR, 54.655)

Here, *sum(e)* is not used in a partitive way. Its meaning in *som tydynges* is not SOME BUT NOT ALL, but rather it is the non-specific, initiatory meaning, used to establish a new plural noun type in discourse: an unspecific, higher-than-one number of *tydynges*. One also finds this

near-article usage of *sum(e)* before mass nouns (40), where it denotes an unspecified amount (rather than unspecified number) of something.

- (40) take a boke and rede, or do **som labour** with thyn hondes
 ‘take a book and read, or do some work with your hands’ (PPCME, CMAELR4, 6.158)

Given the delayed appearance of the near-article function as part of the functional repertoire of *sum(e)*, and considering that it is found in plural and mass noun contexts, just like the partitive quantifier function, it is reasonable to assume that the near-article developed out of the partitive quantifier. This is parallel to the development of the indefinite article *a/an* from the numeral *ān*: in both cases, the primary meaning of quantifying N turned into the more basic and grammatical meaning of establishing the existence of N.

To conclude, *sum(e)* had a variety of functions in historical stages of English. One key function was that of a non-partitive indefiniteness marker for singular nouns alongside *ān* (33)–(34), which is prominent in the Old English period. In addition, there was the near-article function before mass and plural nouns (38)–(40), emerging in Middle English. The latter contrasts with the use of *sum(e)* as a quantifier with a partitive reading (26)–(27), from which the near-article function is diachronically derived.

So far, no empirical study has investigated the development of these functions from Middle English onwards using a quantitative methodology. It is our aim to fill this empirical gap by investigating when the OE individualizing usage with singular count nouns fell out of use as well as when and under what circumstances the near-article function with mass and plural nouns emerged. In section 4 we will describe the former as a case of constructional competition between two viable candidates for the grammatical role of singular indefiniteness marking, *sum(e)* versus *ān*, which was decided in favor of *ān*. The case of near-article emergence with mass and plural nouns, on the other hand, is hypothesized to have followed a constructionalization route via some interesting bridging contexts.

3 Empirical analysis

The functions and developments presented in the previous section generate a number of expectations towards linguistic corpus data for the periods between Old and Present-day English. In the following, some of these predictions will be discussed, alongside appropriate quantitative measures to determine to what extent they are supported by the surviving textual material.

3.1 Data and method

The data for the empirical investigation were supplied by two structurally parallel historical corpora: the *Penn–Helsinki Parsed Corpus of Middle English* (PPCME 2000) and the

Penn–Helsinki Parsed Corpus of Early Modern English (PPCEME 2004).¹¹ Together they cover almost 600 years of historical English (1150 to 1720), including more than 500 text samples, which in total amount to almost 3 million words. In addition to part-of-speech tagging, the corpora are also syntactically parsed, which allows searches for syntactic constituents such as phrase and clause types as well as syntactic function.

We used the open statistical software environment *R* (R Core Team 2018) to extract linguistic strings. In a first step, this meant extracting all NPs from the complete dataset. We further narrowed down the scope of our investigation to NPs functioning either as subjects or objects, to the exclusion of adverbials and other NP types. We did this to factor out the potential skewing effect of NPs that typically display a high amount of idiomaticity, such as adverbial adjuncts (e.g. *some days, some year, ...*). NPs with proper nouns and pronouns as their head were also discarded. From this base dataset – which comprises NPs that have a common noun (CN) as their head – we then extracted subsets including *sum(e)* which will be discussed individually in the following sections. We used *R* to calculate binary logistic regression models for both univariate and multivariate analyses (see Baayen 2008; Gries 2009; Levshina 2015).

3.2 *Determinative some with singular nouns, in contrast to indefinite a(n) and bare nouns*

First, we were interested in the diachronic development of *sum(e)* as an indefiniteness marker with singular nouns. To approach this question, we extracted NPs where singular common nouns were preceded by either *sum(e)* or the individualizing determinative *a(n)*. Note that the results exclude cases where *ān* functions as the numeral *one*. The numeral function and the article function are kept apart by two different POS-tags in the Penn–Helsinki corpora. Additionally, we extracted all singular NPs lacking any determinative before the nominal head (i.e. ‘bare NPs’). Our queries allowed for up to one non-complex adjectival pre-modifier before the nominal head. Thus, the following strings were extracted: [*sum(e)*+(ADJ)+CN_{sg}], [*a(n)*+(ADJ)+CN_{sg}] and [(ADJ)+CN_{sg}]. This means that the results do not include NPs with several adjectives in the pre-head. In that sense we have not analyzed all of the NPs which are determined by *ān* or *sum(e)*, but we consider this limited set to be a good start. No restrictions were placed on the post-head, which means that the singular CN head may be followed by a post-head element (any complementation pattern).

Figure 1 and the results from the binary logistic regression analyses¹² (Appendix tables A1–A3) reveal that the usage of *sum(e)* as a determinative (S) undergoes less dramatic and less monotonic (albeit significant) changes compared to the usage of *a(n)*, which increases dramatically (A). In contrast, the frequency of bare NPs (N) decreases. This is in line with

¹¹ For PPCEME, all text samples were used. For PPCEME, only a third of the text samples were used, viz. the text samples from the first Penn supplement to the original Helsinki corpus (‘p1’).

¹² We ran regression models on each of the three variants separately, following a ‘one versus rest’ approach, rather than running a genuine multinomial regression. For convenience the three variants are plotted together in figure 1.

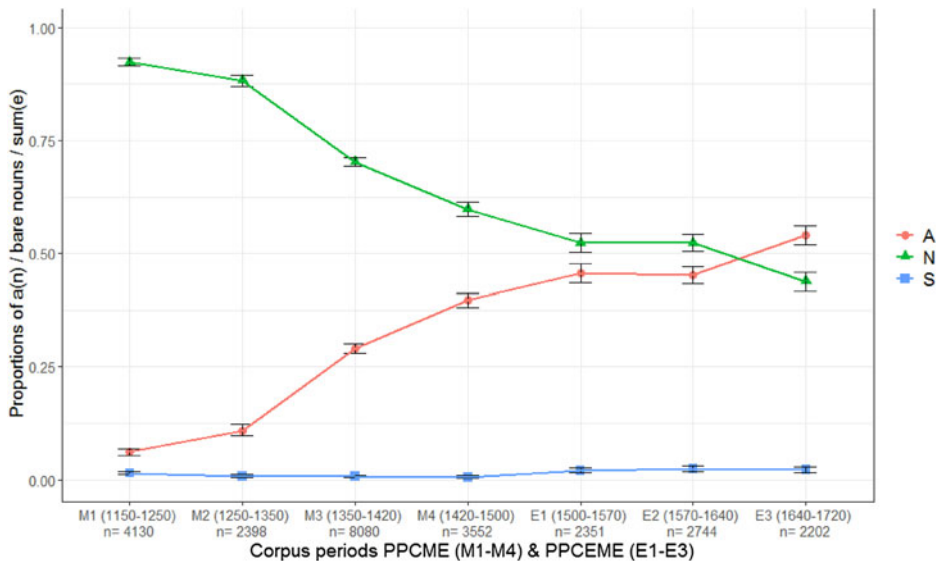


Figure 1. Diachronic development of determinative *sum(e)* (S) with singular nouns, compared to indefinite *a(n)* (A) and bare nouns (N). $N = 25,457$ (sample sizes for individual sub-periods given in graph). Brackets denote 95 percent confidence intervals (Wilson with continuity correction).

previous accounts which state that *ān* as an indefinite article increased dramatically in the ME period (Mustanoja 1960; Rissanen 1967).

At this stage in our analysis, *sum(e)* is included in all its singular functions (i.e. *sum(e)* as an individualizing article but also as a vagueness marker or quantifier with mass nouns). Moreover, the corpora are not lemmatized, and mass nouns (e.g. *water*, *love*, *honor*) are simply tagged as singular nouns. Thus, the results above also include NPs with mass noun heads, which were never eligible for determination with *a/an*.¹³ The corpus is also not tagged for type of reference. This means that not only referential but also non-referential NPs are included (especially in the case of bare NPs). In a project on the shifting strategies of indefiniteness marking, it would have been preferable to use only referential NPs as a base line, but identifying them would have gone beyond the scope of this article.

These limitations result in a relatively large amount of ‘statistical noise’ in our dataset (which was to some extent reduced in our later investigations; see section 3.3). Nonetheless, if our hypothesis that *ān* ousted *sum(e)* as a competitor for indefiniteness marking in singular count nouns as early as late Old English is correct, this should be

¹³ We consider mass nouns to be very different from singular count nouns and conceptually more similar to plural nouns, which is why mass nouns should be extracted from the data pool and treated separately. However, controlling this aspect would have amounted to manually coding thousands of individual examples with, at least for Middle English, highly variable spellings. Barring this labor-intensive option, we decided to ignore the count/mass distinction for this step in our analysis.

reflected in noticeably different diachronic trajectories of the two determinatives in relation to the unmarked competitor in the Middle English and Early Modern English period. This hypothesis is rather unambiguously borne out by the data. Figure 1 makes it clear that the two real competitors for the function in question are $a(n)$ and the bare noun variant, which is reflected in the almost symmetrical patterns of increase and decrease over time: where $a(n)$ soars, decisively and significantly, the share of singular bare nouns falls at an almost identical rate.¹⁴ As was shown in section 2.2, indefiniteness was not marked obligatorily in Old English. No indefinite article existed yet and, in the vast majority of cases, singular (and plural) nominals could occur bare if the context was indefinite. However, this changed in time so that $a(n)$ and bare nouns reached a more or less stable distributional state in Early Modern English, with both accounting for roughly half of the dataset. This can be interpreted as a paradigmatised division of labor: $a(n)$ has become an obligatory indefiniteness marker for singular count nouns in referential NPs, and the bare noun option has become standard for non-referential or mass nouns. The third competitor $sum(e)$ drops out of the competition early on, despite the fact that $\bar{a}n$ and $sum(e)$ both start out from very low rates in the earliest Middle English sources.

Thus, while it is possible, and congruent with our data, that $\bar{a}n$ and $sum(e)$ were still competing for the same constructional space in Old English, later diachronic data makes it evident that $a(n)$ rapidly emerged as the winner from this incipient competition, possibly due to its small (but significant) advantage in terms of frequency in the earliest records.

3.3 Determinative *some* with singular count versus mass nouns

As the results in figure 1 include not only singular count nouns but also mass noun examples, we next consider the relation between singular count and mass noun heads in more detail. In order to circumvent the limitations discussed above, we reduced our sample to occurrences of the form [$sum(e)$ + (ADJ) + CN_{sg}] and manually coded the individual examples for the count/mass distinction. Where the distinction was ambiguous, the mass reading was chosen. Figure 2 depicts the relative proportion of mass nouns in contrast to count nouns in the singular. We expected the share of $sum(e)$ + CN_{mass} to increase over time, assuming that $sum(e)$ loses its OE indefinite marking function for singular count nouns, thereby becoming available as an indefiniteness marker for mass nouns (as well as plural nouns, which are discussed in section 3.4). In other words, we expected a scenario where $sum(e)$ specializes its article-like use to go with mass nouns, in a similar way that $a(n)$ has specialized its use with count nouns. This prediction finds some support in the data, but the development is not as clear as in the case of $a(n)$. Inspecting the error bars in figure 2 reveals a large degree of overlap between the values for individual corpus periods due to the small number of occurrences of *some* with singular nouns overall (N = 309; see figure 2 for sample sizes of individual periods). Notably, at all times the

¹⁴ This visual interpretation is confirmed by a correlation test on the linear relationship between the two (Spearman's ρ , one-tailed): $\rho = -0.96$ ($p < 0.01$).

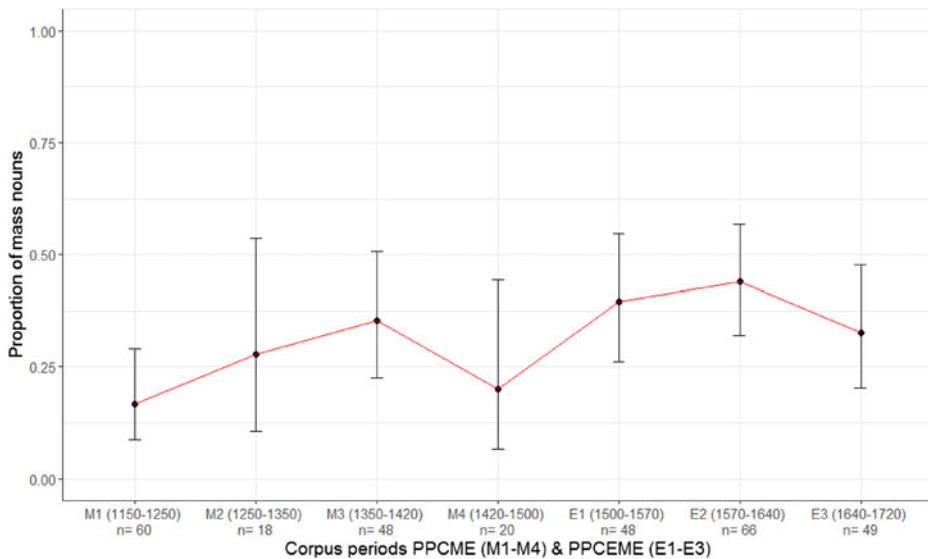


Figure 2. Diachronic development of determinative *sum(e)* with singular mass vs. count nouns. $N = 309$ (sample sizes for individual sub-periods given in graph). Brackets denote 95 percent confidence intervals (Wilson with continuity correction).

use of *sum(e)* with count nouns appears to remain more frequent than its use with mass nouns. Time does emerge as a significant predictor in the binary logistic regression model, but only relatively late and with a modest effect. We thus conclude that while the general trend of determinative *sum(e)* with mass nouns points upwards, the evidence falls short of providing conclusive evidence that *sum(e)* became strongly associated with mass nouns during the time when *a(n)* became tied to count nouns, or at least any such development does not manifest itself clearly in the textual record.

In other words, the quantitative evidence does not allow us to conclude that the frame [*sum(e)* + (ADJ) + CN_{sg}] underwent any sweeping changes during the time under investigation, neither for count nor mass nouns, but rather that frequencies remained rather stable at a very low rate throughout. This however, does not mean that *sum(e)* continued to function as a second article-like element alongside *a(n)* throughout the ME period. A closer look at individual examples suggests that what survives is the vagueness/approximation/evaluative function rather than the article-like use.¹⁵

3.4 Determinative *some* with plural count nouns

Having dealt with the occurrences of *some* before singular noun heads, we continue by investigating the use of *sum(e)* with plural nouns. For this, we extracted all hits for

¹⁵ E.g.: *Than sall þou dispose the for to pray, or for to thynke som gude thoughte ...* 'Then you shall turn to prayer or to thinking some good thought ...' (PPCME, CMROLLTR, 38.787).

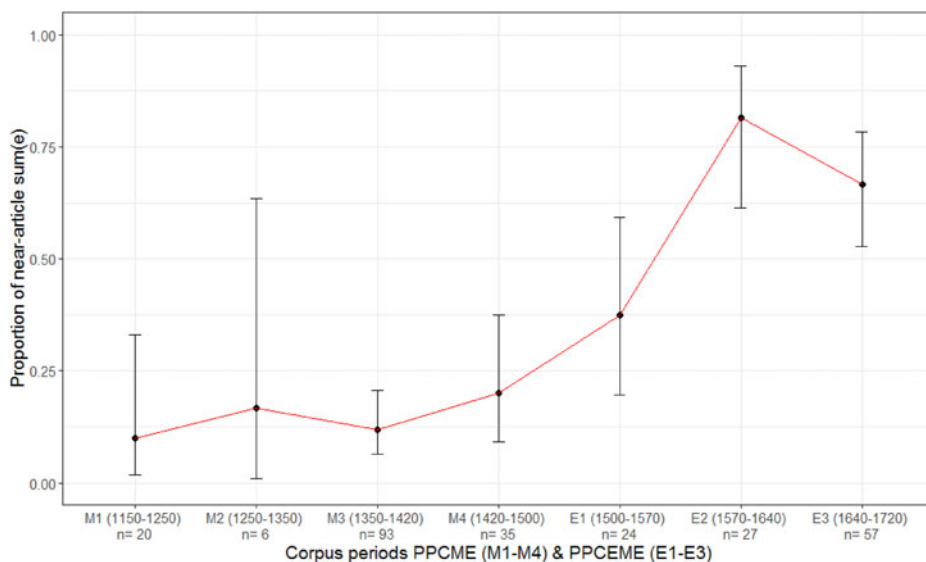


Figure 3. Diachronic development of partitive vs. near-article *sum(e)* with plural nouns. $N = 262$ (sample sizes for individual sub-periods given in graph). Brackets denote 95 percent confidence intervals (Wilson with continuity correction).

[*sum(e)* + (ADJ) + CN_{pl}]. Instead of attempting a ‘global’ analysis as in section 3.2, we proceeded on a slightly different route to establish when the near-article function in the plural became productive. We analyzed the retrieved examples ($N = 262$) with *sum(e)* in qualitative terms, coding them for partitive or near-article use in each individual example. An example was interpreted as partitive when either the semantics of the noun (e.g. *some members*; PPCEME, VICARY-E1-P1, 19.108) or the context of the passage (e.g. *as some pastures breed larger sheep, so do some Rivers [...] breed larger Trout*s; PPCEME, WALTON-E3-P1, 228.149) implied a partitive reading. Wherever a near-article reading was plausible, the example was coded as such. A few cases where no clear decision could be reached were discarded. Initial results are shown in figure 3, the results from the regression analysis can be found in the Appendix (table A5).

There is a strong and statistically significant increase of near-article use in the data. While near-article readings account for only a fraction of the retrieved examples at the beginning of the Middle English period, they increase rapidly in Early Modern English, concluding the investigation period at a level of around two thirds of all plural occurrences. On the face of it, this trend nicely aligns with our assumption that *sum(e)* in plural contexts has over time developed a non-partitive use that is in many ways parallel to the indefinite article *a(n)* in the singular.

However, there are factors that confound this seemingly clear picture of diachronic increase, which are linked to the heterogeneity of the textual material. In addition to time, the rates of near-article and partitive usage may also depend on register. Since grammaticalization processes typically originate in informal language use, we

hypothesize that more informal text types exhibit a higher proportion of near-article usage than more formal text types. At the same time, we recognize that the texts in the two Penn–Helsinki corpora are distributed highly unevenly in this regard. This is due to the limited spread of literacy in medieval times, as well as to the vagaries of textual transmission, which have disadvantaged private, unpublished writings throughout history. Thus, the early periods of PPCME largely consist of texts of a very formal style, such as religious treatises and homilies. More informal and speech-based text types, including private letters, diaries or trial proceedings, only appear in Late Middle and Early Modern English.

To establish, first, whether the assumed correlation between near-article usage and level of formality holds true, and second, whether the observed diachronic increase in the near-article function is real or merely an artifact of the skewed distribution of texts in the corpora, we decided to include formality as an additional predictor. We assigned one of two formality values, ‘formal’ and ‘informal’, to the texts in the corpora.¹⁶

As a third predictor variable, we coded presence or absence of modification in the NP. The rationale underlying this variable is that we expect NPs with pre- or post-modification to have intrinsically partitive semantics. In specifying the noun, a modifier reduces the set of possible referents denoted by the noun to a smaller sub-set sharing the features expressed by the modifier. It is conceivable that the grammaticalization process was even initiated in modified NPs, where *sum(e)* is redundant as a carrier of partitive meaning, thus lending itself to reinterpretation as an indefiniteness marker.

The addition of two more predictors necessitated the construction of a multivariate statistical model in order to control for the influence of each predictor variable on any observed correlation between the other predictors and the output. We therefore built a binary logistic regression model, in which the probability of near-article use serves as the output (dependent) variable (ARTICLE), while PERIOD (nominal scale, with Helmert contrasts), level of FORMALITY (nominal scale) and presence or absence of MODIFICATION (nominal scale) serve as predictor (independent) variables. Since the likelihood of near-article usage in a particular token may also depend on the individual text (or its writer), we included TEXT as a random effect (Baayen 2008: 241–84).¹⁷

On closer inspection of the model statistics (table 2), all three predictors turn out to be significant (PERIOD: $p < 0.05$; FORMALITY: $p < 0.01$; MODIFICATION: $p < 0.001$), meaning they all have a statistically relevant impact on the rate of near-article use in contrast to partitive use. No statistically relevant interactions among the predictors were detected.¹⁸ The direction of the impact is visualized in figure 4, where each panel

¹⁶ The formal category includes homilies, treatises, handbooks, legal texts, histories, philosophical and scientific works, sermons, travelogues, biographies and official letters. The informal category consists of private letters, romances, fiction, drama, diaries, trial proceedings and Bible translations. Since the term ‘informal’ here is meant to include texts that are relatively more speech-based (see Koch & Oesterreicher 1985), trial proceedings and Bible passages were accepted as representatives of informal language use due to the frequent use of direct and indirect speech in both text types.

¹⁷ We thank one anonymous reviewer for suggesting that TEXT be included as a random effect.

¹⁸ We performed ANOVA to test if interactions between the predictors improved the models (Levshina 2015: 267).

Table 2. *Multivariate binary logistic regression model (significance codes: 0.001 '***' / 0.01 '**' / 0.05 '*')*

Formula: ARTICLE ~ PERIOD + FORMALITY + MODIFICATION + (1 + MODIFICATION TEXT)					
AIC: 206.7		logLik: -91.3	Obs: 262		
<i>Random effects</i>					
Groups	Name	Variance	Std. deviation	Corr	
TEXT	(Intercept)	2.219	1.490		
	MODIFICATION unmodified	2.810	1.676	-0.51	
Number of groups: 86					
<i>Fixed effects</i>					
Predictors	Levels	Estimate	Std. error	Z value	p-value
PERIOD (Helmert contrasts)	(Intercept)	0.026188	0.467111	0.056	0.955291
	Contrast level	Ref. level			
	M2	0.491805	1.083973	0.454	0.650041
	M3	0.003718	0.413890	0.009	0.992832
	M4	0.218149	0.307116	0.710	0.477508
	E1	0.291097	0.218143	1.334	0.182062
FORMALITY	E2	0.452625	0.189329	2.391	0.016817*
	E3	0.235448	0.116620	2.019	0.043495*
	formal	Ref. level			
MODIFICATION	informal	2.898330	1.111813	2.607	0.009138**
	modified	Ref. level			
	unmodified	-2.751340	0.741467	-3.711	0.000207***

represents one of the predictor variables. The upper panel shows that the near-article function increases significantly over time even when formality level and modification are controlled, thereby confirming that the observed trend cannot simply be reduced to an imbalance of text types in the corpora. This particular model uses Helmert contrasts for the predictor TIME, so that each level (i.e. corpus period) is compared to the mean of all previous levels, rather than to one reference level that is kept constant. Thus, the exact value of the contrast level changes the further one moves through the graph from left to right on the x-axis. Helmert contrasts are particularly useful for the analysis of diachronic aspects of linguistic variation, especially when the temporal intervals represent unequal amounts of time, as is the case for the periods in the Penn–Helsinki corpora.¹⁹ The y-axis in this graph represents log odds ratios, a (log-transformed) measure of how much more likely it is to encounter the near-article function in a specific corpus period compared to the contrast level. Thus, the log odds estimate for

¹⁹ We thank one anonymous reviewer for suggesting that we use Helmert contrasts.

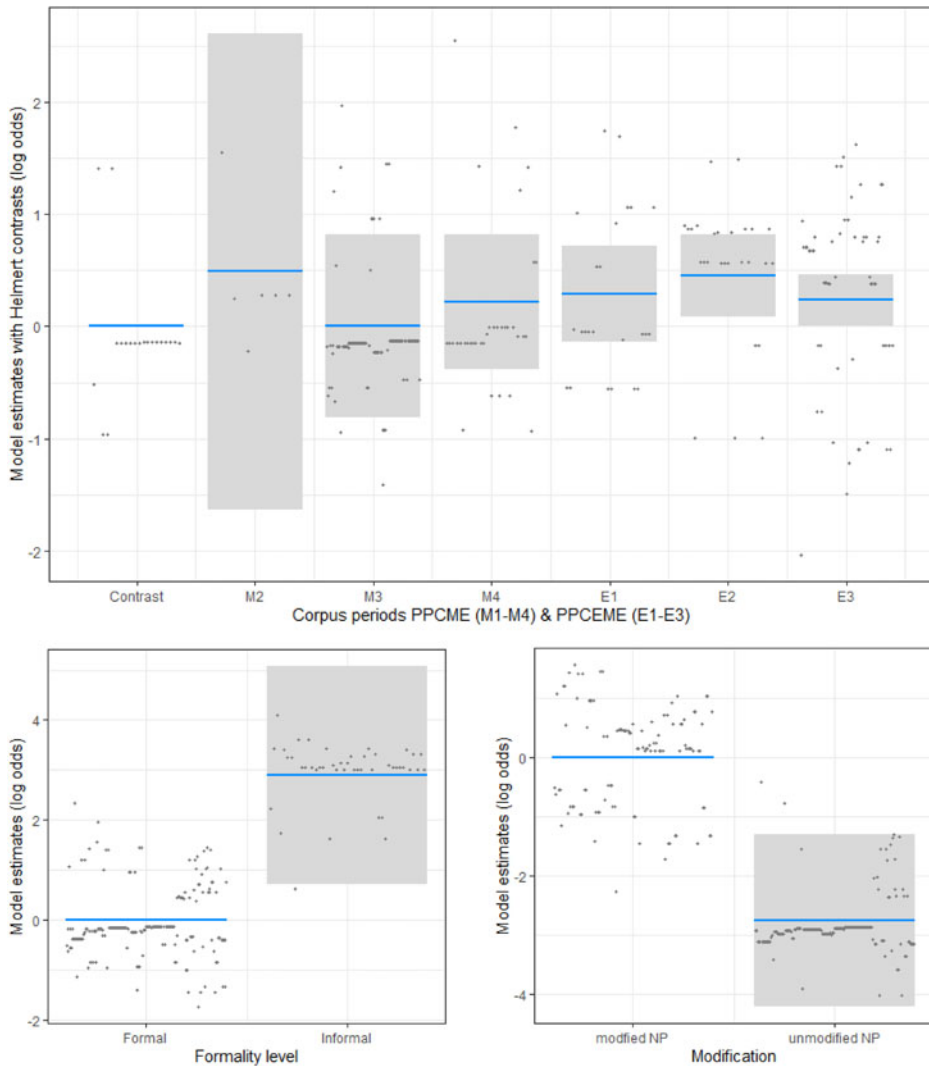


Figure 4. Diachronic development of partitive vs. near-article *sum(e)* with plural nouns. $N = 262$. Grey areas around the estimates denote 95 percent confidence intervals (profile likelihood).

the corpus period E2 (0.45) corresponds to a 57 percent increase of the near-article function in that corpus period compared to the mean of all previous periods. The lower left panel indicates that informal text types have a much higher likelihood to include near-article uses of *sum(e)* than formal ones. Finally, in the lower right panel we see that modified NPs are much more likely to appear with near-article *sum(e)* than with partitive *sum(e)*.

In sum, the multivariate logistic regression model supports our hypothesis that determinative *sum(e)* underwent a process of grammaticalization in the Middle and Early Modern English periods, leading to the emergence of a novel, article-like use with

plural NPs. The fact that our analysis also finds corroborating evidence for our more fine-grained predictions concerning the association of near-article *sum(e)* with informal language and modified NPs represents particularly compelling evidence in this respect.

4 A constructional sketch

4.1 Diachronic Construction Grammar (DCxG)

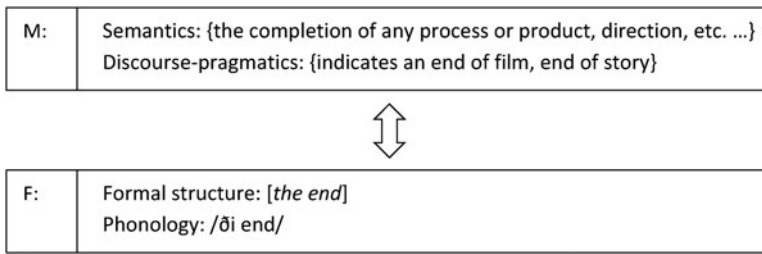
In the past fifteen years, several versions of Construction Grammar have been growing fast as cognitive-functional approaches to language (see e.g. Croft 2001; Tomasello 2003; Croft & Cruse 2004; Goldberg 2006, 2019; Trousdale & Gisborne 2008; Hoffmann & Trousdale 2011, 2013; Steels 2011; Boas 2013; Boogaart, Coleman & Rutten 2014; Hilpert 2014; Diessel 2015, 2019; Ellis, Römer & O'Donnell 2016). Although these models differ substantially from one another, most of them share several basic tenets, some of which we discuss briefly at this point.

Linguistic knowledge is assumed to be usage-based and all linguistic generalizations are derived from the user's experience with language (actual usage events). Syntactic structure emerges through repetition, categorization and conventionalization rather than resulting from an innate matrix (e.g. Bybee 2010; Diessel 2019).

Language consists of constructions; i.e. of conventionalized form-meaning pairings in the sense introduced by de Saussure (Goldberg 2006: 3). A construction is a symbolic sign, which links a formal side to a particular meaning/function via a symbolic correspondence link (Croft & Cruse 2004: 258). Linguistic knowledge about the formal, phonological and orthographic structure of a construction is paired with semantic and discourse pragmatic knowledge that the speaker has acquired about the meaning, function and use of this conventionalized string (e.g. [*the end*]_{Cx} in figure 5).

Regarding their formal shape, constructions can be atomic and substantive (e.g. lexical items like [*car*], [*drive*], [*love*], or function words like [*the*], [*who*]) or atomic and schematic (e.g. an abstract category like [DEM] 'demonstrative' or [CN] 'common noun'). Constructions can also be complex. In that case they are either fully specified (e.g. fixed phrases like [*ladies and gentleman*]_{NP}), semi-specified with substantive and schematic parts (e.g. [*a hell of a CN*]_{NP}, [*a + CN_{sing}*]_{NP_{indef}}), or completely schematic (e.g. NP grounded by a definite determiner [DET_{def} + CN]_{NP_{def}}). Semantically they can be fully compositional and transparent (e.g. [*next year*]) or non-compositional and non-transparent, figurative (e.g. [*SAW_{tensed} logs*]). In other words, complex constructions can have sequential structure with positions that are either fixed and lexically filled or open; their meaning can be semi- or fully idiomatic or completely compositional (Croft & Cruse 2004: 255; Diessel 2015: 312; Smirnova & Sommerer forthcoming).

'Rules' in the traditional generative sense have been replaced by schemas (i.e. grammatical templates which have evolved over concrete tokens) (Croft & Cruse 2004; Diessel 2015; Tomasello 2003; Hoffmann & Trousdale 2013). Langacker (2008: 23) defines schemas as 'abstract templates obtained by reinforcing the commonality inherent in a set of instances. Since grammatical rules are patterns in the formation of

Figure 5. [the end]_{Cx} construction

symbolically complex expressions, they are themselves symbolically complex as well as schematic'. One of the most attractive ideas of Construction Grammar is that we can assign procedural meaning to abstract syntactic templates, independent of the meaning of the words being used in them (Boas 2013: 236) (see figure 6).²⁰

Finally, all the constructions of a language form the 'constructicon': 'a structured inventory, which can be represented by multiple inheritance networks' (Croft & Cruse 2004: 262–5). Constructions are considered as nodes which are organized in taxonomic and meronymic networks of constructional families (Barðdal & Gildea 2015: 23). Linguistic knowledge is represented in two-dimensional networks where lower level constructions are said to inherit features from higher-level constructions through vertical links. Vertical links are instance or inheritance links. If two constructional nodes are connected vertically, the lower node is a more specified version of the higher node and inherits general information from it. Complex abstract schematic constructions can exert influence on more substantive constructions and vice versa. Moreover, horizontal links between constructions on the same level of abstraction exist. Horizontal links express all kinds of paradigmatic relationships between constructions on the same level (for a more detailed account of vertical and horizontal connections and what they express, see Diessel 2015, 2019; Smirnova & Sommerer forthcoming).

In figure 6,²¹ a partial network sketch of definite NP constructions in Present-day English exemplifies the structure of a network. In the mid positions of this network, there are semi-specified schemas like [[*this*]_{DEM} + [CN_{sg}]]_{NPdef} and [[*the*]_{ART} + [CN_{infl}]]_{NPdef}. Influenced by the functional and formal similarity of these strings, it is assumed that the speaker has also abstracted several higher levels, where one finds completely abstract

²⁰ As constructions are based on generalizations rather than actual utterances, input frequency (both type and token frequency) is crucial as it influences the entrenchment of any construction and can also trigger the emergence of abstract grammatical schemas (Hoffmann & Trousdale 2011: 5; also Goldberg 2006: 39, 98–101; Hilpert & Diessel 2016).

²¹ Figure 6 is a preliminary sketch of a possible network of definite NPs in PDE open to discussion. Note that it is partial in the sense that especially at the lower levels not all assumed connections (i.e. existing links between related constructions) have been visually added. For example, the triangles indicate that none of the assumed connections have been sketched.

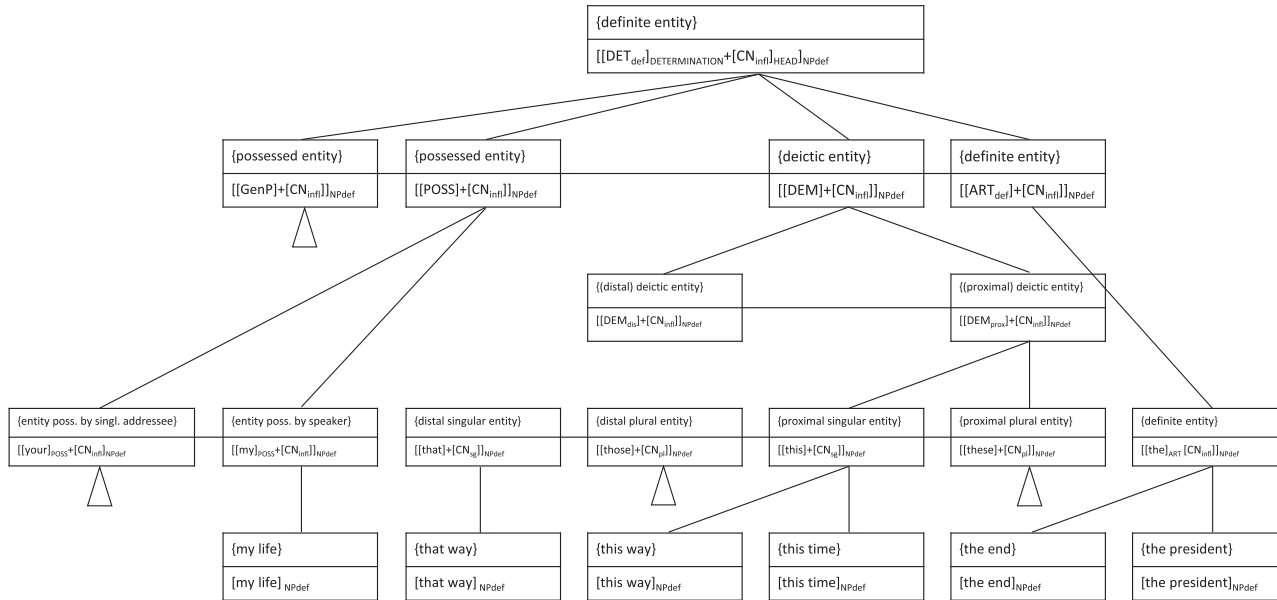


Figure 6. Partial network sketch of definite NPs with common noun head in English

schemas like $[[\text{DEM}] + [\text{CN}_{\text{infl}}]]_{\text{NP}_{\text{def}}}$ or $[[\text{POSS}] + [\text{CN}_{\text{infl}}]]_{\text{NP}_{\text{def}}}$. On the lowest level, one finds fully specified constructions like *[this time]* or *[this way]*, *[the end]* or *[my life]*. These qualify as separate nodes due to their high frequency and specific discourse pragmatic usage. A speaker's network 'can grow "upwards" via schematization, "outwards" via extension and "downwards" as more detailed instances are added' (Evans & Green 2006: 546). Crucially, the formal and functional fate of linguistic forms and constructions is influenced – among other things – by related constructions in the network. This brings us to language change.

4.2 *Constructional competition and network reconfiguration*

Usage-based cognitive Construction Grammar also aims to incorporate a diachronic/variationist perspective (e.g. Israel 1996; Bergs & Diewald 2008; Hilpert 2013, 2018; Traugott & Trousdale 2013; Trousdale 2014; Coussé, Andersson & Olofsson 2018; Sommerer 2018; Van Goethem, Norde, Coussé & Vanderbauwhede 2018; Zehentner 2019; Smirnova & Sommerer forthcoming). In current versions of DCxG, all types of linguistic change are being reconceptualized as 'network changes'. It has been suggested that the network can change:

- (i) via node creation and node loss ('constructionalization' and 'constructional death'),
- (ii) via node-internal changes ('constructional change'), and
- (iii) via node-external changes, i.e. constructional network reconfiguration (Hilpert 2018; Smirnova & Sommerer forthcoming)

Diachronically, 'constructionalization' (i.e. the emergence and entrenchment of a new form-meaning pairing) but also the potential marginalization or death of a construction can, among other things, be influenced by discourse-pragmatic needs (functionally driven), by form-driven frequency effects and/or cognitive preferences like processing efficiency or analogical reasoning skills (e.g. Fischer 2007; Hoffmann & Trousdale 2011; Sommerer 2012, 2015; De Smet 2013).

The network does not only change when new nodes are added but also when node-external links between constructions are rearranged. Hilpert (2018) calls those 'connectivity changes' in which the network undergoes some rewiring. Links between existing constructions may fade and disappear or new links may emerge (Torrent 2015; Lorenz forthcoming).

Grammaticalization is reconceptualized as a special case of 'grammatical constructionalization' (e.g. Trousdale 2014; Coussé, Andersson & Olofsson, 2018). A new form-meaning pairing constructionalizes which is semantically less heavy and more procedural than the one of the source construction (Trousdale 2014: 113).

Finally, an important concept is that of 'constructional competition' (e.g. Zehentner 2019; Sommerer forthcoming), i.e. two constructions compete until one ousts the other:

[I]n many cases, the old form drops out of use and the new form goes on to become the canonical, or even the only, way to code the distinction in question. (Barðdal & Gildea 2015: 38)

Of course, it is also possible that two forms reach an equilibrium and coexist, if the constructions find respective niches (see also Fonteyn & Heyvaert 2018; Zehentner & Traugott *forthcoming*). Investigating the various (changing) functions of *sum(e)* in Middle English offers an excellent opportunity to show in more detail what is meant by such general statements.

4.2.1 Demise of the individualizing usage with singular count nouns

It was shown earlier in this article that in Old English *sum(e)* and *ān* were infrequently used in front of singular nouns to ‘individualize’ (in contrast to their relatively frequent usage as a quantifier or numeral). However, after a phase of functional coexistence in Old English, *sum(e)* was ousted from this function by the numeral *ān*. We argue that the observable development is a result of constructional competition, which unfolded in the following way: in West Germanic and early Old English, definiteness and indefiniteness were not marked obligatorily. In referential NPs, the common noun did not need to combine with a determinative (i.e. a grounding element in the sense of Langacker (2008)). Broadly speaking, this means that referential NPs were constructed in exactly the same way as non-referential NPs: *I see abbot in the cloisters* (referential) versus *He was elected abbot* (non-referential). However, many CNs in definite and indefinite contexts often collocate with definite or indefinite determinatives (demonstratives, possessives, numerals, quantifiers) which indirectly express (in)definiteness next to other semantic content (e.g. possession, spatial deixis, number etc.). Still, any marking of definiteness or indefiniteness happens ‘parasitically’ and is optional (see table 3).

This stage of non-marking puts a heavy decoding pressure on the listener. Often nothing in the input overtly helps the listener to decode the referential status of the message. The speaker/listener has to guess from context whether the noun phrase is referential or non-referential or whether the referent is conceptually definite or indefinite. That is why we argue that at one point in time the speakers/listeners – influenced by frequency and analogy effects – change their marking strategy. As a first step, the marking of definite contexts becomes obligatory in all referential cases between early and late Old English. Speakers adopt a new default strategy to mark singular, plural count and mass nouns obligatorily. One reason to ground the referent by an overt element and to code intertextual (anaphoric) relations overtly is that it is cognitively and communicatively efficient for the listener (Hawkins 2004).²²

In constructional terms, this new obligatory and overt marking strategy corresponds to the emergence of an abstract definite NP schema with a determination slot that has to be filled: $[[\text{DET}_{\text{def,infl}}]_{\text{DETERMINATION}} + [\text{CN}_{\text{infl}}]_{\text{HEAD}}]_{\text{NPdef}} \leftrightarrow \{\text{definite entity}\}$. The existence of this construction leads to the recruitment of the OE demonstrative *se* as an obligatory default marker of definiteness, triggering its grammaticalization into the definite article *the* (Sommerer 2015; 2018). What can be observed is a case of

²² See Sommerer (2018) for a detailed account of potential reasons why this abstract definite construction emerged.

Table 3. *Shifting strategies of (in)definiteness marking in English*

	Referential NPs					
	Overt obligatory definiteness marking		Overt obligatory indefiniteness marking		Non-referential NPs	
	Singular	Plural/Mass	Singular	Plural/Mass	Singular	Plural
Early OE	–	–	–	–	–	–
Middle OE	(+)	(+)	–	–	–	–
Late OE	+	+	–	–	–	–
Early ME	+	+	(+)	–	–	–
Middle ME	+	+	+	–	–	–
Late ME/EME	+	+	+	(+)/–	–	–

grammatical constructionalization, where a new node constructionalizes $[se_{infl}]_{ART} \leftrightarrow \{\text{definiteness marker}\}$ that is more ‘grammatical’ than its source node: $[se_{infl}]_{DEM} \leftrightarrow \{\text{marker of situational or intertextual deixis}\}$.

Late Old English sources also reveal that at that point, indefiniteness marking is not yet obligatory. What helps speakers to interpret a message is the following opposition: definiteness is marked overtly; indefiniteness is indicated by leaving the determination slot empty (see table 3). This is the stage at which OE *sum* but also OE *ān* are sometimes used in an individualizing, non-partitive function.

However, around early Middle English the textual sources suggest that speakers shifted to the overt marking of indefiniteness in singular NPs as well. Influenced by the already emerged definite template, an indefinite abstract schema develops:

$$[[\text{DET}_{\text{indef, infl}}]_{\text{DETERMINATION}} + [\text{CN}_{\text{count, sg, infl}}]_{\text{HEAD}}]_{\text{NP}_{\text{indef}}}$$

Again this schematic template formalizes a new coding strategy; a newly entrenched procedural routine, which is to mark indefinite contexts overtly as well. The semantically bleached numeral *ān* is recruited as the default slot filler to be used in NPs with a singular noun head. $[\bar{a}n]_{\text{NUM}} > [\bar{a}n]_{\text{ART}_{\text{indef}}}$. Figure 7 shows this step of network reconfiguration.

The speakers analogically extend the definite template to the indefinite realm. We again assume that speakers introduce overt marking of indefiniteness because it helps them distinguish between referential and non-referential NPs, increases processing efficiency and is communicatively more efficient (again see Sommerer 2018).

It would have been possible for OE *sum(e)* to be recruited for the job of a default indefiniteness marker for singular CNs as well, but ultimately *ān* is preferred. One reason why *ān* might have been preferred is that it was simply more frequent as a pre-head dependent from the beginning (see figure 1). This is why at this point, *sum(e)* sheds its function as an individualizing element and speakers no longer produce

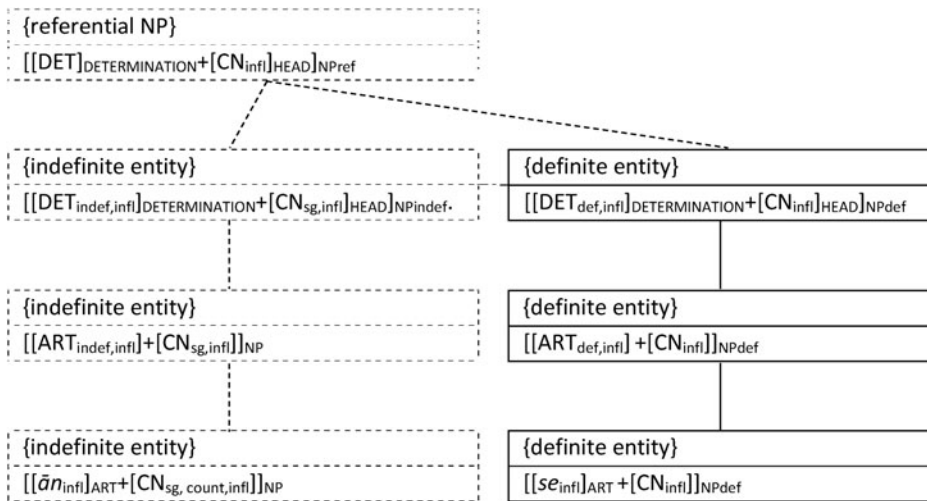


Figure 7. Emergence of abstract indefinite CNPs schema and indefinite article in early ME (Sommerer 2018: 286)²³

examples like *sum wif hatte Uenus* ‘there was a woman called Venus’ (DOEC, WHom 12 [0026 (77)]). Two forms share the same function for some time but at one point one form takes over completely. This, however, does not mean that *sum(e)* loses its partitive use or its other functions. It continues to be used as a vagueness/approximation marker and a quantifier.

4.2.2 Emergence of new indefinite plural article node

Our empirical analysis also reveals that by the mid fourteenth century, speakers mark indefiniteness with singular nouns very consistently but interestingly not yet with plural and mass nouns. However, after 1530 and especially in Early Modern English the non-partitive function of *sum(e)* in front of plural and mass nouns increases significantly, eventually outpacing the original partitive reading in those contexts. Here, we argue that as a next step a new abstract template for plural and mass nouns develops – $[[ART_{indef,infl}]+[CN_{pl/mass,infl}]]NP_{indef}$ – and a *sum(e)* node constructionalizes as an indefinite near-article for plural and mass nouns, alongside the indefinite article *a(n)* for singular count nouns (see figure 8). In more traditional parlance, ME *sum(e)* starts to grammaticalize into an indefinite near-article: e.g. $[sum_{infl}]_{QUANT} > [sum_{infl}]_{ART}$.

As we have argued throughout, we classify this *sum(e)* as ‘near-article’ because it is not obligatory, so that even in PDE it is still grammatical to leave the determination slot empty (e.g. *Cars are driving on the road*). In indefinite NPs with plural and mass nouns, the nominal can still occur as a bare noun: $[[CN_{pl/mass,infl}]]NP_{indef} \leftrightarrow \{\text{indefinite entities}\}$.

²³ Dotted lines indicate emerging constructions and links; solid lines indicate already established constructions at that point.

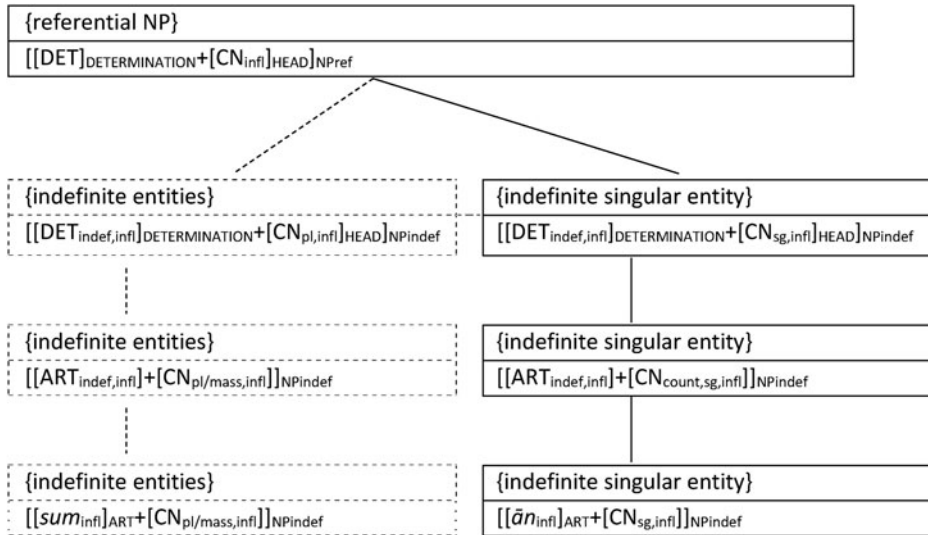


Figure 8. Network reorganization in indefinite CNPs in late ME/EME

The empirical analysis in section 3.4 also suggests that *sum(e)* takes up its new function in some contexts earlier than in others. For example, it was shown that it spreads earlier in informal texts. This can be accounted for in a constructional model by adding discourse-pragmatic knowledge to the node:

Sem: {indefinite entities}
 Dis-prag: {feature of informal style}
 $[[sum_{infl}]_{ART} + [CN_{pl/mass,infl}]]_{NPIndef}$

At the same time, the data also showed that, when *sum(e)* starts to spread with plural CNs, it is used primarily in constructions with pre- and or post-head modification. This can be expressed by the postulation of more fine-grained templates, which function as bridging contexts. Obviously, the modifier restriction is given up soon and *sum(e)* extends to the non-modified cases.

$$[[sum_{infl}]_{ART} + [ADJ] + [CN_{pl/mass,infl}]]_{NPIndef}$$

$$[[sum_{infl}]_{ART} + ([ADJ]) + [CN_{pl/mass,infl}] + [Complement/Modifier]]_{NPIndef}$$

$$[[sum_{infl}]_{ART} + [CN_{pl/mass,infl}]]_{NPIndef}$$

Seen from a different angle, it also seems possible to argue that *some* in its non-partitive, article-like function ‘completes’ the article paradigm of English.

Table 4. *(In)definiteness marking in Present-day English*

	DEFINITE		INDEFINITE	
	Count	Non-count	Count	Non-count
SINGULAR	the book	the ink	a book	–/(some) ink
PLURAL	the books		–/(some) books	

Whereas for definite NPs, the article *the* is used not only for singular but also for plural/mass heads, we have a division-of-labor situation for the indefinite contexts. *A(n)* is used for singular contexts, *some* for plural and mass contexts. The addition of *some* as an overt element which can function as a marker of indefinite contexts completes the English article paradigm (see table 1, reproduced here as table 4 for the reader's convenience).

English has thereby reached a stage in which referential contexts can be marked overtly by a specific element. Here, we would like to make two important remarks: obviously, at all historical stages where we have claimed that (in)definiteness marking has become obligatory, one can find examples of definite and indefinite constructions which do not take an article. Some groups of nouns resist the overt marking process longer than others (e.g. unique nouns, etc.).

At the same time there are several cases in Present-day English where article usage has been extended to the non-referential or generic domain (e.g. *She plays the piano*; *A lion is a dangerous animal*). This means that the development depicted in table 3 is not as clear-cut as it seems. However, we are concerned with the majority strategy and conceptualize constructions deviating from it as semi-specified or fully specified constructions on the lowest levels of the network, which block inheritance from above.

5 Conclusion

In this article, we analyzed the development of ME *sum(e)* from a constructional point of view using a quantitative methodology. It was shown that the usage of *sum(e)* with singular nouns drops in Middle English as a reaction to the recruitment of the OE numeral *ān* as the indefinite article (i.e. constructional competition). The indefinite, near-article function in plural CNPs only develops and starts to spread in late Middle English (mid fourteenth century) (i.e. constructionalization). This happens first in informal texts and in certain complex constructions. *Sum(e)* became an indefinite near-article for plural and mass nouns due to a shifting strategy of definiteness marking in English: from covert to overt and obligatory marking of referential definite and indefinite NPs. Originally lacking an indefinite article category, English only developed one as a result of the complete systemic reorganization of (in)definiteness marking in late OE and ME.

Authors' addresses:

Department of English and American Studies
University of Vienna
Spitalgasse 2-4
1090 Vienna
Austria
lotte.sommerer@univie.ac.at
klaus.hofmann@univie.ac.at

References

- Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Barðdal, Jóhanna & Spike Gildea. 2015. Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications. In Barðdal, Smirnova, Sommerer & Gildea (eds.), 1–50.
- Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.). 2015. *Diachronic Construction Grammar*. Amsterdam and Philadelphia: John Benjamins.
- Bergs, Alexander & Gabriele Diewald (eds.). 2008. *Constructions and language change*. Berlin: Mouton de Gruyter.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.
- Boas, Hans. 2013. Cognitive Construction Grammar. In Hoffmann & Trousdale (eds.), 233–55.
- Boogaart, Ronny, Timothy Colleman & Gijsbert Rutten (eds.). 2014. *Extending the scope of Construction Grammar*. Berlin: Mouton de Gruyter.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Carlier, Anne. 2007. From preposition to article: The grammaticalization of the French partitive. *Studies in Language* 31(1), 1–49.
- Chesterman, Andrew. 1993. Articles and no articles. In Andreas Jucker (ed.), *The noun phrase in English: Its structure and variability*, 13–24. Heidelberg: Universitätsverlag C. Winter.
- Christophersen, Paul. 1939. *The articles: A study of their theory and use in English*. Oxford: Oxford University Press.
- Coussé, Evie, Peter Andersson & Joel Olofsson (eds.). 2018. *Grammaticalization meets Construction Grammar: Opportunities, challenges and potential incompatibilities*. Amsterdam and Philadelphia: John Benjamins.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William & Alan D. Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- De Smet, Hendrik. 2013. *Spreading patterns: Diffusional change in the English system of complementation*. Oxford: Oxford University Press.
- Diessel, Holger. 2015. Usage-based construction grammar. In Ewa Dąbrowska & Dagmar Divjak (eds.), *Handbook of cognitive linguistics*, 295–321. Berlin: De Gruyter Mouton.
- Diessel, Holger. 2019. *The grammar network. How linguistic structure is shaped by use*. Cambridge: Cambridge University Press.

- DOEC = Antonette diPaolo Healey, John Price Wilkin & Xin Xiang. 2009. *Dictionary of Old English Web Corpus*. University of Toronto. <https://tapor.library.utoronto.ca/doecorpus/> (accessed 8 May 2019).
- Downing, Angela & Philip Locke. 2006. *English grammar: A university course*, 2nd edn. London and New York: Routledge.
- Ellis, Nick C., Ute Römer & Matthew B. O'Donnell. 2016. *Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of Construction Grammar*. Malden, MA: Wiley-Blackwell.
- Evans, Vyvyan & Melanie Green. 2006. *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- Fischer, Olga. 2007. *Morphosyntactic change: Functional and formal perspectives*. Oxford: Oxford University Press.
- Fonteyn, Lauren & Liesbet Heyvaert. 2018. Category change in the English gerund: Tangled web or fine-tuned constructional network? In Van Goethem, Norde, Coussé & Vanderbauwhede (eds.), 149–77.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, Adele E. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton, NJ: Princeton University Press.
- Gries, Stefan Th. 2009. *Statistics for linguistics with R: A practical introduction*. Berlin: De Gruyter Mouton.
- Halliday, Michael A. K. & Christian M. I. M. Matthiessen. 2014. *An introduction to functional grammar*, 4th edn. London: Routledge.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Heltveit, Trygve. 1977. Aspects of the syntax of quantifiers in Old English. *Linguistics* 31, 47–94.
- Hilpert, Martin. 2013. *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.
- Hilpert, Martin. 2014. *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press.
- Hilpert, Martin. 2018. Three open questions in diachronic construction grammar. In Coussé, Andersson & Olofsson (eds.), 21–39.
- Hilpert, Martin & Holger Diessel. 2016. Entrenchment in Construction Grammar. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning*, 57–74. Berlin: De Gruyter Mouton.
- Hoffmann, Thomas & Graeme Trousdale. 2011. Variation, change and constructions in English. *Cognitive Linguistics* 22(1), 1–23.
- Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford handbook of Construction Grammar*. Oxford: Oxford University Press.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Israel, Michael. 1996. The way constructions grow. In Adele E. Goldberg (ed.), *Conceptual structure, discourse and language*, 217–30. Stanford, CA: CSLI.
- Israel, Michael. 2000. Some and the pragmatics of indefinite construal. *Annual Meeting of the Berkeley Linguistics Society* 25(1), 169–82.
- Jacobsson, Bengt. 2002. The indefinites *some* and *any* in linguistic theory and actual usage. *Studia Neophilologica* 74(1), 1–14.
- Jespersen, Otto. 1949. *A modern English grammar on historical principles. Part 7*. Copenhagen: E. Munksgaard.
- Kingston, John. 2007. The phonetics–phonology interface. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 401–34. Cambridge: Cambridge University Press.

- Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36, 15–43.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A basic introduction*. New York: Oxford University Press.
- Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam and Philadelphia: John Benjamins.
- Lorenz, David. Forthcoming. Converging variations and the emergence of horizontal links: *to*-contraction in American English. In Sommerer & Smirnova (eds.).
- Mitchell, Bruce. 1985. *Old English syntax*. Oxford: Oxford University Press.
- Mustanoja, Tauno F. 1960. *A Middle English syntax*, part I. Helsinki: Société Néophilologique.
- OED = Oxford English Dictionary*. 1989. 2nd edn. Oxford: Oxford University Press.
- Payne, John & Rodney Huddleston. 2002. Nouns and noun phrases. In Huddleston & Pullum, 323–523.
- PPCEME = Anthony Kroch, Beatrice Santorini & Lauren Delfs. 2004. *The Penn–Helsinki Parsed Corpus of Early Modern English* (PPCEME). University of Pennsylvania.
- PPCME = Anthony Kroch & Ann Taylor. 2000. *The Penn–Helsinki Parsed Corpus of Middle English* (PPCME2). University of Pennsylvania.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. Harlow: Longman.
- R Core Team 2018. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. www.R-project.org
- Rissanen, Matti. 1967. *The uses of one in Old and Early Middle English*. Helsinki: Modern Language Society.
- Ropers, Karl. 1918. *Zur Syntax und Stilistik des Pronominalgebrauchs bei Aelfric*. Dissertation, University of Kiel.
- Sahlin, Elisabeth. 1979. *Some and any in spoken and written English*. Uppsala: Acta Universitatis Upsaliensis.
- Selkirk, Elisabeth. 1996. The prosodic structure of function words. In James L. Morgan & Katherine Demuth (eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 187–214. Mahwah, NJ: Lawrence Erlbaum Associates.
- Smirnova, Elena & Lotte Sommerer. Forthcoming. The nature of the node and the network: Open questions in Diachronic Construction Grammar. In Sommerer & Smirnova (eds.),
- Sommerer, Lotte. 2012. Investigating article development in Old English: About categorization, gradualness and constructions. *Folia Linguistica Historica* 33, 175–213.
- Sommerer, Lotte 2015. The influence of constructions in grammaticalization: Revisiting category emergence and the development of the definite article in English. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic Construction Grammar*, 107–33. Amsterdam: John Benjamins.
- Sommerer, Lotte. 2018. *Article emergence in Old English: A constructionalist perspective*. Berlin: De Gruyter Mouton.
- Sommerer, Lotte. Forthcoming. Constructionalization, constructional competition and constructional death: Investigating the demise of Old English POSS DEM constructions. In Sommerer & Smirnova (eds.).
- Sommerer, Lotte & Elena Smirnova (eds.). Forthcoming. *Nodes and networks: Advances in Diachronic Construction Grammar*. Amsterdam and Philadelphia: John Benjamins.
- Steels, Luc (ed.). 2011. *Design patterns in fluid Construction Grammar*. Amsterdam: John Benjamins.
- Süsskand, Peter. 1935. *Geschichte des unbestimmten Artikels im Alt- und Frühmittelenglischen*. Halle: Niemeyer.

- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Torrent, Tiago T. 2015. On the relation between inheritance and change: The construction network reconfiguration hypothesis. In Barðdal, Smirnova, Sommerer & Gildea (eds.), 173–212.
- Traugott, Elizabeth C. 1992. Syntax. In Richard M. Hogg (ed.), *The Cambridge history of the English language*, vol. 1: *From the beginnings to 1066*, 168–289. Cambridge: Cambridge University Press.
- Traugott, Elizabeth C. & Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- Trousdale, Graeme. 2014. On the relationship between grammaticalization and constructionalization. *Folia Linguistica* 48(2), 557–78.
- Trousdale, Graeme & Nikolas Gisborne (eds.). 2008. *Constructional approaches to English grammar*. Berlin: Mouton de Gruyter.
- Van Goethem, Kristel, Muriel Norde, Evie Coussé & Gudrun Vanderbauwhede. 2018. Category change from a constructional perspective: Introduction. In Van Goethem, Norde, Coussé & Vanderbauwhede (eds.), 3–12.
- Van Goethem, Kristel, Muriel Norde, Evie Coussé & Gudrun Vanderbauwhede (eds.), *Category change from a constructional perspective*. Amsterdam and Philadelphia: John Benjamins.
- Wülfing, Johann E. 1901 [1894]. *Die Syntax in den Werken Alfreds des Großen*. Bonn: Hanstein.
- Zehentner, Eva. 2019. *Competition in language change: The rise of the English dative alternation*. Berlin: De Gruyter Mouton.
- Zehentner, Eva & Elizabeth C. Traugott. Forthcoming. Constructional networks and the development of benefactive ditransitives. In Sommerer & Smirnova (eds.).

Appendix

Table A1

Formula: SOME ~ PERIOD					
AIC: 3385.8	Obs: 25457	C: 0.673			
Predictors	Levels	Estimate	Std. error	Z value	p-value
	(Intercept)	−4.35109	0.06299	−69.074	<2e-16***
PERIOD	Contrast level				
(Helmert contrasts)	M2	−0.30647	0.13225	−2.317	0.0205*
	M3	−0.12394	0.06171	−2.008	0.0446*
	M4	−0.13161	0.06118	−2.151	0.0315*
	E1	0.18166	0.03454	5.259	1.45e-07***
	E2	0.15161	0.02451	6.185	6.20e-10***
	E3	0.09471	0.02290	4.136	3.53e-05***

Table A2

Formula: A(N) ~ PERIOD					
AIC: 28142	Obs: 25457	C: 0.741			
Predictors	Levels	Estimate	Std. error	Z value	p-value
PERIOD (Helmert contrasts)	(Intercept)	-0.906284	0.017653	-51.338	<2e-16***
	Contrast level				
	M2	0.313440	0.046043	6.808	9.93e-12***
	M3	0.505828	0.017387	29.092	<2e-16***
	M4	0.371658	0.011687	31.801	<2e-16***
	E1	0.271549	0.009707	27.974	<2e-16***
	E2	0.179024	0.007359	24.327	<2e-16***
	E3	0.177442	0.006703	26.473	<2e-16***

Table A3

Formula: NOUN ~ PERIOD					
AIC: 28666	Obs: 25457	C: 0.741			
Predictors	Levels	Estimate	Std. error	Z value	p-value
PERIOD (Helmert contrasts)	(Intercept)	0.815423	0.017054	47.81	<2e-16***
	Contrast level				
	M2	-0.239082	0.043237	-5.53	3.21e-08***
	M3	-0.466654	0.016538	-28.22	<2e-16***
	M4	-0.348993	0.011367	-30.70	<2e-16***
	E1	-0.269667	0.009556	-28.22	<2e-16***
	E2	-0.180455	0.007262	-24.85	<2e-16***
	E3	-0.177290	0.006684	-26.52	<2e-16***

Table A4

Formula: MASS ~ PERIOD					
AIC: 388.63	Obs: 309	C: 0.648			
Predictors	Levels	Estimate	Std. error	Z value	p-value
PERIOD (Helmert contrasts)	(Intercept)	-0.84892	0.14582	-5.822	5.83e-09***
	Contrast level				
	M2	0.32696	0.31501	1.038	0.2993
	M3	0.22723	0.14542	1.563	0.1181
	M4	-0.08276	0.15139	-0.547	0.5846
	E1	0.14303	0.07407	1.931	0.0535
	E2	0.12523	0.05191	2.412	0.0159*
	E3	0.02083	0.04932	0.422	0.6727

Table A5

Formula: ARTICLE ~ PERIOD

AIC: 265.24	Obs: 262	C: 0.862			
Predictors	Levels	Estimate	Std. error	Z value	<i>p</i> -value
	(Intercept)	-0.79112	0.22765	-3.475	0.00051***
PERIOD	Contrast level				
(Helmert contrasts)	M2	0.29389	0.66247	0.444	0.65731
	M3	-0.03516	0.24539	-0.143	0.88606
	M4	0.13805	0.15514	0.890	0.37354
	E1	0.25792	0.11047	2.335	0.01956*
	E2	0.50402	0.09633	5.232	1.67e-07***
	E3	0.24738	0.05483	4.512	6.42e-06***