Lauren Fonteyn and Stefan Hartmann*

Usage-based perspectives on diachronic morphology: A mixed-methods approach towards English *ing*-nominals

Abstract: This paper illustrates how different methodological approaches can be combined to reveal complex patterns of constructional variation and change in the diachronic development of English *ing*-nominals. More specifically, we argue that approaching the data from a schema-based (rather than morpheme-based) perspective shows that nominal gerunds in English, from the sixteenth to the nineteenth century, have undergone a semantic drift towards more "nouny" construal variants. This hypothesis is supported not only by raw frequency counts, but also by association measures and by a detailed analysis of hapax legomena.

Keywords: word-formation change, usage-based linguistics, collostructional analysis, historical linguistics, corpus linguistics

1 Introduction

This paper investigates the diachronic development of English deverbal nominalizations in *-ing* and illustrates how combining different methodological approaches can reveal a diachronic semantic shift from more "verby" to increasingly nominal readings. We discuss how (i) raw frequency counts, (ii) productivity measures, and (iii) explorative methods such as collostructional analysis can be combined to track diachronic changes not only in the *formal* features of *ing*-nominalizations, but also developments pertaining to the *functional* side of the word-formation pattern. In Section 2, we will first give a brief overview of previous research on the English gerund. Section 3 discusses how different theoretical and methodological perspectives, i. e. a morpheme-based vs. a schema-based account, can lead to very different results in the analysis of morphological data. Section 4 presents a diachronic corpus analysis of nominal gerunds on the basis of the Penn-Helsinki corpora. Combining an explorative analysis of the most frequent tokens and the tokens most strongly attracted to the individual corpus periods according to diachronic distinctive collexeme analysis with a systematic analysis of all tokens in the corpus, which have been handcoded for their reading variants, we show that non-eventive *ing*-nouns become significantly more frequent over time. In addition, a closer look at the hapax legomena in the dataset suggests that non-eventive readings become much more prominent even among newly-coined derivatives. Section 5 offers a brief conclusion, in which we argue that the combination of different methods allows for identifying patterns of change that might otherwise remain undetected.

2 The English gerund

In English, the most commonly studied type of deverbal nominalizations in *-ing* is the so-called English gerund, a pattern in which the 'situational' semantics of the base verb is retained. As such, even though gerunds have nominal external distribution (functioning as subjects, direct objects, or prepositional complements), they are essentially still constructions that serve to refer to an event or a state, as in (1). In this function, they are distinct from prototypical nominals, which serve to refer to *objects* (cf. e. g. Langacker 2008: 34).

---

*Corresponding author: Stefan Hartmann, German Department, University of Hamburg, Von-Melle-Park 6, 20146 Hamburg, Germany, E-mail: stefan.hartmann@uni-hamburg.de

Lauren Fonteyn, Department of Linguistics, KU Leuven, Leuven, Belgium

(1)   *Labour will ban **the killing of foxes** and encourage **the killing of human foetuses**.* (BNC)

There is also a fairly large group of *ing*-nominalizations that do have a more concrete denotation (Hopper and Thompson 1984; Grimshaw 1990; Koptjevskaja-Tamm 1993; Alexiadou 2001; Bauer et al. 2013; Andersen 2007). These 'concrete' *ing*-nominals impose a denotational shift on the base verb (Jespersen 1946: 87; Marchand 1969: 303; Urdang et al. 1982: 185; Grimshaw 1990) and commonly refer to the resulting (abstract) object of an action, as in (2), or something more generally connected with the verbal action, as in (3):

(2)   (a) *In June 1845 he wants to buy **a painting of a bear** for his room and entitle it 'Portrait of Gustave Flaubert'* (BNC)
       (b) *(…) it is an undeniable fact that in the 1990s most hitherto self-proclaimed socialist or communist countries are rapidly coming to an **understanding** with transnational capitalism (…)* (BNC)

(3)   *The linen area was unheated, and though they were well protected from the ravages of the wind she was aware of **her damp clothing** chilling her skin (…)* (BNC)

As noted by Panagl (1987), there is a cross-linguistically well-attested diachronic relation between nominalizations referring to situations (or 'nouns of action') and nominalizations referring to more concrete object-like entities associated with a situation:

> (…) nouns of action generally show an inherent tendency toward categorial change of meaning. This development (…) tends to proceed through the level of resultative noun (nomen acti) and in many cases reaches the level of concrete noun. (Panagl 1987: 146)

The drift described by Panagl (1987) concerns a cross-linguistic tendency towards particular semantic changes of a nominalization schema, in particular the loss of semantic features or the extension of these features through metaphor or metonymy (e. g. Himmelmann 2004). Aronoff (1976: 19, 43) regards such developments as instances of semantic drift in word formation patterns. Like Panagl (1987), Aronoff observes that action nominalizations can acquire additional 'non-regular' meanings through semantic specialization, so that for example the word *transmission* underwent a meaning change from 'the act of transmitting' to 'the part of a car that transmits the power of the engine to the wheels'. This is similar to the development of certain deverbal *ing*-nominals, which acquired specialized (or metonymically related) meanings revolving around the object, source, and result of the verb they were based on, e. g. *painting*.

Interestingly, some scholars have suggested that nominalizations in -*ing* "characteristically display eventive [i. e. situational] readings and are less prone to (…) semantic drift" than other nominalization strategies in English such as zero-affixation (Bauer et al. 2013: 208), or related nominalization strategies in different languages (Demske 2002). In an elaborate corpus-based study of German *ung*-nominalizations, Demske (2000, 2002) argues that, for instance, the *ung*-nominalization *Heizung* in Present-Day German can only have reference to a concrete object (i. e. a 'heating installation') and can no longer be used to refer to the 'process of heating' (also see Hartmann 2016: 264f.). Similarly, *Lesung* (from *lesen* '(to) read') in Present-Day German does not refer to the 'process of reading', but rather to a kind of formal event or lecture. In older stages of German, however, *ung*-nominals did quite commonly refer to the process denoted by the base verb. The change that occurred in the system of German *ung*-nominalizations can be explained as follows: *ung*-nominals, which in earlier stages of German were productively used to refer to situations, gradually acquired new meanings (cf. Aronoff 1976; Sauer 2004: 1626). These new meanings were acquired through the semantic extension mechanisms of metaphor and metonymy (cf. Panther and Thornburg 2001; Booij 2010: 78), going from the default prototypical meaning of 'situation' or 'event' to derived meanings such as concrete locations, objects, and even agents involved in these situations (Hartmann 2014a, 2014b, 2014c). Subsequently, then, the word formation pattern in -*ung* was reanalyzed and the new meaning

variants (e. g. location, object, person associated with V-*ung*) were "abstracted away from the lexicalized word-formation products and transferred to the word-formation pattern itself" (Hartmann 2014c: 135; Scherer 2006). From the seventeenth century onwards, *ung*-nominals denoting situations started to decline (Hartmann 2014b), resulting in increased overall 'object-like' denotation of the *ung*-nominalization pattern. In other words, the probability that a word formation product in -*ung* had non-eventive semantics increased over time. Demske (2002: 68) refers to the diachronic development of *ung*-nominalization as a "nominalization process with 'nominalization' taken literally". Importantly, then, it is said that "no such nominalization process is attested with *ing*-nouns in the history of English" (Demske 2002: 88).

We will argue, however, that the nominal *ing*-pattern does undergo a semantic shift, but detecting this shift in corpus data is not a trivial matter. More specifically, we will argue that even a brief exploration of historical corpus data indicates that the most frequent members of the category of *ing*-nominals in fact have non-eventive (rather than eventive) meanings. Combining these raw frequency counts using diachronic distinctive collexeme analysis (Hilpert 2006), one can find suggestive evidence for a potential semantic shift in nominal *ing*-constructions. The existence for such a shift, then, will be further substantiated by means of a more elaborate manual analysis of the data set. Taking a closer look at the hapax legomena, i. e. the formations only attested once in the corpus, we can also show that nominal gerunds with non-eventive meanings are not merely the product of lexicalization processes (such as e. g. *saying* 'process of saying sth.' > 'proverb', *feeling* 'state/process of feeling sth.' > 'emotion') but become more prevalent in occasional formations as well. At the same time, we will point out that the results of this semantic analysis greatly differ depending on whether we take a morpheme-based perspective, including all deverbal nominalizations in -*ing*, or a schema-based perspective, limiting the scope to *ing*-nominalizations with nominal internal syntax. Arriving at a more accurate and nuanced picture of the development of English *ing*-nominals thus requires a "cumulative" approach to the available data, in which several perspectives and methods complement each other, rather than a selective one.

3 Morpheme-based or schema-based view?

One of the most intriguing aspects of English *ing*-nominals is that, unlike German *ung*-nominals, they can occur in a nominal as well as in a more clause-like construction. Examples (1)-(3) above illustrate the so-called nominal gerund, while (4) is an example for the more clause-like use of *ing*-forms.

(4)   *Charles does not hunt because he enjoys **killing foxes**.* (BNC) (compare: *the killing of foxes*)

The *ing*-form in (4) is a so-called verbal gerund.[1] Like other *ing*-nominals, the verbal gerund shows signs of external recategorization from a verbal to a nominal structure in that it adopts the distribution of a noun phrase: verbal gerunds can function as the subject or object of a clause, or as a prepositional complement. Yet, the structural differences between the verbal gerund in (4) and the *ing*-nominals discussed above are obvious: the *ing*-nominals display the internal syntax of a noun phrase, taking determiners, adjectives, and an *of*-phrase to express their subject or direct object, whereas verbal gerunds have the internal syntax of a clause, with clausally-realized participants and adverbial modification. The verbal gerund – unlike the *ing*-nominal with noun phrase-like internal structure – never displays non-eventive semantics: it is always used to refer to a situation or event.

---

1 Note that the notion of "verbal gerund" does not include the participial use of *ing*-forms, as in *She is writing a paper*. Although it is often assumed that gerunds and participles are morphologically identical, De Smet (2010) points out that language users do seem to make a difference between them; for instance, there is evidence that the phonological variation between /in/ and /iŋ/ is conditioned grammatically, reflecting the distinction between gerunds and participles fairly well.

Given the differences between nominal and verbal gerunds, the question emerges whether they should be treated as two separate constructions or as one single construction used in different environments (cf. e. g. Alexiadou 2013). From a morpheme-based perspective, the latter analysis would seem natural. On this view, the morpheme is the central unit of linguistic morphology, and morphology itself is seen as "the set of principles for combining morphemes into words." (Booij 2012: 8) From a usage-based and constructionist point of view, by contrast, a schema-based view seems more appropriate.[2] On this view, (at least) two different constructional schemas sharing the same morpheme -*ing* can be posited for the English gerund, one for the nominal gerund ([V-*ing*]$_N$]) and one for the verbal gerund ([[V-ing]$_V$ + DO]$_{NP}$), given their clear differences in form, function, and distribution.

As we will illustrate in Section 4, studies of the semantics of English *ing*-nominals can have varying results based on whether the researcher takes a morpheme-based perspective, including all deverbal nominalizations in -*ing* regardless of whether they are nominal *ing*-constructions or verbal gerunds, or a schema-based perspective, which makes a difference between nominal and verbal gerunds. In our study, we will, by and large, take the latter view and limit our scope to constructions aligning with the nominal *ing*-construction schema. However, we will additionally also take a look at verbal *and* nominal gerunds from a morpheme-based perspective, and we will argue that combining both perspectives can contribute significantly to understanding the diachronic development of the *ing*-construction.

4 Detecting semantic shift: Exploring the data using raw frequencies, productivity measures, and diachronic distinctive collexeme analysis

To investigate whether the nominal *ing*-schema was subject to semantic drift, we extracted the entire set of nominal *ing*-forms from the PPCEME and PPCMBE (Kroch et al. 2004; Kroch et al. 2010), i. e. all deverbal forms in -*ing* with nominal distribution taking nominal internal syntax, including constructions with object-like denotation (*the old building*) as well as those with eventive or situational denotation (*the coming of Christ*). The data are divided into 4 subsequent time slices, two for Early Modern English (E) and two for Late Modern English (L), i. e. period E1 (1570–1640), E2 (1640–1710), L1 (1710–1780) and L2 (1780–1850). All attestations were lemmatized manually. The token frequencies for each time slice are presented in Table 1.

Table 1: Token frequencies of nominal gerunds in Early and Late Modern English.

	E1 1570–1640	E2 1640–1710	L1 1710–1780	L2 1780–1850
ing-nominal	1,140	720	510	386

As pointed out by e. g. Demske (2002), Sleeman and Brito (2010), and Hartmann (2014b, 2014c), eventive and non-eventive (object-like) readings are often metaphorically or metonymically related. A consequence of this close relation between eventive and non-eventive readings is that, in some cases, the deverbal nominalization is ambiguous and cannot be decisively be interpreted as being one or the other (Sleeman and Brito 2010: 200). In most cases, however, the lexicogrammatical context in which the nominalization occurs helps to arrive at the precise interpretation (Bauer et al. 2013). For instance, the eventive reading of nominalizations is most easily available when the nominalization displays the full argument structure of the corresponding verb (Borer 2012), as in *Mr Shamir's building of new settlements in the occupied territories*.

---

2 Cf. e. g. Hilpert 2014: 76 f. for arguments in favor of viewing word-formation patterns as morphological constructions rather than "rules").

In other cases, contextual elements like qualitative adjectives can aid in arriving at an object-like reading, as properties such as 'old' (as in *Did you ever notice the old building at Forter?*) or' damp' (as in *she was aware of her damp clothing chilling her skin*) are not normally ascribed to events or situations.

In order to gain a first impression of the data, it can be helpful to take a look at the most frequent *ing*-nominals in each period. Table 2 shows the 10 most frequent items for each of the four time slices. From a usage-based perspective, highly frequent exemplars can be assumed to contribute significantly to the meaning of the more abstract schema that is abstracted away from the individual instances, although the precise influence of high token frequency is still a matter of considerable debate (cf. e. g. Bybee 2010; Blumenthal-Dramé 2012).

Table 2: The ten most token-frequent ing-nominals in each of the four time periods covered by PPCEME and PPCMBE. Nominals with non-eventive meaning are displayed in boldface, nominals that allow for both eventive and non-eventive readings are underlined.

E2		E3		L1		L2	
beginning	99	**being**	58	**beginning**	85	meeting	64
coming	66	coming	55	**meaning**	31	teaching	43
learning	62	**beginning**	50	**being**	21	**meaning**	26
offering	45	**saying**	35	**understanding**	19	**beginning**	16
understanding	44	**understanding**	31	**rigging**	19	writing	16
blessing	43	**building**	28	meeting	18	building	16
lodging	39	learning	26	coming	17	passing	14
writing	37	**blessing**	26	learning	17	**drawing**	14
meaning	34	meeting	22	**coating**	17	**understanding**	13
saying	29	**lodging**	19	**blessing**	14	coming	13

At least for Early Modern and Late Modern English, the claim that nominalizations in -*ing* characteristically display eventive meanings seems problematic. As the results listed in Table 2 show, the exemplars of the nominal *ing*-construction are in fact commonly associated with non-eventive meanings, as in (5), or at least often allow non-eventive readings alongside eventive ones, as in (6).

(5)   a. (...) *in **the beginning of the fourth century** the Saxons were not alone on the ocean* (1799, PPCMBE)
       b. *His brain was clouded; **his understanding** was in a maze;* (1805, PPCMBE)
       c. (...) *this man will die in Gaol, or on a dunghill, without **a being** to lament him or a wife assist him* (1808, PPCMBE)
       d. *Teachers should also explain, by word of mouth, **the meaning of every lesson**, what it contains* (1743, PPCMBE)

(6)   a. *Cain brought of the fruite of the ground, **an offering vnto the Lord**.* (1611, PPCEME) [ambiguous]
       b. *It is a handsome building* (180x, PPCMBE)

However, the frequencies in Table 2 do not suggest that the nominal *ing*-construction has shifted towards more concrete denotations over time. They merely show that some of the most frequent nominal *ing*-forms have non-eventive meanings. If we only take the most frequent items into account, the proportion of non-eventive readings does not seem to change.

If we take a slightly different perspective in exploring the data, focusing not on the *ing*-forms that are the most frequent but on those which are unusually or *distinctively* frequent in a particular period, we might in fact find indications that a semantic shift of some sort did occur. This can be done using Diachronic Distinctive Collexeme Analysis (henceforth DDCA; Hilpert 2006), which is a diachronic adaptation of Gries and Stefanowitsch's (2004) distinctive collexeme analysis. Essentially, the diachronic adaptation of this

method can be used for comparing the attraction of lexical items to a single construction over sequential periods of time. In other words, DDCA compares the frequency of each lexical item in one time period to its frequency in other investigated periods, and to the frequency of the other lexical items occurring in the construction during the same period. Hence, DDCA can abstract away from items that frequently occur in every period under investigation and thus helps to detect or highlight items that are *distinctive* for a particular period. Items are judged as "distinctive" if they have a higher relative frequency in one period than in the others and if they are significantly more frequent than expected in a particular period.

DDCA has previously been adopted in a few studies looking to detect semantic changes of (grammaticalizing) verbal constructions by investigating potential changes in their collocational preferences. Hilpert (2006: 248ff.) shows the benefits of adopting DDCA using the development of the auxiliary *shall*, and points out that merely looking at absolute frequencies of collocations "sometimes obscure[s] differences that hold across sets of [historical] data". If one merely considers the raw frequencies of the verbs that occur with the auxiliary *shall*, one finds that the general verbs *be, have, do,* etc. are among the most frequent complements of shall in each period, and no striking differences between periods can be detected. Using a DDCA for studying the development of *shall*, however, allows "the identification of a semantic change that follows the trajectory of subjectification" (Hilpert 2006: 254), and thus aids the analyst in looking for patterns of change.[3]

The application of DDCA remains, however, controversial. In particular, DDCA has been criticized for treating historical periods as "quasi-linguistic categories" (Stefanowitsch 2006: 257). More specifically, Stefanowitsch (2006) points out that historical periods are neither psychological nor natural categories. The fact that historical periods are not psychological categories entails that, unlike synchronic DCA, DDCA cannot be taken to reflect psychological associations in the mind of the language user (cf. Coussé 2014). Therefore, "a claim to psychological reality cannot be maintained under a historical application of collostructional analysis" (Hilpert 2008: 44). However, this has never been the goal of DDCA. While DDCA adopts the *method* of distinctive collexeme analysis, the results of course have to be analyzed in a different way: not as psychologically realistic patterns of association and dissociation between lexical items and constructions, but rather as a "seismograph" of diachronic changes in these association patterns across generations of speakers. Stefanowitsch's second criticism is much more important, namely that period boundaries are necessarily arbitrary and that the results of DDCA strongly depend on how one carves up the data. However, this problem is not specific to DDCA, and as in the case of other quantitative methods, it can be encountered – but of course never entirely solved – by using maximally balanced corpora. As the individual periods of the Penn corpora are well-balanced in terms of token frequencies and text types, among other factors, it seems legitimate to operationalize the time slices for which the corpus is balanced as admittedly arbitrary, but heuristically valuable and methodologically justified categories in DDCA.

Taking into account the benefits of and objections to adopting DDCA to assess semantic changes within a construction, we suggest that it can serve as a valuable *explorative* tool to detect semantic changes. Consider the results presented in Table 3, which provides a ranked list of the ten most distinctive *ing*-forms for each individual period, followed by their absolute frequency and their "distinctiveness"[4]:

The results presented in Table 3 suggest an interesting tendency. In period E2, the ten most distinctive *ing*-forms mainly have eventive readings, as illustrated in (7):

(7)   a. *Also it is knowen most certainly what great good is done in* **curing of such rotten gums and sore mouthes**. (1596, PPCEME)
    b. *As for* **sending of Letters to my Lord of Northumberland**, *(...) all is false.* (1571, PPCEME)
    c. *I (...) was at moch charge in* **keping of my horse**, *my conies.* (c? 1602, PPCEME)
    d. *she suffred him (...) to plowe and* **gathering of stickes** *and such like.* (1602, PPCEME)

---

3 For other applications of DDCA, see Hilpert (2008) and Coussé (2014).
4 As usual in collostructional analysis, the collostructional strength, or "distinctiveness", is calculated by log10-transforming the p-value of a Fisher Exact Test.

Table 3: Results DDCA of nominal ing-formations.

E2	N	Coll.	E3	N	Coll.	L1	N	Coll.	L2	N	Coll.
offering	45	7.93	**wedding**	15	4.41	rigging	19	14.38	teaching	43	24.87
dealing	25	6.62	**breeding**	19	4.19	coating	17	12.87	**meeting**	64	20.12
curing	14	4.34	**saying**	35	3.92	beginning	85	9.58	**drawing**	14	7.42
moving	14	3.65	enlarging	6	3.52	gilding	6	4.54	passing	14	5.97
sending	15	3.42	coming	55	2.55	**firing**	8	3.65	**reasoning**	13	5.37
earning	9	3.41	distilling	4	2.3	**collecting**	4	3.03	**feeling**	13	5.37
keeping	16	2.48	drinking	5	1.82	**hailing**	4	3.03	**swelling**	10	3.66
lodging	39	2.3	arming	3	1.76	**concealing**	4	2.39	**opening**	9	3.58
gathering	8	2.28	charring	3	1.76	**fasting**	4	2.39	printing	5	3.43
publishing	6	2.27	founding	3	1.76	**planting**	4	2.39	speaking	6	2.89

Three of the most distinctive *ing*-formations in E2, however, do not (solely) function as nominalizations of actions or events, but denote physical objects. As illustrated in (6) above, the lexeme *offering* in a few cases can still be interpreted as the act of offering something, but at the same time it also allows for a resulting object reading (i. e. 'the thing which is offered'). Similarly, *earning* in (8) in all cases signifies the thing that is earned and *lodging* (9) denotes 'a place of residence', rather than '(an instance of) to lodge (i. e. 'reside')':

(8)   (...) *this is* **the choisest and best earning which can possible be made by any Hus-wife.** (1615, PPCEME)

(9)   *At night within two miles of* **my lodging**, *I was faine to wade ouer the Riuer of Annan in Scotland* (1630, PPCEME)

In period E3, we again find a large number of actional/eventive *ing*-forms in the list of most distinctive collexemes, such as *coming* and *distilling*:

(10)   *And St. Jude out of an ancient prophecy of Enoch expresly mentions this as one reason of* **the coming of the Lord**. (a1671, PPCEME)

(11)   *Whereas it is found by Experience that* **the drawing distilling and makeing of Brandies or Spiritts from Malt** *is one Cause of the present dearnesse of Corne* (1690, PPCEME)

Interestingly, the three top-listed *ing*-forms – *wedding, breeding* and *saying* – again do not have actional or eventive semantics. Rather, *wedding* (12), refers to a ceremony or festivity, *breeding* (13) refers to the abstract concept of 'upbringing' or 'education', and *saying* (14) refers to the thing that is said:

(12)   *Being thus met at* **the Wedding**, *the Bride came to distribute her Favours amongst her friends.* (1684–1687, PPCEME)

(13)   *I mean those who are of better* **breeding**, *because they pretend to understand the laws of behaviour and the decencies of conversation better than other men.* (a1671, PPCEME)

(14)   *I end this with* **the saying of Laelius in Cicero**: *Amicitia non debet consequi utilitatem, sed amicitiam utilitas.* (1662, PPCEME)

In the first Late Modern period L1, the importance of non-eventive *ing*-forms still increases, with *rigging, coating, beginning* and *gilding* being the most distinctive collexemes:

(15) *(...) they unbent the split Topsail from the Yards, and found* **the running Rigging** *much damaged* (1744, PPCMBE)

(16) *(...) because, when the water is poured out of the phial, there still remains* **a thin coating of the fluid,** *which might be thought to contain the power of giving a shock.* (1769, PPCMBE)

In period L2, finally, as many as six of the ten most distinctive *ing*-forms have non-eventive denotation:

(17) *The brewery scheme is quite at an end: at* **a meeting of the subscribers** *last week it was by general, and I believe very hearty, consent dissolved.* (180X, PPCMBE)

(18) *However* **the reasoning of man** *might have led him to such a hope;* (1793, PPCMBE)

(19) *I should feel almost at home even on board a ship, were it not for* **an anxious feeling** *....* (1836, PPCMBE)

(20) *I obtained many premiums in the different classes, and once the head premium for* **my drawing of the Ariadne**, *the well-known fine antique.* (1826, PPCMBE)

(21) *It is no unusual thing for* **a swelling in the roof of the mouth** *to supervene* (1814, PPCMBE)

(22) ***(...)*** *buttons all the way down the coat, but only one at the waist buttoned; the hilt of the sword through* **the opening of the skirt**; (1826, PPCMBE)

In sum, the results of the DDCA suggest that it is at least worth pursuing the idea that the nominal *ing*-construction became increasingly associated with non-eventive readings. To this end, all tokens in the corpus have been handcoded in context for their semantics, using a coarse-grained distinction between eventive and non-eventive reading variants. The results corroborate the hypothesis that non-eventive meanings, over time, have become more prevalent in the nominal *ing*-construction than eventive reading variants. While the frequency of the latter has dropped considerably, non-eventive *ing*-nominals remain fairly frequent (cf. Figure 1).

Harking back to the issues discussed in Section 3, viz. the importance of the perspective taken on the data, it should be emphasized that very different results would be obtained if all *ing*-formations, regardless of whether they are nominal or verbal, were taken into account. The right panel of Figure 2 illustrates that a purely morpheme-based view would largely obscure the diachronic changes. If nominal and verbal gerunds are lumped together, the eventive reading seems to remain the predominant one by far.

While the morpheme-based perspective obscures the increasing prominence of non-eventive *ing*-nominals, leaving the verbal gerund entirely out of the equation would also skew the findings. The mere fact that the development of nominal *ing*-constructions in English is similar to that of German *ung*-nominals does not mean that the English *ing*-nominal *in general* undergoes a denotational shift of the same scale as the one that affected German *ung*-nominals. Non-eventive *ing*-nominals in English are not nearly as prominent as they are in German, and even in present-day English their overall frequency is only modest compared to that of verbal gerunds. The crucial difference between German *ung*-nominals and English *ing*-nominals is that they are part of a different language network and are surrounded by different 'neighboring' constructions. In German, the event-referring uses of *ung*-nominals gradually came to be replaced by nominalized infinitives, which in German (and other Germanic languages such as Dutch) cover a different, and somewhat wider, functional domain than in English (Hartmann 2014a, 2014b, 2014c). In English, the domain of 'event reference' is dominated by the verbal gerund, which is syntactically different from
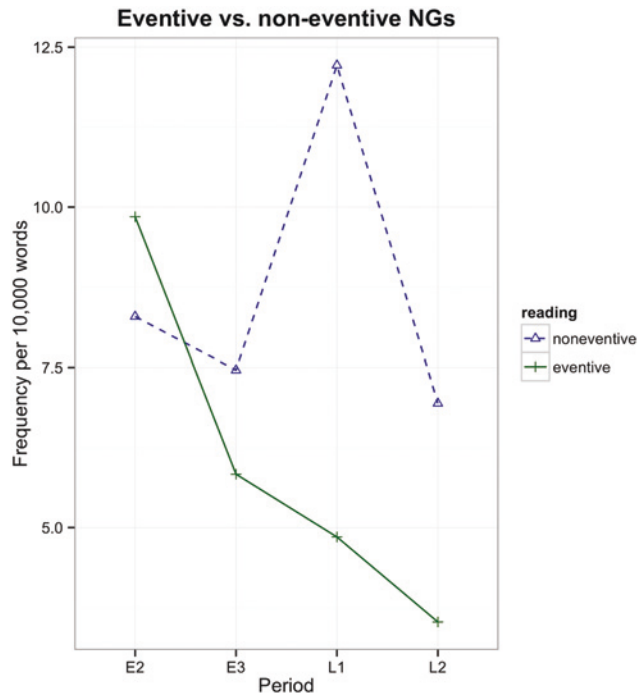
Eventive vs. non-eventive NGs

Figure 1: Normalized frequencies of eventive vs. non-eventive ing-nominals. The Y-axis indicates the normalized frequency of the construction per 10,000 words (calculated by [number of occurrences]/[totalnumber of words in the subcorpus]*10,000).
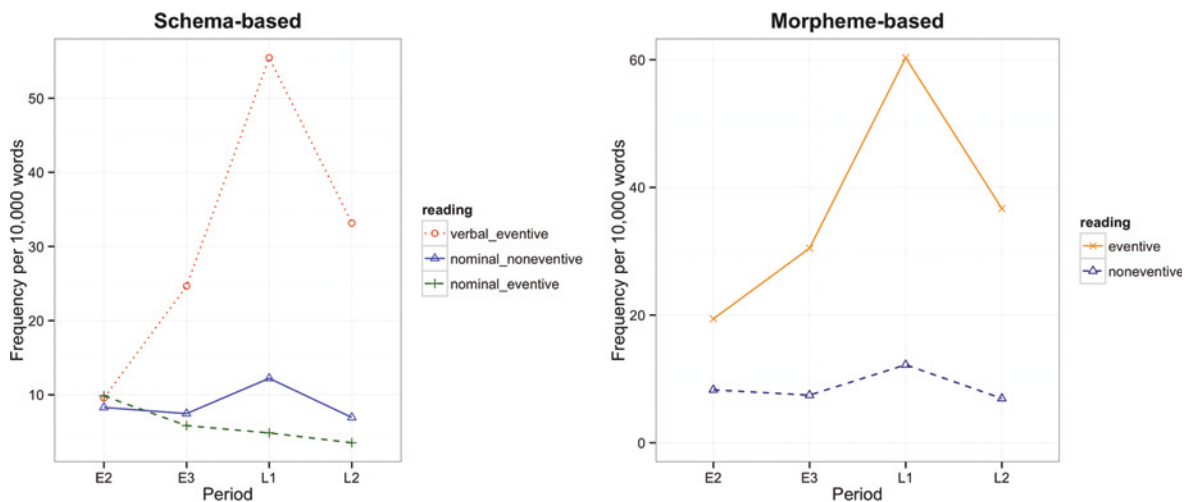
Schema-based

Morpheme-based

Figure 2: Normalized frequencies of verbal vs. nominal-eventive vs. nominal-noneventive ing-formations; normalized frequencies of eventive and noneventive ing-formations, lumping together nominal and verbal gerunds.

nominal *ing*-constructions, but is both historically and phonologically very closely associated with them. As such, the formal identity between the *ing*-morpheme in nominal and verbal *ing*-nominals can be assumed to evoke a close association between the two types, causing semantic attraction between them. This might explain why eventive readings are still much more common in English nominal gerunds than in their German or Dutch counterparts, even in productive uses of the nominal *ing*-construction.

This leads us to the next issue: Any analysis relying solely on token frequencies can potentially be misleading. As lexicalization and high token frequency tend to accompany each other (cf. e. g. Hartmann 2016), it is perfectly possible that the results shown in Figure 1 above are due to a few very frequent

lexicalized tokens which have assumed non-eventive readings over time. Therefore, it makes sense to check whether non-eventive readings become more frequent in occasionally coined *ing*-nominals as well. To do so, we can take a closer look at the so-called hapax legomena, i. e. the lemmas occuring only once in the corpus. Hapaxes have played a key role both in approaches comparing the productivity of different word-formation patterns (e. g. Baayen and Lieber 1991) and in attempts to compare the productivity of one pattern at different points in time (e. g. Scherer 2006). More precisely, these studies have made use of a productivity measure proposed by Baayen (1992, 1994, 2009, among others), namely potential productivity. The potential productivity of a construction in a given corpus can be obtained by dividing the number of hapax legomena belonging to that construction by the total number of tokens instantiating the construction in question ($P = V_1/N$). For instance, in period E2 of our corpus data, 62 out of 1140 *ing*-nominals are hapaxes, which yields a productivity value of $P = 62 / 1140 = 0.054$. As Figure 3 shows, the potential productivity of the English nominal gerund construction does not seem to change considerably over the time period under investigation.[5] However, the proportion of hapaxes with a non-eventive reading variant increases considerably from period E3 to period L1. If we aggregate the hapaxes from periods E2 and E3 on the one hand and from periods L1 and L2 on the other, a Fisher Exact Test confirms that non-eventive hapaxes can be found highly significantly more often in the Late Modern English periods than in the Early Modern English ones (p<0.001, odds ratio = 11.79).
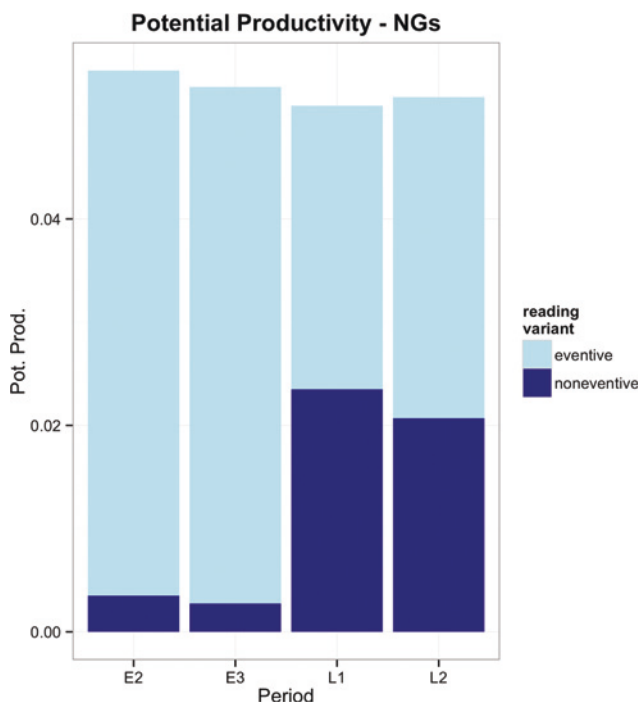

Figure 3: Hapax legomena in -ing in proportion to all tokens belonging to the nominal ing-construction in each period of PPCEME and PPCMBE (potential productivity), with relative frequencies of hapaxes exhibiting eventive vs. non-eventive readings. The productivity values for eventive and for non-eventive ing-nominals were calculated using the total number of nominal ing-formations as denominator. Thus, they add up to the overall productivity value of the construction.

---

5 It should be noted that comparing the potential productivity values is somewhat problematic in this case as the token frequency of *ing*-nominals decreases drastically over time (see Table 1). As Säily (2011: 125) points out, potential productivity is "dependent on the size of the corpus, both in terms of token frequency and the number of running words in the corpus". As Gaeta and Ricca (2006) point out, comparing productivity values with very different *N*s can be quite misleading. For our purposes, however, the proportion of eventive vs. non-eventive nominals among the hapax legomena is more important than the actual productivity values. Consequently, this issue can be neglected here.

5 Conclusion

Regarding the diachronic development of English *ing*-nominals, it has been claimed that no semantic "drift" from eventive to more concrete or object-like meanings can be observed (cf. Demske 2002: 88). However, by approaching the data from a variety of perspectives (morpheme-based vs. schema-based view, diachronic distinctive collexeme analysis, and semantic analysis of all tokens in aggregate as well as the hapax legomena), we have shown that non-eventive meanings do seem to become a more prominent part of the nominal *ing*-construction. In fact, in the final Late Modern period (L2, 1780–1850), the nominal *ing*-construction is most often used with non-eventive semantics. We have argued that these developments can only be detected if one takes a schema-based perspective, which is methodologically operationalized by treating verbal and nominal gerunds as instances of different constructions and therefore analyzing them separately. From a morpheme-based perspective, which tends to lump together nominal and verbal gerunds (unless nominal and verbal *-ing* are regarded as two different homophonic suffixes), the eventive reading seems to stay the predominant interpretation of *ing*-forms, and it even seems to become more prevalent over time. Yet, the morpheme-based perspective is crucial in understanding why the denotational shift of English *ing*-nominals from eventive to non-eventive meanings is only a modest one compared to the shift that affected, for instance, German *ung*-nominals. To detect and understand the actual developments, then, the analyst should realize that to choose is to lose; it is only by approaching the data from various perspectives that one will arrive at the most accurate and nuanced picture.

References

Alexiadou, Artemis. 2001. *Functional structure in nominals. Nominalization and ergativity*. Amsterdam & Philadelphia: John Benjamins.

Alexiadou, Artemis. 2013. Nominal vs. verbal -ing constructions and the development of the english progressive. *English Linguistics Research* 2(2). 126–140.

Andersen, Øivin. 2007. Deverbal nouns, lexicalization and syntactic change. *Nordic Journal of Linguistics* 30(1). 55–86.

Aronoff, Mark. 1976. *Word formation in generative grammar*, Vol. 1. (Linguistic Inquiry, Monographs). Cambridge: MIT Press.

Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford Reference guide to English morphology*. Oxford: Oxford University Press.

Blumenthal-Dramé, Alice. 2012. *Entrenchment in usage-based theories: What corpus data do and do not reveal about the mind*. Vol. 83. (Topics in English Linguistics). Berlin & New York: De Gruyter.

Booij, Geert E. 2010. *Construction morphology*. Oxford: Oxford University Press.

Booij, Geert E. 2012. *The grammar of words: An introduction to linguistic morphology*, 3rd edn. Oxford: Oxford University Press.

Baayen, Harald. 1992. Quantitative aspects of morphological productivity. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1991*, 109–149. Dordrecht: Kluwer.

Baayen, R. Harald. 1994. Productivity in language production. *Language and Cognitive Processes* 9(3). 447–469.

Baayen, R. Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics*, 899–919. (HSK 29.2). Berlin & New York: De Gruyter.

Baayen, Harald & Rochelle Lieber. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29. 801–843.

Borer, Hagit. 2012. In the event of a nominal. In Martin Everaert, Marijana Marelj & Tal Siloni (eds.), *The theta system. Argument structure at the interface*, 103–149. Oxford: Oxford University Press.

Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Coussé, Evie. 2014. Lexical expansion in the HAVE and BE perfect in Dutch A constructionist prototype account. *Diachronica* 31. 159–191.

Demske, Ulrike. 2000. Zur Geschichte der ung-Nominalisierung im Deutschen: Ein Wandel morphologischer Produktivität. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 122. 365–411.

Demske, Ulrike. 2002. Nominalization and argument structure in early new High German. In Ewald Lang & Ilse Zimmermann (eds.), *Nominalisations*, 67–90. (ZAS Papers in Linguistics, 27). Berlin: ZAS.

De Smet, Hendrik. 2010. English -ing-clauses and their problems: The structure of grammatical categories. *Linguistics* 48(6).

Gaeta, Livio & Davide Ricca. 2006. Productivity in Italian word-formation. *Linguistics* 44(1). 57–89.

Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on "alternations." *International Journal of Corpus Linguistics* 9(1). 97–129.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge: MIT Press.

Hartmann, Stefan. 2014a. The diachronic change of German nominalization patterns: An increase in prototypicality. In Gabriella Rundblad, Aga Tytus, Olivia Knapton & Chris Tang (eds.), *Selected papers from the 4th UK cognitive linguistics conference*, 152–171. London: UK Cognitive Linguistics Association.

Hartmann, Stefan. 2014b. "Nominalization" taken literally: A diachronic corpus study of German word-formation patterns. *Italian Journal of Linguistics* 26(2). 123–155.

Hartmann, Stefan. 2014c. Constructing a schema: Word-class changing morphology in a usage-based perspective. In Martin Hilpert & Susanne Flach (eds.), *Yearbook of the German Cognitive Linguistics Association, Vol. 2*, 235–252. Berlin & New York: De Gruyter.

Hartmann, Stefan. 2016. *Wortbildungswandel. Eine diachrone Korpusstudie zu deutschen Nominalisierungsmustern*. Berlin & New York: de Gruyter.

Hilpert, Martin. 2006. Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2(2). 243–256.

Hilpert, Martin. 2008. *Germanic future constructions: A usage-based approach to language change*. (Constructional Approaches to Language, 7). Amsterdam & Philadelphia: John Benjamins.

Hilpert, Martin. 2014. *Construction grammar and its application to English*. Edinburgh: Edinburgh University Press.

Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticization: Opposite or orthogonal? In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What Makes grammaticalization?*, Vol. 158, 21–42. (Trends in Linguistics. Studies and Monographs). Berlin, New York: De Gruyter.

Hopper, Paul J. & Sandra A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60. 703–752.

Jespersen, Otto. 1946. *A modern English grammar on historical principles. Part V, Vol. IV: Syntax*. Copenhagen: Ejnar Munksgaard.

Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.

Kroch, Anthony, Beatrice Santorini & Lauren Delfs. 2004. The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME). Department of Linguistics, University of Pennsylvania. http://www.ling.upenn.edu/ppche-release-2016/PPCEME-RELEASE-3.

Kroch, Anthony, Beatrice Santorini & Ariel Diertani. 2010. Penn Parsed Corpus of Modern British English. http://www.ling.upenn.edu/hist-corpora/PPCMBE-RELEASE-1/index.html.

Langacker, Ronald W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

Marchand, Hans. 1969. *The categories and types of present-day English word formation: A synchronic-diachronic approach*. München: Beck.

Panagl, Oswald. 1987. Productivity and diachronic change in morphology. In Wolfgang U. Dressler (ed.), *Leitmotifs in Natural Morphology*, 127–151. (Studies in Language Companion Series 10). Amsterdam & Philadelphia: John Benjamins.

Panther, Klaus-Uwe & Linda Thornburg. 2001. A conceptual analysis of English -er nominals. In Martin Pütz, Susanne Niemeier & René Dirven (eds.), *Applied Cognitive Linguistics*, 149–200. (Cognitive Linguistics Research 19.2). Berlin & New York: De Gruyter.

Säily, Tanja. 2011. Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. *Corpus Linguistics and Linguistic Theory* 7(1). 119–141.

Sauer, Hans. 2004. Lexicalization and demotivation. In Geert E. Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds.), *Morphology*, 1625–1636. (HSK 17.2). Berlin & New York: De Gruyter.

Scherer, Carmen. 2006. Was ist Wortbildungswandel? *Linguistische Berichte* 205. 3–28.

Sleeman, Petra & Anna Maria Brito. 2010. Aspect and argument structure of deverbal nominalizations: A split vP analysis. In Artemis Alexiadou & Monika Rathert (eds.), *The syntax of nominalizations across languages and frameworks*, Vol.23, 199–217. (Interface Explorations). Berlin & New York: De Gruyter.

Stefanowitsch, Anatol. 2006. Distinctive collexeme analysis and diachrony: A comment. *Corpus Linguistics and Linguistic Theory* 2(2). 257–262.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Urdang, Laurence, Alexander Humez & Howard Zettler. 1982. *Suffixes and other word-final elements of English*. Detroit: Gale Research.