

Up and down the substantivization cline: Response to Bekaert & Enghels

Abstract

In their target paper, Bekaert & Enghels (B&E) (this issue) show convincingly that deverbal nominalizations differ with regard to their ‘nouniness’: while some capture the semantics of their base verbs quite faithfully, others approximate the semantic prototype of a noun as they refer to more concrete concepts. This is also reflected in their distributional behavior. This response paper relates B&E’s considerations to a different word-formation pattern in another language and to diachrony. Drawing on a corpus study of German *ung*-nominalization, it is shown how intracategorical heterogeneity can evolve diachronically, and the implications of this study as well as the research presented in the target paper for our understanding of categorial shift are discussed.

1. Introduction

In their target paper, Bekaert & Enghels (henceforth B&E) raise a number of interesting theoretical and empirical issues in the domain of word-formation. Drawing on a set of eight deverbal nominals in Spanish, they propose a transcategorical cline for Spanish nominalization, ranging from event nominals (e.g. *creación* ‘creation’, referring to the process of creating) to state nominals (e.g. *preocupación* ‘preoccupation’, referring to a state of mind) and referential nominals (e.g. *construcción* ‘building’, referring to an object). Their cline is a simplified version of Malchukov’s (2004, 2006) typologically oriented nominalization hierarchy. In the present paper, I wish to further develop this line of thought by discussing the ‘substantivization cline’ from a diachronic point of view and by addressing the question of how nominalization as a synchronic word-formation process relates to diachronic ‘nominalization’, i.e. the phenomenon that a word-formation product or word-formation pattern tends to assume more prototypically ‘nominal’ features over time. As an example, I will discuss the history of German *ung*-nominalization as sketched in, among others, Demske (2000, 2002) and Hartmann (2016, 2018). My review of earlier research will be complemented by a corpus study on the basis of the German Text Archive (*Deutsches Textarchiv*, DTA for short). In addition, I will discuss how ‘nominalization’, both synchronically and diachronically, can be accounted for in a usage-based model of language. Thus, the main aim of the present paper is to explore the origins of the intracategorical heterogeneity that B&E empirically demonstrate for Spanish nominalizations and that also characterizes German *ung*-nominalization. In addition, I will address unresolved theoretical and methodological questions with regard to the ‘nouniness’ of word-formation patterns and word-formation products.

The remainder of this paper is structured as follows. Section 2 discusses the ‘substantivization cline’ from a cognitive-linguistic and usage-based perspective. B&E’s classification scheme is compared to similar clines that have been proposed for German *ung*-nominalization by Ehrich & Rapp (2000) and by Hartmann (2014). Section 3 investigates the diachronic development of *ung*-nominalization on the basis of three different corpora. In

Section 4, the results are discussed with regard to their implications for the understanding of categorial shift in synchrony and diachrony.

2. ‘Nouniness’, ‘verbiness’, and the meaning-as-use hypothesis

As B&E show for Spanish, deverbal nominalizations differ in their degree of ‘nouniness’ – while some tend to capture the semantics of their base verb quite faithfully, others approximate the semantic prototype of a noun as they denote concrete entities or, more generally, what Langacker (e.g. 2008) refers to as ‘things’. Langacker (2008, p.105) defines a thing as “any product of grouping and reification”. While nouns prototypically denote things in this broad sense, a verb typically denotes a process, i.e. “a complex relationship that develops through conceived time and is scanned sequentially along this axis” (Langacker 2008, p. 112). Similar ideas are put forward in Croft’s (2001, p. 87) universal-typological part-of-speech model. Croft defines prototypical objects (roughly comparable to Langacker’s ‘things’) as nonrelational, stative, and permanent, while he defines actions (comparable to Langacker’s ‘processes’) as relational, processual, and transitory. When applied to nominalization (and verbalization), the ‘process’ and ‘thing’ prototypes, summarized in Table 1 (adopted from Fonteyn 2016), can be conceived of as endpoints on a continuum. While some nominalizations are closer to the ‘processual’ endpoint, others are more ‘thing’-like in that they construe the denoted entity as more holistic and/or decoupled from the participants of the event. This is what the parameter of relationality (or conceptual dependence) refers to: conceptualizing an event requires conceptualizing its participants, while a prototypical object can be conceptualized independently of its participation in any specific event (Langacker, 2008, p. 104). In many cases, these participants have to be overtly expressed. Nominalizations of transitive verbs differ in the degree to which the event participants can be omitted, and the omission of event participants usually entails a more concrete reading: compare, for example, the more processual *Bestellung von Spielsachen* ‘**order(ing) of toys**’ (Jean Paul, *Sämtliche Werke*, 1806, DTA) with the more concrete *Sie hatten eine Bestellung [...] zu überbringen* ‘They had to deliver an **order**’ (Achim von Arnim, *Der tolle Invalide*, 1818, DTA).

Table 1.

‘Process’ vs. ‘thing’ construal according to Langacker (1987), adopted from Fonteyn (2016, p. 76)

‘Process’ construal	‘Thing’ construal
relational / conceptually dependent	non-relational / conceptually independent
temporal / unfolding (sequential scanning)	atemporal / holistic (summary scanning)
⇒ dynamic	⇒ stative
⇒ transitory	⇒ persistent

The parameters mentioned in Table 1 are partly mirrored in B&E’s categories [\pm dynamic], [\pm temporal], and [\pm referential], which are operationalized by taking the collocations of the individual nominalizations into account. The concept of referentiality is connected to the notion of relationality mentioned in Table 1. As Croft (2001, p. 87) points out, nouns are prototypically used for reference, while verbs are prototypically used for predication. Reference

is inherently non-relational. Predication, by contrast, is inherently relational in that it refers to “what the speaker intends to *say about* what he is talking about (the referent)” (Croft, 1991, p. 52, emphasis original). Dynamicity and temporality, then, relate to the parameters subsumed under the heading of ‘temporal/unfolding’ and ‘atemporal/holistic’ in Table 1.

In order to operationalize their parameters, B&E use formal diagnostics that can be retrieved from the derived nominal’s larger context. For instance, a verb like *tener lugar* ‘take place’ occurring with the nominal points to an event construal, while a phrase like *un estado de* ‘a state of’ indicates that the derived nominal denotes a state. A similar approach has been proposed by Ehrich & Rapp (2000) for German *ung*-nominalizations: they use a four-way distinction that partly intersects with B&E’s three-way categorization scheme. While B&E distinguish between event nominals (EvNom), state nominals (StateNom), and referential nominals (RefNom), Ehrich & Rapp (2000) distinguish between process nominals, event nominals, result states, and result objects. The difference between process and event nominals is that the processes denoted by the latter, but not those denoted by the former, have an inherent result state (e.g. *Verfolgung* ‘pursuit’ vs. *Vernichtung* ‘annihilation’; cf. Ehrich & Rapp, 2000, p. 251). Ehrich & Rapp’s (2000) four categories are exemplified in (1), with attestations for the nominal *Absperrung* ‘cordon(ing off)’ from the German Reference Corpus (DeReKo, Kupietz et al., 2010).

- (1) a. Das Gemeindebauamt half bei der **Absperrung** der Strecke mit. (‘The municipal construction office helped **cordoning off** the route.’) (A08/MAI.05067) [process]
- b. Nach der **Absperrung** der besetzten Gebiete und des Gaza-Streifens hat Israel bislang 59 000 Ausländer angeworben, die die palästinensischen Arbeiter ersetzen sollen. (‘After **cordoning off** the occupied territories and the Gaza Strip, Israel has so far recruited 59,000 foreigners to replace the Palestinian workers.’) (U95/FEB.11701) [event]
- c. Die **Absperrung** eines Teiles der Straße bleibt aber bestehen. (‘The **cordoning off** of a part of the street will remain, however.’) (K98/MAR.17142) [result state]
- d. Ein Kunde hatte eine **Absperrung** zerstört. (‘A customer had destroyed the **cordon.**’ (P05/OKT.04221) [result object]

Ehrich & Rapp (2000) propose a variety of tests that disambiguate between the possible readings. For instance, processes can be combined with modifiers like *tagelang* ‘for (many) days’, while states can be combined with *bestehend* ‘existing’. Unlike B&E, however, they do not apply their diagnostics to actual corpus data.

Note that *Absperrung* in (1) shows the full semantic potential envisaged by Ehrich & Rapp, in that it allows for various construals from a process nominal to a result object reading.¹ However, the semantic potential of *ung*-nominals is even more fine-grained. Consider in this respect a nominal like *Lesung* ‘reading’, which tends to occur in ‘processual’ or ‘eventive’ contexts with *stattfinden* ‘take place’, the German equivalent of *tener lugar* mentioned by B&E, as a diagnostic for EvNoms. Compared to its base verb *lesen* ‘to read’, however, the meaning

¹ Ehrich & Rapp’s own example in their constructed example sentences is *Beklebung* ‘gluing/posting’.

of *Lesung* is restricted to a specific type of reading, namely a reading as a public event. In order to take such cases – which are common among German *ung*-nominals – into account, Hartmann (2014) extended the type of categorization scheme used by, for instance, Demske (2000) or Ehrich & Rapp (2000), which distinguishes between e.g. process, object, and person readings, with Langacker's (e.g. 2008) notion of 'bounded region'. This concept accommodates the observation that nominals like *Lesung* seem to be 'nounier' than typical action nouns as they profile an event as a whole, rather than the process. This is also reflected in their valence patterns: **Die Lesung des Buches* 'the reading of the book' seems odd, while *die Überwachung des Verdächtigen* 'the surveillance of the suspect' does not. As some *ung*-nominals can also refer to locations (e.g. *Kreuzung* 'crossroads', *Ausgrabung* 'excavation (site)'), Hartmann's (2014) classification scheme distinguishes bounded regions in time from bounded regions in space.

Regardless of the classification scheme one adopts, it has to be kept in mind that categories like EvNom, StateNom, and RefNom, or 'bounded region in time' and 'bounded region in space' can only delimit important 'landmarks', as it were, in a very heterogeneous semantic space. This heterogeneity can partly be attributed to the historical development of individual word-formation products. For instance, at the beginning of the 20th century, *Lesung* could still refer to the process of reading a text, rather than an event where a person reads out a text aloud. Ultimately, as argued in Hartmann (2016), paraphrasing the famous slogan probably going back to Hugo Schuchardt (cf. Campbell, 2013, p. 188), each word(-formation product) has its own history. If we take the usage-based assumptions seriously that "[m]eaning is use" (Tomasello, 2009, p. 69) and that "context, meaning, reference and use [...] co-determine each other" (Fischer, 2010, p. 46), any word-formation product will almost necessarily develop nuances of meaning that set it apart from its base, as it is used in different contexts and exhibits different collocational preferences. The behavior of individual word-formation products, however, feeds back into language users' constructional knowledge about the superordinate word-formation pattern. Accordingly, intracategorical heterogeneity can be considered an evolving variable – the semantic/functional potential of a word-formation construction reflects language change and keeps evolving along with the changing usage patterns of its instantiations. Taking diachrony into account can therefore offer an important additional perspective on the 'substantivization cline'.

3. A diachronic approach to nominalization

It has often been noted that *ung*-nominals in earlier stages of German differ considerably from present-day *ung*-derivatives. For example, Demske (2002, p. 68) argues that

in earlier periods of German *ung*-nominals are recategorizations of verbs only in a syntactic sense, sharing with their verbal stems linking properties and sortal interpretation. Regarding the latter properties, deverbal nouns evolve a more noun-like character not until a fairly recent stage in the history of German.

Drawing on a fairly small newspaper corpus, Demske (2000; see also Demske, this issue) shows that the word-formation pattern of *ung*-nominalization takes on more prototypically “nominal” features over time. The formal diagnostics she uses partly overlap with those operationalized by B&E, namely determination, pluralization, and adjectival modification. In addition, she observes that in Early New High German, *ung*-nominals are frequently used as complements of prepositions, as in (2) (from Demske, 2000, p. 380).

- (2) Diese wochen hat man alhie **in grabung** deß Grunds zu S. Petro ein Kreuzlein oder heyligthumb [...] gefunden. (‘This week, **in digging** the ground of St Peter’s [cathedral], a cross or sanctuary has been found.’) (Relation des Jahres 1609)

This [P NOM] construction² denotes anteriority, simultaneity, posteriority, or – in some cases – causality (see Demske, 2000, pp. 379–387). Importantly, however, the use of *ung*-nominals in this construction almost always entails a highly processual construal (see Hartmann, 2016, pp. 192–196). This also becomes clear in the examples from the DTAbaby corpus in (3) and (4).

- (3) allermaffen der günstige Lefer **in durchlefung** gedachter Reyße [...] zu erfehen haben wird. (‘as the interested reader will see **in reading** the aforementioned [lit. thought-of] journey’)
(Hulsius, Schiffahrt, 1649, DTAbaby)
- (4) wie etwa David **bei heimführung** der bundeslade mit aller macht tanzete (‘as David was dancing vigorously [lit. with all might] **when taking home** the Ark of the Covenant’)
(Geier, Heinrich Schütz, 1672, DTAbaby)

The *ung*-nominals occurring in the [P NOM] construction tend to refer to dynamic events which are processual (rather than stative) and transitory (rather than permanent). In addition, they tend to make the relationality exhibited by the base verb transparent by encoding what would be the (obligatory) direct object of the verb in a genitival complement, e.g. *in durchlefung gedachter Reyße* in (3) above. Thus, the decline of this construction lends support to the hypothesis that *ung*-nominalization becomes “nounier” over time.

Hartmann (2016) has replicated Demske’s (2000) results on the basis of two corpora – the *Mainz Early New High German Corpus* (MzENHG, Kopf 2016) and the *GerManC* corpus (Durrell et al., 2007) – and has extended the scope of investigation to more recent stages in the history of German. The remainder of this section summarizes the results of his study and adds findings from a third corpus, the German Text Archive (*Deutsches Textarchiv*, DTA). First, I will very briefly introduce the three corpora; then, the aggregated results from all three corpora will be presented and discussed (for a more detailed discussion of the data, see Hartmann 2016, 2018).

² The [P NOM] construction not only occurs with *ung*-nominals, but also with other nominalizations (see Hartmann 2016); in this paper, however, I will only be concerned with the sub-construction [P V-*ung*].

The MzENHG corpus is largely based on a corpus compiled by Bergmann & Nerius (1998) for investigating the development of sentence-internal capitalization in German. It comprises 80 texts from 1500 to 1710, and it is balanced for eight thirty-year periods and two text types (clerical texts and a variety of other texts subsumed under the broad heading of ‘non-fiction’, e.g. scientific texts or travel reports). Altogether, the version of the corpus used for the present study (see Hartmann 2016, pp. 126–128 for details) consists of 348,839 tokens. The GerManC corpus comprises 336 texts. It is balanced for three 50-year time periods from 1650 to 1800 and seven text types. Altogether, it contains 677,538 tokens. The DTA is a 100-million-word corpus covering the time span from 1600 to 1900.³ As it is not very well-balanced for periods and text types, I will use a sample of 270 texts (916,789 tokens) balanced for six fifty-year time periods and three text types (science, fiction, and so-called ‘functional texts’ [*Gebrauchsliteratur*], the latter comprising various text types from cookbooks to sermons). This balanced sample of the DTA will be referred to as *DTAbaby* (see Hartmann, 2018, for more details). The fact that the three corpora partly overlap allows for comparison between the data: if the data obtained for the same time period from two different corpora differ considerably, at least the results from one corpus might be a corpus artefact. If the differences are marginal, the results can be considered all the more convincing as they mutually reinforce each other.

In total, 3,178 *ung*-nominals (tokens) were attested in the MzENHG corpus (741 types), 7,039 tokens (1,192 types) were retrieved from the GerManC corpus, and 10,946 tokens (1396 types) were attested in the *DTAbaby*. The *ung*-nominals in all three corpora were annotated for the formal diagnostics mentioned above: (i) whether they occur with a determiner (and if so, which type of determiner they take); (ii) whether they are part of a [P NOM] construction; (iii) whether they are modified by an adjective, and (iv) whether they are pluralized. Fig. 1 shows the results.

³ The corpus used for the present study is based on the 2014 version of the DTA (available at http://media.dwds.de/dta/download/dta_komplett_2014-03-10.zip, last accessed on 8 August 2017). In the meantime, the corpus has grown and now also features text from the late 15th century and the 16th century.

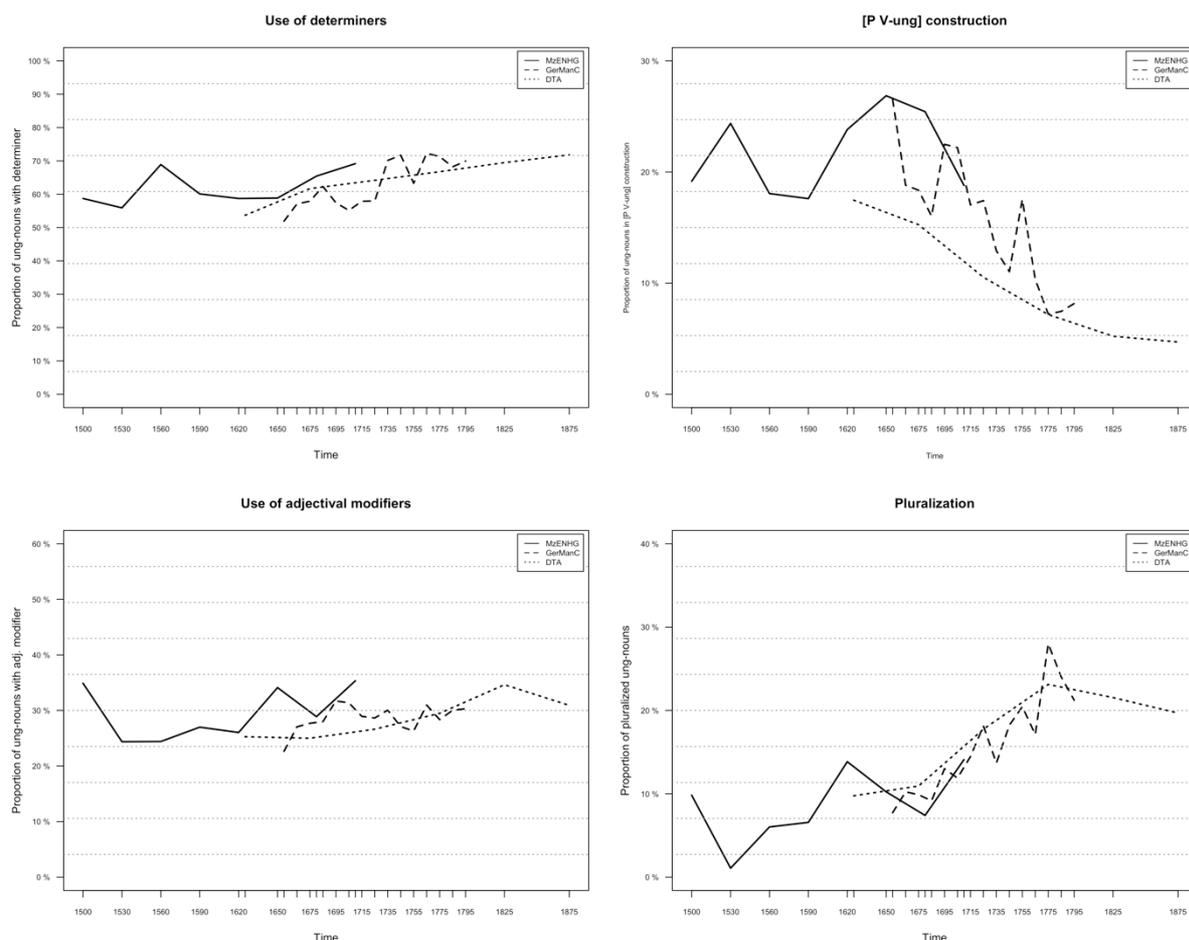


Fig. 1. Results of the corpus analysis on the basis of three different corpora

Let us discuss the individual diagnostics and their results in turn, starting with the use of determiners. Langacker (e.g. 1991, 2004) describes determiners as grounding elements, relating a discourse referent to the speech event, its participants, and the immediate context. More specifically, grounding pertains to whether an instance of a type, or a set of instances, is uniquely apparent to both speaker and hearer (see Langacker, 1991, p. 53). Note that nominal grounding – at least to a certain degree – always entails a count-noun construal (see Vogel, 1996, p. 131; Smirnova & Mortelmans, 2010, p. 113). Due to its inherent boundedness, a count noun is closer to the prototype (or “conceptual archetype”, as Langacker, 2004 calls it) of a noun (viz. a tangible, visible object; see Hopper & Thompson, 1984; Langacker, 2004; Fonteyn, 2016) than a more abstract mass noun or a very ‘verby’ action nominal, which is considered less individuated (see, e.g., Janda, 1996, p. 338; Enger & Nessel, 2011, p. 194). Using the signed-rank correlation coefficient Kendall’s Tau, which Hilpert & Gries (2009) have argued to be an appropriate measure for assessing the significance of frequency changes in diachronic corpora, we can identify the increase of *ung*-nominals with a determiner as significant both in the GerManC corpus ($\tau=0.6$, $T=84$, $p<0.01$) and in the DTA *baby* corpus ($\tau=1$, $T=15$, $p<0.01$).

As B&E point out, it makes sense to differentiate between different kinds of determiners. A study of the data in the DTA *baby*, for instance, reveals that among those *ung*-nominals that occur with a determiner, the indefinite and especially the definite article have become more prominent (Fig. 2). This may provide a clue to the semantics of the respective

nominalizations, especially with regard to their degree of ‘nouniness’. As Langacker (2008, p. 129) points out, “[o]nly a count noun permits the indefinite article.” The definite article entails an even higher degree of individuation as “it implies that the relevant scope of consideration contains only one relevant instance of the specified type” (Langacker, 2008, p. 285). A number of different quantifiers and other grounding elements that characteristically occur with more mass-noun-like elements are subsumed in the “other” category in Fig. 2, e.g. *solch* ‘such’, *jene* ‘that’, *diese* ‘this’, *allerhand/allerlei* ‘all sorts of’ or *alle* ‘all’. Many of these quantifiers tend to occur with mass nouns. Thus, the observation that *ung*-nominals in later periods are more likely to occur with definite or indefinite determiners can be interpreted as further evidence in favor of the hypothesis that those variants of derived nominals that have a more concrete reading (and can thus be regarded as more prototypical nonus) have become more prominent over time. The observation that the proportion of prenominal genitives and possessive pronouns has decreased can be attributed to the fact that the default position of the genitive in German has gradually shifted from prenominal to postnominal (see Nübling et al., 2013, pp. 107–110).

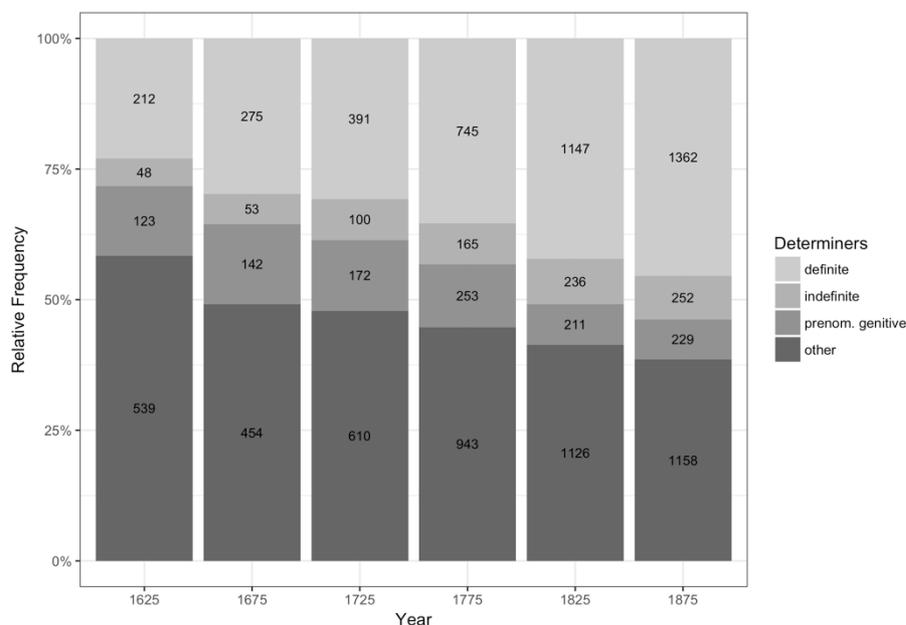


Fig. 2. Distribution of determiners for *ung*-nominals with a determiner in the DTAbaby corpus

As the upper right panel of Fig. 1 shows, the proportion of *ung*-nominals in [P NOM] constructions decreases significantly both in the GerManC corpus ($\tau=-0.68$, $T=17$, $p<0.001$) and in the DTAbaby corpus ($\tau=-1$, $T=0$, $p<0.01$). No significant changes can be detected for the use of adjectival modifiers (Fig. 1, lower left panel). However, Hartmann (2016, p. 226) observes for the MzENHG and GerManC data that ‘nounier’ *ung*-nominals seem to be more likely to occur with an adjective. A very simple distinctive collexeme analysis (Gries & Stefanowitsch, 2004) can be used to identify the *ung*-nouns that occur with or without an adjectival modifier above chance level in the aggregated data from MzENHG, GerManC, and DTAbaby. Distinctive collexeme analysis is typically applied to investigate which lexemes preferentially occur in the open slot of two (or more) semantically similar grammatical constructions, in this case, which lexemes preferentially occur in the verb slot in ditransitive vs. caused-motion constructions (*John gave me the book* vs. *John gave the book to me*). In the

present study, however, the method is used in a purely exploratory fashion, as we are not, strictly speaking, comparing two constructions, let alone two semantically similar constructions. Instead, what is used as input for the distinctive collexeme analysis is the aggregated list of all 21,163 lemmatized tokens from the three corpora with a binary annotation for the presence or absence of a modifier. As Table 2 shows, *ung*-nominals which preferentially occur in (or at least allow for) a variant with a more concrete reading tend to be identified as ‘attracted’ collexemes, i.e. they occur significantly more often with a modifier than expected by chance. For instance, *Verordnung* ‘order’ can refer not only to the event of passing a law or regulation, but also to the (abstract) entity, i.e. the law itself, as well as to a concrete object, i.e. a written text. *Zeitung* could refer to a piece of news but has now come to be used exclusively in the concrete sense of ‘newspaper’. *Abteilung* ‘division’ usually refers to a chapter of a book in our data,⁴ while *Sammlung* ‘collection’ can refer to a concrete assembly of collected things. Among the negatively attracted (‘repelled’) collexemes, some refer to states (e.g. *Verzweiflung* ‘despair’, *Ermangelung* ‘lack’), but some are also highly processual, e.g. *Erborenwerdung* ‘being born’ or *Eroberung* ‘conquest’.

Table 2.

Results of a collostructional analysis comparing *ung*-nominals with and without adjectival modification. The left half of the table shows the top collexemes that tend to occur with an adjectival modifier significantly more often than would be expected by chance, while the right half shows lemmas that are seldom or never modified by an adjective in the corpus data

“attracted” to adjectival modification				“repelled” from adjectival modification			
Lemma	Freq.	Freq. with adj.	coll.strength (LogLik)	Lemma	Frequency	Freq. with adj.	coll.strength (LogLik)
Anzeigeung ‘report’	103	62	40.98	Ansehung ‘view’	93	1	65.31
Verordnung ‘order’	89	50	26.84	Erhaltung ‘preservation’	100	6	70.23
Zeitung ‘newspaper’	174	84	26.4	Vermeidung ‘avoiding’	41	1	28.79
Einrichtung ‘institution’	77	43	22.64	Auferstehung ‘resurrection’	56	5	39.33
Verfassung ‘constitution’	102	50	16.65	Aufrichtung ‘raising / erecting’	20	0	14.05
Abteilung ‘division (also: chapter)’	53	30	16.46	Verzweiflung ‘despair’	55	5	38.63
Erscheinung ‘appearance’	144	65	15.2	Versöhnung ‘reconciliation’	18	0	12.64
Vorstellung ‘imagination / performance’	119	55	14.32	Erstattung ‘refund’	17	0	11.94
Untersuchung ‘examination’	141	63	14.03	Ermangelung ‘lacking’	16	0	11.24

⁴ This variant has fallen out of use in Present-day German. Nowadays, *Abteilung* refers to a department of, for instance, an institution or a shop.

Bedeutung 'meaning'	125	57	13.97	Weissagung 'prohecy'	25	1	17.56
Anstrengung 'effort'	25	16	12.48	Ausweisung 'expulsion'	15	0	10.53
Empfindung 'feeling'	153	66	12.32	Heiligung 'sanctification'	15	0	10.53
Bewegung 'movement'	309	120	11.74	Eroberung 'conquest'	31	2	21.77
Anschauung 'view'	35	20	11.29	Scheidung 'divorce'	42	4	29.5
Sammlung 'collection'	35	20	11.29	Befindung 'adjudging'	14	0	9.83
Verhandlung 'trial / negotiation'	47	25	11.2	Aufhebung 'abolition'	29	2	20.37
Ordnung 'order'	584	210	10.63	Abschaffung 'abolition'	13	0	9.13
Richtung 'direction'	87	40	10.16	Übertretung 'trespassing'	13	0	9.13
Durchdringung 'permeation'	4	4	9.7	Erborenwerdung 'being born'	12	0	8.43
Ölung 'oiling'	4	4	9.7	Verhütung 'prevention'	12	0	8.43

Interestingly, the fact that *Anzeigung* 'report' is identified in Table 2 as the collexeme most attracted to adjectival modification can almost exclusively be attributed to the earliest corpus data. For a more detailed investigation of how the adjectival modifiers are distributed over the time periods covered by the corpus, we can use diachronic distinctive collexeme analysis (DDCA, Hilpert, 2006). DDCA applies the logic of distinctive collexeme analysis as described above to the comparison of data not from different constructions, but from different time periods. Applying this method to the adjectives co-occurring with *ung*-nominals, we find that *redlich* 'righteous' and *genügsam* 'sufficient', which tend to co-occur with *Anzeigung*,⁵ are both strongly attracted to the first of the four centuries into which the aggregated data can be subdivided (16th, 17th, 18th, and 19th century). Other adjectives attracted to the 16th century include *fleißig* 'diligent', which tends to be used with event nominals such as *Verhörung* 'interrogation' and *Übung* 'practicing', and *gross* 'big/large', which emphasizes the 'mass-noun' character of the modified nominal, e.g. *groß Irrung* 'big mistake', *groß Lästerung* 'big heresy/blasphemy', or *groß Fäulung* 'big rot(ting)'. The adjectives that are attracted to the last of the four centuries are very different from these and include *wissenschaftlich* 'scientific', *künstlerisch* 'artistic', *bedeutend* 'significant', *amtlich* 'official', and *politisch* 'political'. Apart from reflecting cultural developments (and potentially differences in the composition of the corpus texts), these changing tendencies arguably also mirror the diachronic evolution of *ung*-nominals towards a more count-noun-like construal; e.g. *amtliche Stellung* 'official position', *politische Abhandlung* 'political tractatus', *künstlerische Gestaltung* 'artistic design'. In sum,

⁵ 39 out of 67 occurrences of *Anzeigung* co-occur with either one of these adjectives: *redlich* co-occurs 26 times with *Anzeigung* and only 5 times with one out of four other lemmas, while *genügsam* co-occurs 13 times with *Anzeigung* and 12 times with one out of 11 other lemmas.

even though the frequency development of *ung*-nominals with adjectival modification does not indicate any significant changes, a closer look at the interaction between modifiers and *ung*-nominals reveals interesting tendencies that are in line with the general hypothesis that the word-formation pattern becomes ‘nounier’ over time.⁶

A final indicator of ‘nouniness’ is pluralization. As Werner (2012, p. 7) points out, only count nouns allow for pluralization. If an abstract noun is used in the plural, this entails a count-noun construal, as Vogel (1996, p. 115) illustrates using the example of *beauty*: while the noun itself refers to an abstract feature, its pluralized form *beauties* refers to several entities (in this case, persons) characterized by this feature. The proportion of pluralized *ung*-nominals increases across all three corpora, despite some fluctuation that is particularly obvious in the MzENHG corpus and the GerManC corpus. However, the increase is only significant in the GerManC corpus ($\tau=0.77$, $T=93$, $p<0.001$). In the DTAbaby corpus, the increase is quite steady until the late 18th century, at which point the proportion of pluralized *ung*-nouns starts to slightly decrease again. Overall, however, it is clear that we find more pluralized *ung*-nominals at the end of the period covered by the DTAbaby corpus than at the beginning of that period.

In sum, then, all four formal diagnostics support the hypothesis that the place of the word-formation pattern on the ‘substantivization cline’ has shifted diachronically. Compared to earlier stages of German, *ung*-nominals in the 19th century – which is the latest period investigated in the diachronic study – tend to exhibit more characteristics of prototypical (count) nouns.

4. Discussion and conclusion

The main aim of the present paper was to expand B&E’s study (this issue) of intracategorical heterogeneity in two respects. On the one hand, I have argued that the idea of a ‘substantivization cline’ fits in well with cognitive-linguistic and usage-based approaches, which posit that meaning emerges from language use. On the other hand, I have presented a case study that taps into the diachronic origins of intracategorical heterogeneity, using the example of German *ung*-nominalization. The semantic change from processual to more concrete meanings that can be witnessed in the case of *ung*-nominalization is a cross-linguistically well-attested pattern, which Panagl (1987, p. 147) sees as an instance of ‘drift’ in the sense of Sapir (see also Fonteyn & Hartmann, 2016). The diachronic development of *ung*-nominalization can be accounted for in terms of a ‘substantivization cline’ with more relational,

⁶ Note that this is, first and foremost, a claim about the construction, i.e. about language users’ *knowledge* about the word-formation pattern, which can neither be separated from the instances of the construction nor from the contexts in which they occur. Importantly, I do not claim that the word-formation pattern, over time, tends to yield word-formation products that are inherently ‘nounier’ than the words coined by the same pattern in earlier stages. However, the word-formation products are *used* in a more noun-like fashion. This, in turn, can be assumed to be an important part of language users’ knowledge about the individual word-formation products. If the tendency to be used in a more “nouny” way applies to many word-formation products coined according to the same pattern, it seems reasonable to assume that it also becomes part of the users’ knowledge about the pattern itself, which in turn may help explain why coinages that are typically used in a more verb-like fashion, like *Murmelung* ‘muttering’ and *Schweigung* ‘silence’, have fallen out of use and seem ungrammatical (i.e. not compatible with the word-formation construction any more) today.

conceptually dependent construals on one end and more ‘thing’-like construals on the other, as indicated in Table 1. As diagnostics for this development, I largely relied on changes in distributional preferences already pointed out by Demske (2000), which in turn bear some similarities to B&E’s diagnostics. However, I have only briefly touched upon a major question that merits further discussion: does the ‘substantivization’ process only happen on the level of individual word-formation products (as argued e.g. by Müller 2016) or does it happen on the level of the construction, i.e. the word-formation pattern, as well?

This issue is closely related to the more general question of what actually undergoes categorial shift: the word-formation product, the word-formation pattern, or both? Ultimately, this question concerns the relation between complex words and their bases. While much of the metaphorical inventory employed in the standard terminology of linguistic morphology intimates that word-formation is very much a matter of on-line processing (*nominalizations* are *derived* from their bases; complex words come about through *processes* of word-formation), the usage-based perspective suggests that word-formation products are often firmly entrenched lexical items in their own right. However, the connections between complex words and their bases differ in strength: word-formation products can be semantically autonomous and weakly connected to their base, or they can be closely tied to the meaning of their base. Innovative meanings of derivatives can be transferred to other word-formation products via analogy, but they can also be reanalyzed as part of the semantic potential of the word-formation pattern and come to be used productively (see, e.g., Barðdal, 2008, on analogical extension at the lower end of the productivity cline, and Scherer, 2006, on semantic reanalysis). Therefore, it is not always possible to clearly distinguish between semantic change affecting individual word-formation products on the one hand and change affecting entire word-formation patterns on the other (see, e.g., Müller, 2016). A usage-based account explicitly accommodates this indistinctness. On this view, categorial shift is a consequence of language use on at least two different levels: on the micro-level (tied to individual utterances, hence inherently synchronic), it follows from how a linguistic unit is used in a specific context. On the macro-level (abstracting over individual utterances and hence, ultimately, inherently diachronic), it follows from the usage preferences of a linguistic unit. Both levels are tightly intertwined, as the usage preferences on the macro-level constrain the uses on the micro-level, while the macro-level construction is an abstraction over micro-level constructs (i.e. concrete instantiations of a construction; see Traugott & Trousdale, 2013, p.2).

Quantitative corpus studies by necessity take a macro-level perspective, as they abstract over individual utterances. However, both synchronic and diachronic corpus studies can provide clues to potential ‘triggers’ of change on the micro-level. Both the results from B&E’s analysis of a present-day word formation system and the diachronic analysis of a word-formation pattern reported on in the present paper suggest that the ‘feedback loop’ between micro- and macro-level also entails a co-evolution of constructs and constructions – hence, categorial shifts that affect individual word-formation products can have repercussions on the word-formation pattern as a whole. However, it is an open question to what degree individual word-formation products actually contribute to language users’ knowledge about the corresponding word-formation construction. It seems fair to assume (i) that different word-formation products contribute to language users’ knowledge about the superordinate construction to different degrees, depending on factors like their formal and semantic

transparency, their frequency, the frequency of their bases, their occurrence in constructional idioms, among others, and (ii) that it might be necessary to posit various sub-constructions for one word-formation pattern which differ in productivity and are subject to different word-formation constraints (see also the discussion about the ‘right’ level of abstraction for a construction in Hilpert, 2013). Empirical studies addressing these hypotheses in more detail would be an important desideratum for future research. Potentially, they could shed new light on how the semantic potential of a word-formation pattern develops, thus also contributing to our understanding of categorial shift.

The results discussed in Section 3 show that the distributional preferences of *ung*-nominalization have changed significantly. In all, these observations seem to indicate an increase in ‘nouniness’. However, from a usage-based point of view, cause and effect cannot be easily told apart, as constructional change always involves a cyclic relationship between instance and generalization: our knowledge about a construction, e.g. a word-formation pattern, comes about via generalization and abstraction over individual instances; these instances are in turn formed in accordance with what we know about the pattern, which includes knowledge about its constraints and its distributional preferences. As such, any change affecting the instances of a pattern potentially has ramifications for the pattern itself. This, in turn, means that some of the phenomena that were used as diagnostics for the increasing ‘nouniness’ of *ung*-nominalization in the present paper could also be causally involved in this development. For instance, the rise of the [P NOM] construction in Early New High German seems to contribute significantly to the ‘verby’ nature of *ung*-nominalization in that period. Conversely, the decline of this construction may have contributed to the decline of ‘verby’ *ung*-nominalization as well.

The interrelation between formal and functional aspects lies at the heart of B&E’s paper. An important merit of their study is that they apply their diagnostics to corpus data in order to test whether the formal/distributional behavior of the nouns they are investigating reflects their semantics. Their results suggest that this is actually the case. From a usage-based point of view, this finding is highly important as it adds more empirical evidence to the hypothesis that the meaning of a linguistic unit is both reflected and (partly) determined by its distribution (see, e.g., Firth, 1957; Levshina & Heylen, 2014; Perek, 2016). Any study that adds support to this idea makes research that fundamentally relies on the ‘meaning-as-use’ hypothesis less conjectural. A study of German *ung*-nominalization, or of other nominalization patterns, along the lines of B&E would be a highly welcome addition to the existing research. Given the heterogeneity of processual *ung*-nominals discussed above, it might also be interesting to complement B&E’s approach with diagnostics that help identify instances that entail a ‘bounded region’ construal, as discussed in Section 1. For the same reason, it would seem promising to apply bottom-up methods such as distributional semantics (e.g. Levshina & Heylen, 2014) to identify clusters in the data that may help us arrive at a more fine-grained data-driven categorization of derivatives. Such an analysis may also prove helpful in identifying potential sub-constructions of a word-formation pattern. Such bottom-up methods could also inform the theoretical discussion about intracategorial heterogeneity, as they might help to decide at which point sub-constructions can be posited that can be considered more or less independent from the parent construction, which in turn could shed new light on the question of whether a specific reading variant is part of the construction itself or rather the result of lexeme-specific processes of semantic change.

Acknowledgments

I would like to thank Hubert Cuyckens and Liesbet Heyvaert as well as two anonymous reviewers for providing helpful comments and raising some important questions. The usual disclaimers apply.

Sources and References

Corpora

DeReKo	=	<i>German Reference Corpus</i> , available via https://cosmas2.ids-mannheim.de/cosmas2-web/
DTA	=	<i>Deutsches Textarchiv</i> , http://deustextarchiv.de/
DTA <i>baby</i>	=	balanced sample of DTA, as described in the text
GerManC	=	<i>GerManC corpus</i> http://www.llc.manchester.ac.uk/research/projects/germanc/germancplus/
MzENHG	=	<i>Mainz Early New High German Corpus</i>

References

- Barðdal, Jóhanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam/Philadelphia: John Benjamins.
- Bergmann, Rolf & Dieter Nerius. 1998. Die Entwicklung der Großschreibung im Deutschen von 1500 bis 1710. Heidelberg: Winter.
- Campbell, Lyle. 2013. *Historical Linguistics: An Introduction*. 3rd ed. Cambridge: MIT Press.
- Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago: University of Chicago Press.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- 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. doi:10.1515/bgsl.2000.122.3.365.
- 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). Berlin: ZAS.
- Durrell, Martin, Astrid Ensslin & Paul Bennett. 2007. The GerManC Project. *Sprache und Datenverarbeitung* 31. 71–80.

- Ehrich, Veronika & Irene Rapp. 2000. Sortale Bedeutung und Argumentstruktur: -ung-Nominalisierungen im Deutschen. *Zeitschrift für Sprachwissenschaft* 19. 245–303. doi:10.1515/zfs.2000.19.2.245.
- Enger, Hans-Olav & Tore Nessel. 2011. Constraints on diachronic development: the Animacy Hierarchy and the Relevance Constraint. *Sprachtypologie und Universalienforschung* 64. 193–212.
- Firth, John Rupert. 1957. *Papers in Linguistics 1934-1951*. Oxford: Oxford University Press.
- Fischer, Kerstin. 2010. Quantitative Methods in Cognitive Semantics: Introduction to the Volume. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*, 43–59. Berlin, New York: De Gruyter.
- Fonteyn, Lauren. 2016. *Categoriality in Language Change. The Case of the English Gerund*. PhD Thesis: KU Leuven.
- Fonteyn, Lauren & Stefan Hartmann. 2016. Usage-based perspectives on diachronic morphology: A mixed-methods approach towards English ing-nominals. *Linguistics Vanguard* 2(1). doi:10.1515/lingvan-2016-0057.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending Collostructional Analysis: A Corpus-Based Perspective on “Alternations.” *International Journal of Corpus Linguistics* 9(1). 97–129.
- Hartmann, Stefan. 2014. 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. 2016. *Wortbildungswandel. Eine diachrone Studie zu deutschen Nominalisierungsmustern*. Berlin, Boston: De Gruyter.
- Hartmann, Stefan. 2018. Derivational morphology in flux: A case study of word-formation change in German. *Cognitive Linguistics* 29(1). 77–119.
- Hilpert, Martin. 2006. Distinctive Collexeme Analysis and Diachrony. *Corpus Linguistics and Linguistic Theory* 2(2). 243–256.
- Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.
- Hilpert, Martin & Stefan Th Gries. 2009. Assessing Frequency Changes in Multistage Diachronic Corpora: Applications for Historical Corpus Linguistics and the Study of Language Acquisition. *Literary and Linguistic Computing* 24(4). 385–401.
- Hopper, Paul J. & Sandra A. Thompson. 1984. The Discourse Basis for Lexical Categories in Universal Grammar. *Language* 60. 703–752.
- Janda, Laura A. 1996. Figure, Ground, and Animacy in Slavic Declension. *The Slavic and East European Journal*, 40(2). 325–355.
- Kopf, Kristin. 2016. *Mainzer (Früh-)Neuhochdeutschkorpus*. Mainz: Johannes Gutenberg-Universität Mainz.
- Kupietz, Marc, Cyril Belica, Holger Keibel & Andreas Witt. 2010. The German Reference Corpus DeReKo: A primordial sample for linguistic research. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, 1848–1854. Valletta: European Language Resources Association. http://www.lrec-conf.org/proceedings/lrec2010/pdf/414_Paper.pdf.
- Langacker, Ronald W. 1987. Nouns and Verbs. *Language* 63. 53–94.

- Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar. Vol. 2: Descriptive Application*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2004. Remarks on Nominal Grounding. *Functions of Language* 11(1). 77–113.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Levshina, Natalia & Kris Heylen. 2014. A Radically Data-driven Construction Grammar: Experiments with Dutch causative constructions. In Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the Scope of Construction Grammar*. Berlin, New York: De Gruyter.
- Malchukov, Andrej L. 2004. *Nominalization/Verbalization: Constraining a Typology of Transcategorial Operations*. Munich: LINCOM.
- Malchukov, Andrej L. 2006. Constraining Nominalization: Form/Function Competition. *Linguistics* 44(5). 973–1009.
- Müller, Peter O. 2016. Wortbildungswandel oder Bedeutungsbildung? Zur Entstehung und Interpretation sekundärer Wortbildungsbedeutungen. In Sarah Kwekkeboom & Sandra Waldenberger (eds.), *PerspektivWechsel oder: Die Wiederentdeckung der Philologie. Bd. 1: Sprachdaten und Grundlagenforschung in Historischer Linguistik*, 309–332. Berlin: Erich Schmidt Verlag.
- Nübling, Damaris, Antje Dammel, Janet Duke & Renata Szczepaniak. 2013. *Historische Sprachwissenschaft des Deutschen: Eine Einführung in die Prinzipien des Sprachwandels*. 4th ed. Tübingen: Narr.
- Panagl, Oswald. 1987. Productivity and Diachronic Change in Morphology. In Wolfgang U. Dressler (ed.), *Leitmotifs in Natural Morphology*, 127–151. Amsterdam and Philadelphia: John Benjamins.
- Perek, Florent. 2016. Using distributional semantics to study syntactic productivity in diachrony. A case study. *Linguistics* 54(1). 149–188.
- Scherer, Carmen. 2006. Was ist Wortbildungswandel? *Linguistische Berichte* 205. 3–28.
- Smirnova, Elena & Tanja Mortelmans. 2010. *Funktionale Grammatik: Konzepte und Theorien*. Berlin, New York: De Gruyter.
- Tomasello, Michael. 2009. The Usage-Based Theory of Language Acquisition. In Edith Laura Bavin (ed.), *The Cambridge Handbook of Child Language*, 69–87. Cambridge: Cambridge University Press.
- Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
- Vogel, Petra M. 1996. *Wortarten und Wortartenwechsel: Zur Konversion und verwandten Erscheinungen im Deutschen und in anderen Sprachen*. Berlin, New York: De Gruyter.
- Werner, Martina. 2012. Genus, Derivation und Quantifikation: Zur Funktion der Suffigierung und verwandter Phänomene im Deutschen. Berlin, New York: De Gruyter.