

Evaluative Compounding in Dutch: Productivity and Competition in Simile-Based Constructions

Kristel Van Goethem, F.R.S.-FNRS & UCLouvain, Belgium

Abstract

*Although compounding is one of the most common word-formation processes, its role in evaluative formation is often considered marginal. Yet, in some languages, compounding serves as an important source of evaluative morphology. This paper investigates evaluative compounding in Dutch through two case studies, focusing on its productivity and interaction with corresponding phrasal constructions. The first case study summarizes findings from previous research on nominal ameliorative compounds and their binominal phrasal counterparts (e.g., *een prachtvrouw* ‘lit. a beauty-woman’ vs. *een pracht van een vrouw* ‘a beauty of a woman’). The second case study provides a new corpus-based study of adjectival evaluative compounds expressing intensification through comparison – so-called “similes” or “relative compounds” (Hoeksema 2012) – and their phrasal equivalents (e.g., *apetrots* ‘lit. monkey-proud’ vs. *zo trots als een pauw* ‘as proud as a peacock’).*

Our findings reveal consistent tendencies: in general, evaluative phrasal patterns are more flexible, productive and creative, whereas evaluative compounds, though frequently attested in Dutch, are often (semi-)lexicalized and may involve grammaticalized affixoids. Furthermore, phrasal constructions favor animate fillers, while evaluative compounds attract inanimate fillers, suggesting a functional division of labour between both patterns. At the level of specific expressions, however, the relationship becomes far more complex, and may be characterized by overlap, competition, or divergence.

Keywords: *evaluative morphology; compounding; simile constructions; Dutch; corpus study.*

1 Introduction

Štekauer (2015) identifies several word-formation processes that can convey evaluative meaning, including affixation, reduplication, compounding, and sound symbolism. Among these, affixation is described as “by far the most common strategy employed in evaluative morphology” (Štekauer 2015: 47). Consequently, research on evaluative morphology (e.g., Grandi & Körtvélyessy 2015) has predominantly focused on derivational processes, such as diminutive suffixes and intensifying prefixes (e.g., Dressler & Barbaresi 1994; Rainer 2015; Manova et al. 2023). Interestingly, Štekauer (2015: 51) points out an intriguing paradox: although compounding – alongside suffixation – is one of the most widespread word-formation processes cross-linguistically, its role in evaluative morphology is often limited. In many languages, its use is restricted and typically involves combinations with specific lexical items.

This paper investigates evaluative compounds in Dutch. Following a brief cross-linguistic overview of evaluative compounding (Section 2), the focus shifts to Dutch, where we examine both its productivity and its interaction with evaluative phrasal constructions that express similar meanings. The analysis is based on two case studies. The first explores nominal ameliorative compounds (e.g., *een droomhuis* ‘a dream house’) and their interaction with Dutch evaluative binominal phrases (e.g., *een droom van een huis* ‘a dream of a house’). As this case

study has already been addressed in a previous publication (Van Goethem 2023), we will only summarize the relevant findings in Section 3. The second case study presents new research and centers on adjectival simile constructions, which can appear either as compounds (e.g., *beresterk* ‘lit. bear-strong; very strong’) or as phrasal expressions (e.g., *zo sterk als een beer* ‘lit. as strong as a bear’). Section 4 introduces the topic of simile constructions and reviews relevant literature. Sections 5 and 6 present the methodology and results of a corpus-based study comparing the structural and semantic properties, as well as the productivity, of both types of simile constructions in Dutch. Section 7 provides a detailed discussion of the findings, exploring the interaction between productivity and semantic coherence of the simile-based constructions. Section 8, finally, presents the concluding remarks, offering our corpus-based insights into the interplay between evaluative compounds and evaluative phrases in Dutch.

2 Evaluative compounding

Although evaluative compounding is surprisingly rare, noteworthy examples can still be found across various languages. Štekauer (2015: 51–52) presents instances of different types, including diminutive compounds (1), augmentative compounds (2), and elative compounds which express a high degree (3).

- (1) Bahasa Indonesia: *anak ayam* ‘lit. child hen; chicken’
- (2) Nêlêmwa: *axomoo naabuc* ‘lit. mother mosquito; big mosquito’
- (3) Dutch: *beresterk* ‘lit. bear strong; very strong’,

Dutch elative compounds have been extensively examined by Hoeksema (2012), who defines them as follows:

By the term elative compound, I will refer to any compound XY in which X serves as a modifier that is used to denote a very high degree of the property associated with head element Y. (...) [T]his means that elative compounds are possible only when Y is gradable. (Hoeksema 2012: 98)

The literature on such compounds typically refers to adjectival simile compounds like *ijskoud* ‘ice-cold’ and *doodziek* ‘lit. dead ill’ (see e.g., Booij 2010; Hoeksema 2012; Hendrikx et al. 2017; van der Wouden & Foolen 2017), which is understandable given that adjectives are often gradable and thus lend themselves well to expressing a high evaluative degree. Nevertheless, Hoeksema (2012: 115–117) also provides examples of nominal elative compounds with a gradable nominal head, such as *bloedvaart* ‘lit. blood rush’, *wolfshonger* ‘lit. wolf hunger’, and *doodzonde* ‘lit. dead sin’.

In the nominal compounds central to our first case study (e.g., *een droomhuis* ‘a dream house’, *een prachtdochter* ‘lit. a beauty-daughter’) (see Section 3), the first component also expresses an evaluation, namely a positive one, but the nominal head does not need to be gradable (e.g., *huis* ‘house’, *dochter* ‘daughter’). Furthermore, corpus data reveal that nominal evaluative compounding in Dutch is not limited to ameliorative semantics (4); it also

encompasses pejorative (5), augmentative (6), and diminutive (7) meanings (all examples drawn from nlTenTen20).¹

- (4) *De kanosport is een **prachtsport** (...) ‘Canoeing is a magnificent sport (lit. beauty-sport) (...)’*
- (5) *Zijn vriendin laat het er echter niet bij zitten, zo’n **snotjong** leest haar toch echt de les niet. ‘His girlfriend doesn’t let it slide; some brat (lit. snot-boy) like that certainly isn’t going to lecture her.’*
- (6) *Er hoeft maar iets te gebeuren en de file erachter groeit al snel uit tot een **monsterfile**. ‘It only takes a small incident, and the traffic jam behind it quickly turns into a monster jam.’*
- (7) *Renault Clio is de **dwergauto** die uitgebracht werd sinds 1990. ‘The Renault Clio is the dwarf car that has been released since 1990’*

This indicates that evaluative compounds in Dutch are not confined to elative meanings but instead span a broad range of semantic types.

A second observation is that the first components of these evaluative compounds often acquire a more abstract and semantically bleached meaning, which qualifies them as *affixoids* (see Booij 2010; Amiot & Dugas 2020). This phenomenon reveals the blurred boundary between compounding and affixation. Amiot & Dugas (2020), for instance, mention the French prefixoid *bébé-* ‘baby’, which has become grammaticalized to convey the meaning ‘small’ in compounds such as *bébé-voiture* ‘very small car’ and *bébé-entreprise* ‘very small/young company’. A comparable development is found in Dutch with *reuze-*, derived from the noun *reus* ‘giant’, which functions as an intensifier in adjectival compounds like *reuzeleuk* ‘very nice’ and even *reuzeklein* ‘very small’ (cf. Norde & Van Goethem 2014; Van Goethem & Hiligsmann 2014; Van Goethem 2020).

3 Previous research: the Evaluative Binominal Construction

The main objective of Van Goethem (2023) was to re-examine the so-called Evaluative Binominal Construction (EBC) through a detailed corpus-based analysis. While Dutch expressions like *een schat van een kind* ‘lit. a treasure of a child; a sweet child’ have been widely discussed in earlier Dutch literature (cf. Paardekooper 1956; Everaert 1992; Foolen 2004; Verhagen 2005; Haeseryn et al. 2019; Broekhuis 2020), they had not previously been examined from a corpus-linguistic perspective. Moreover, this type of phrasal expression had not yet been systematically compared to its potential morphological counterpart in the form of a binominal compound. As illustrated in examples (8–11), such a correspondence is sometimes possible, but not systematically (11).

- (8) *een droom van een huis* ‘a dream of a house’ vs *een droomhuis* ‘a dream house’

¹ See also the paper by Hoeksema on pejorative and meliorative affixoids in Dutch in this special issue.

- (9) *een pracht van een dochter* ‘a beauty of a daughter’ vs *een prachtdochter* ‘lit. a beauty-daughter’
- (10) *een wonder van een vrouw* ‘a wonder of a woman’ vs *een wondervrouw* ‘a wonder woman’
- (11) *een schat van een kind* ‘lit. a treasure of a child; a sweet child’ vs **een schatkind* ‘lit. a treasure-child’

This case study focused on the four positively connotated evaluative nouns illustrated in the examples (8–11), and their function within both evaluative compounds and phrasal expressions. The methodological details are discussed in Van Goethem (2023); in the present paper, we will limit ourselves to reporting the main findings, focusing on their formal characteristics, semantic tendencies, and productivity.

First, the study revealed that the syntactic EBC patterns tend to allow more complex slot fillers, including compounds and full noun phrases. For example, constructions such as *een schat van een Engelse nanny* ‘a sweetheart (lit. treasure) of an English nanny’ illustrate the flexibility of this pattern. In contrast, morphological EBC patterns typically involve simplex slot fillers, as in examples like *een wonderfiets* ‘a wonder bike’, indicating a more constrained structural pattern.

Additionally, a clear semantic distinction emerged between the two types. Syntactic EBC patterns predominantly feature animate slot fillers, often referring to PEOPLE or ANIMALS (e.g., *een droom van een moeder* ‘a dream of a mother’; *een schat van een labrador* ‘lit. a treasure of a labrador; a sweet labrador’). Morphological EBC patterns, on the other hand, are more frequently associated with inanimate slot fillers, such as PLACES or ABSTRACT CONCEPTS (e.g., *een droomappartement* ‘a dream apartment’; *een prachtkans* ‘a beautiful opportunity’).

Finally, the syntactic EBC shows greater productivity, particularly in the patterns [*een droom/pracht/wonder van een N*]. This is evidenced by the presence of numerous hapax legomena, including novel expressions such as *een droom van een stofzuiger* ‘a dream of a vacuum cleaner’ and *een pracht van een ribeye steak* ‘a beauty of a ribeye steak’. In contrast, the morphological EBC appears less productive and more lexicalized, sometimes even leading to semi-fixed expressions such as *een wonderkind* ‘a prodigy child’.

More generally, some constructions turned out to function as close alternates with respect to semantics and productivity, such as [*een pracht-N*] and [*een pracht van een N*], while others exhibit notable divergence in usage, as seen in [*een droom-N*] versus [*een droom van een N*].

In what follows, we will examine a new case study focusing on simile-based constructions in both compounding and syntax, with the aim of determining whether the observed tendencies can be generalized beyond the initial EBC data.

4 Simile constructions

Israël et al. (2004) argue that “simile” constitutes a distinct rhetorical figure, characterized by three core properties: it involves a comparison, the comparison is made explicit (in contrast to metaphor), and it typically links entities that are not conventionally seen as comparable, thereby creating a figurative rather than a literal comparison. Crucially, similes often convey

superlative evaluation, a phenomenon the authors refer to as the “Superlative Source Constraint”. This means that similes prompt the interpretation of their source concept as an idealized exemplar or prototype of the property being compared. For instance, the expression *her argument was as clear as glass* implies an exceptionally high degree of clarity. Moreover, Israël et al. argue that simile is not restricted to canonical constructions such as *as X as Y* or *X like Y*; a wide range of syntactic patterns may serve this evaluative function. Example (12), cited by Israël et al. (2004: 125), illustrates how similes can be embedded in more creative and less conventional structures while still fulfilling their evaluative and comparative role.

- (12) *And my husband and I basically view skiing as an invitation to suicide*

Cross-linguistic research has shown that both phrasal and morphological simile constructions coexist in many European languages, though their structure, productivity, and interchangeability vary considerably (Schlücker 2019, ed.). Coexistence is illustrated in examples (13–16) from English, German, Finnish and Italian.

- (13) English: *as red as blood* vs *blood red* (Finkbeiner & Schlücker 2019)
- (14) German: *so weich wie Butter* ‘as soft as butter’ vs *butterweich* ‘butter-soft’ (Schlücker 2019)
- (15) Finnish: *hidas kuin etana* ‘slow as a snail’ vs *etanan hidas* ‘lit. snail_{gen} slow’ (Hyvärinen 2019)
- (16) Italian: *bianco come la neve* ‘as white as snow’ vs *bianco neve* ‘snow white’ (Masini 2019)

Nevertheless, the interchangeability of these forms is not always guaranteed. For example, Danish allows both *dum som snot* ‘dumb as snot’ and *snotdum* ‘lit. snot-dumb’, while Swedish permits *mjuk som smör* ‘soft as butter’ but not *smörmjuk* ‘butter-soft’. These findings suggest that while simile constructions are widespread, their formal realization and degree of productivity are language-specific and subject to lexical and syntactic constraints.

Dutch simile constructions also include both phrasal and morphological types, each with distinct structural and functional properties. Phrasal similes, as described by Broekhuis (2015), include both (semi-)fixed expressions (e.g., *zo sterk als een beer* ‘as strong as a bear’) and more creative formations (e.g., *zo stoned als een garnaal* ‘lit. as stoned as a shrimp’). These constructions are syntactically constrained and cannot occur in attributive positions (e.g., **een zo sterke als een beer jongen* ‘lit. a as strong as a bear boy’). In contrast, morphological similes fall under the category of elative compounds, which are compounds in which the modifier intensifies the property associated with the head (see Section 2). Hoeksema (2012) further distinguishes between three types: stereotyped comparisons (e.g., *pijlsnel* ‘lit. arrow-fast’), analogical extensions of stereotyped comparisons (e.g., *beresterk* ‘lit. bear-strong’ → *beregezellig* ‘lit. bear-cosy’), and causative compounds (e.g., *doodmoe* ‘lit. dead-tired’), which may also allow for analogical extensions (e.g., *doodgewoon* ‘lit. dead normal’). It can already be noted that, in Dutch as well, the correspondence between phrasal and morphological simile constructions is not systematic. While *zo sterk als een beer* aligns with *beresterk*, similar

phrasal constructions such as *zo sterk als een paard* ‘as strong as a horse’ do not yield acceptable morphological equivalents (e.g., **paardesterk* ‘lit. horse-strong’).

5 Corpus data and methods

For this study on morphological and phrasal simile-based constructions, we selected ten Dutch adjectives that describe human characteristics, whether physical, mental, or emotional (see Table 1). The selection was guided by three criteria: (i) the semantic criterion (human properties) to ensure contextual comparability across the patterns, (ii) sufficient frequency of the adjectives in the corpus, and (iii) frequent occurrence of the adjectives in simile constructions.²

Table 1: Selection of head adjectives

1. <i>bang</i> ‘afraid’	1. <i>mooi</i> ‘beautiful’
2. <i>blij</i> ‘happy’	2. <i>sterk</i> ‘strong’
3. <i>dom</i> ‘dumb’	3. <i>stil</i> ‘quiet’
4. <i>gelukkig</i> ‘happy, lucky’	4. <i>trots</i> ‘proud’
5. <i>gek</i> ‘crazy’	5. <i>ziek</i> ‘sick’

The corpus study was conducted using a random sample of 2,000 instances for each adjective in each pattern – syntactic and morphological – extracted from the Dutch nlTenTen20 web corpus via SketchEngine (Kilgarriff et al. 2014). The corpus queries used to retrieve syntactic and morphological similes, based on the example of the adjective *blij* ‘happy’, are shown in (17) and (18). These queries were obtained through trial and error in order to achieve optimal results.³ As illustrated in (18), frequently occurring irrelevant results, such as the prefixed form *superblij* ‘super happy’, were excluded from the analysis by a specific operation inserted in the query.

- (17) Syntactic simile (‘as happy as a’):
 [lemma = "zo"] [lemma = "blij"] [lemma = "als"] [lemma = "een"]

²As noted in Section 4, a wide range of syntactic patterns can express a similitive function, not only the [*as* Adj *as* NP] construction. More specifically, similes are not restricted to adjectival constructions but may also have a verbal head, as rightly observed by a reviewer of this paper (e.g., *beven als een rietje* ‘to tremble like a leaf (lit. like a reed)’, *slapen als een baby* ‘to sleep like a baby’). However, for the purpose of this study, we focus on adjective-based simile constructions in order to enable comparison with their adjectival compound counterparts.

³The corpus query specified the indefinite article *een* as part of the simile pattern in order to obtain optimal results. Indeed, when the article was omitted, a significant amount of noise appeared in the corpus data (e.g., *Deze mensen waren net zo bang als ik* ‘These people were just as afraid as I was’; *Waarom ben je dan zo gelukkig, als ik vragen mag?* ‘Then why are you so happy, if I may ask?’). We acknowledge, however, as rightly pointed out by a reviewer, that this query did not capture the full range of possible simile constructions, including those with a noun phrase starting with the definite article or with a bare noun (e.g., *zo lelijk als de nacht* ‘lit. ‘as ugly as the night’, *zo wit als krijt* ‘lit. as white a chalk’). Moreover, the inclusion of *zo* ‘as’ in the query followed the same rationale, namely to maximize relevant results and minimize false positives (e.g., *Hij is telkens ziek als er een belangrijke voetbalwedstrijd is geweest* ‘He is falling ill every time an important football match has taken place’), even though this query excludes alternative formulations without *zo* (e.g., *trots als een pauw* ‘proud as a peacock’).

- (18) Morphological simile (*happy):
[lemma = ".+blij" & lemma != "superblij" & tag = "adj.*"]

Subsequently, each of the 20 random samples was manually reviewed until 100 relevant concordance lines were identified per simile type, resulting in a total of 2,000 relevant occurrences. During this manual screening, irrelevant examples, such as literal comparisons (see example 19) and non-simile compounds (see example 20), were excluded from further analysis.

- (19) *Laminaatvloeren zijn tegenwoordig net zo mooi als een houten vloer [...]* ‘Laminate floors are nowadays just as beautiful as wooden floors [...]
- (20) *Vorig jaar vond ik het wel erg speciaal op zee jarig te zijn maar nu voelde ik me een beetje zeeziek [...]* ‘Last year, I found it quite special to celebrate my birthday at sea, but this time I felt a bit seasick [...]

All relevant concordance lines were annotated in Excel files according to the properties listed in Table 2.⁴

Table 2 : Annotation grid

Structural properties	Part of Speech of the simile’s reference point (N/NP, A, V, NP) Syntactic function (Attributive, Predicative, Adverbial)
Semantic properties	Semantic category of the simile’s reference point (PERSON, ANIMAL, OBJECT, PLACE, CONCEPT, PROPERTY, ACTION)
Productivity	Annotation of the lemma of the reference point, to calculate two productivity ratios (see Section 6.3): - Type-Token Ratio (TTR) - Hapax Legomena-Token Ratio or Potential Productivity (PP)

6 Results

In this section, we will report the results related to the structural properties (6.1), semantic features (6.2) and productivity (6.3) of the syntactic and morphological simile constructions.

6.1 Structural properties

First, we compare the syntactic and morphological similes with respect to the POS of the item expressing the reference point, and its complexity.

⁴ The annotation was carried out during the master’s research seminar LNED2731 “Construction Grammar: A corpus-based exploration of Dutch constructions” (Spring 2025, Université catholique de Louvain). Each student and the instructor (author of this paper) was assigned an adjective, which was analyzed within both the morphological and syntactic simile constructions. The annotation criteria were discussed during the course, and the actual annotation took place during class sessions. Difficult examples and semantically ambiguous cases were addressed collectively to ensure consistent annotation across the different adjectives.

Overall, when merging all results from the syntactic similes in the dataset, we notice that in 80% of the cases, the pattern [zo Adj *als een*] is followed by a single noun (e.g., *zo blij als een kind* ‘as happy as a child’). In the remaining 20% of the cases, the nominal slot of this sequence takes the form of a noun phrase (e.g., *zo bang als een schildpad in het donker* ‘as scared as a turtle in the dark’).

In the morphological similes, we observe mainly [N+A] compounds (78%) (e.g., *bloedmooi* ‘lit. blood-beautiful; very beautiful’), followed by [A+A] compounds (20%) (e.g., *dolgelukkig* ‘lit. crazy-happy; madly happy’, and only rarely [V+A] and [Prep + A] compounds (e.g., *knettergek* ‘lit. crackling-crazy; completely crazy’; *doorziek* ‘lit. through-sick; very sick’).

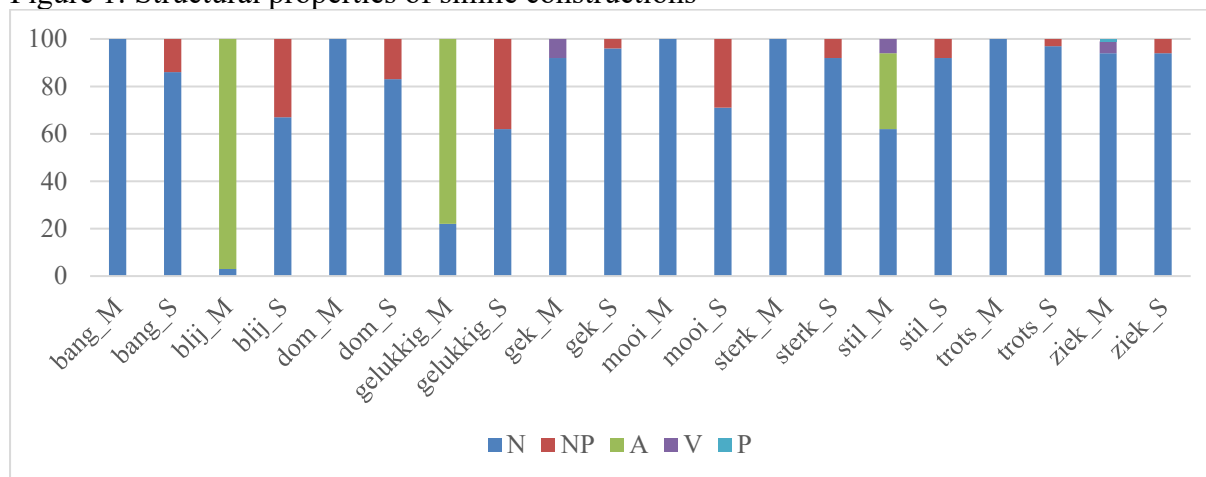
Figure 1 provides a detailed overview of the structural properties for each item (the extension _M refers to the morphological similes and _S to the syntactic patterns). The figure clearly illustrates variation across items. For example, the proportion of noun phrases (NPs) in the syntactic pattern ranges from 3% to 38%. Some creative examples with extended noun phrases are given in (21–22).

(21) *Wij zijn net zo blij als een uitgehongerde hond die een piepklein kluifje krijgt toegeworpen* ‘We are as happy as a starving dog being tossed a tiny bone’

(22) *Of ben ik alleen wat ik weet van mijzelf: Onrustig, vol heimwee, zo ziek als een gekooide vogel, snakkend naar lucht (...)* ‘Or am I only what I know of myself: restless, full of longing, as sick as a caged bird, gasping for air (...)’

From Figure 1, it can also be drawn that the formal structure of morphological similes appears to be adjective-dependent. Whereas the [N+A] structure is predominant for most items, the synonymous adjectives *blij* and *gelukkig* ‘happy’ most frequently occur in the [A+A] pattern, especially in *dolblij* and *dolgelukkig* ‘lit. crazy-happy’.

Figure 1: Structural properties of simile constructions



Regarding the syntactic function of simile constructions, phrasal patterns are almost exclusively used in a predicative position (see Figure 2). However, examples such as (23), where the phrasal simile appears postnominally without a copular verb, have been classified as attributive uses – a function that Broekhuis (2015) considers ungrammatical (cf. Section 4).

- (23) *Vakantiehuis op een plek zo mooi als een parel* ‘Holiday home in a place as beautiful as a pearl’

In contrast, morphological similes exhibit a more balanced distribution between predicative (66%) and attributive use (30%), suggesting that simile compounds are syntactically more flexible than their phrasal counterparts. While certain compound adjectives show a strong preference for predicative use (e.g., *bang* ‘afraid’, *blij* ‘happy’, *gelukkig* ‘happy, lucky’, *trots* ‘proud’), others frequently occur in attributive position (e.g., *sterk* ‘strong’, *gek* ‘crazy’, *dom* ‘dumb’). For *sterk* ‘strong’, attributive use even reaches up to 70% of its occurrences; this is largely explained by the metaphorical use of the compound *ijzersterk* ‘iron-strong’ in prenominal position (see example 24).

- (24) *Boeiende cases, keiharde spanning en actie, ijzersterke acteerprestaties en tientallen bekende gastacteurs, dat zijn nog steeds de pittige ingrediënten van Zone Stad (...)* ‘Fascinating cases, rock-hard tension and action, iron-strong acting performances and dozens of well-known guest actors—those are still the spicy ingredients of Zone Stad (...)’

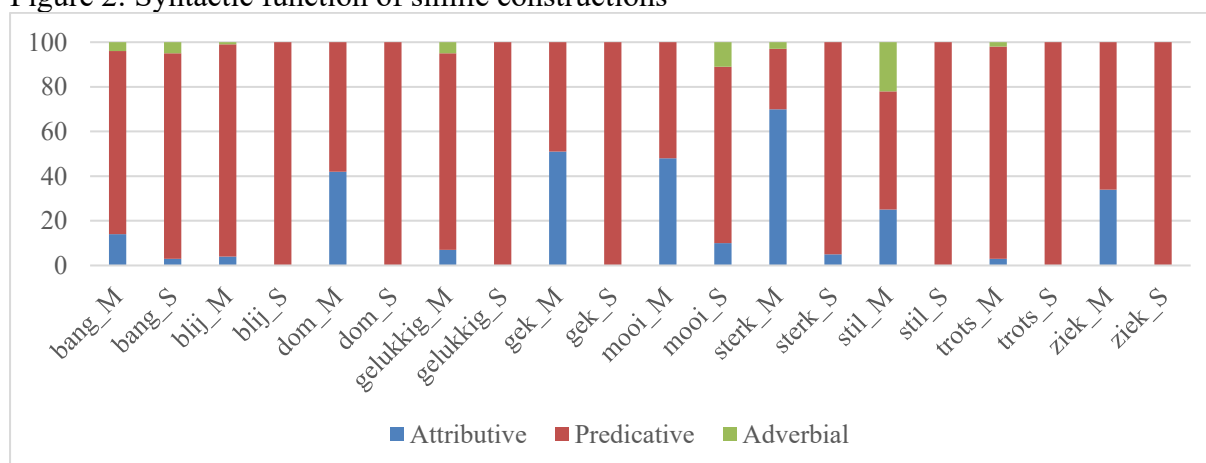
For simile compounds with *gek* ‘crazy’, we observe an interesting syntactic split: *knotsgek* ‘lit. club-crazy’ is predominantly used in attributive position (e.g., *een knotsgekke wedstrijd* ‘a completely crazy match’), whereas *stapelgek* ‘lit. pile-crazy’ typically occurs in predicative position (e.g., *Alie was stapelgek op Emma* ‘Alie was completely crazy about Emma’).

Finally, in both types of similes, adverbial use can also be marginally observed (see examples 25–26).

- (25) *Ik schaats zo mooi als een patrijs* ‘I skate as beautifully as a partridge’

- (26) *Niet in de laatste plaats vanwege de hybride aandrijving die muisstil varen op een elektromotor mogelijk maakt.* ‘Not least because of the hybrid drive, which makes whisper-quiet (lit. mouse-quiet) sailing on an electric motor possible.’

Figure 2: Syntactic function of simile constructions



6.2 Semantic properties

A comparison of the semantic profiles of the reference points in syntactic and morphological similes reveals another interesting pattern: syntactic similes predominantly combine with animate reference points, with the categories PERSON and ANIMAL accounting for up to 72%. In contrast, morphological similes typically involve non-animate reference points, as the categories OBJECT, CONCEPT, and PROPERTY together represent 85% of the dataset examples.

A detailed overview of the semantic categories associated with each item is provided in Figure 3.

Figure 3: Semantic categories of simile constructions

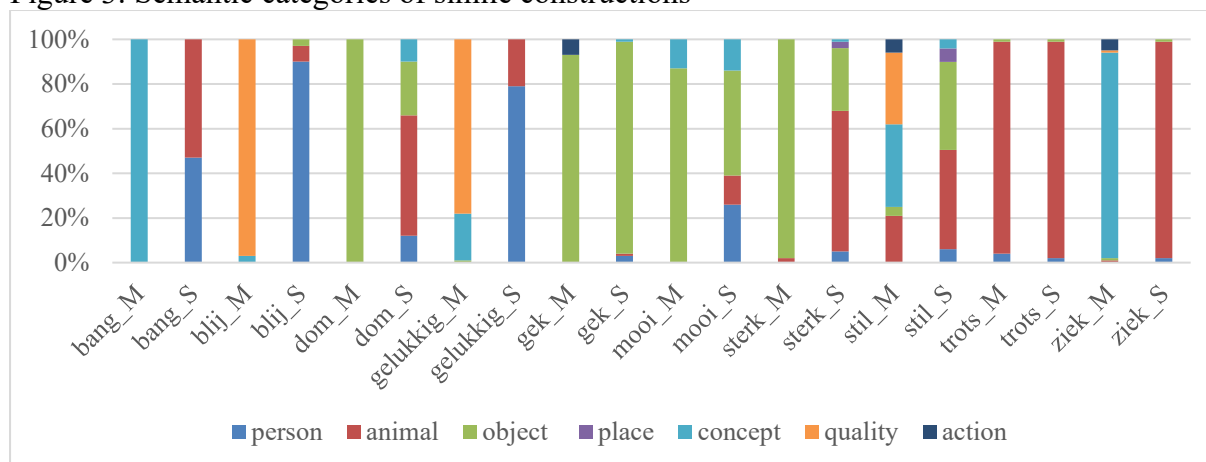


Figure 3 reveals notable differences across all 10 adjectives, and across morphological and syntactic patterns, even when based on the same adjective. For example, the adjectives *bang* ‘afraid’ and *blij* ‘happy’ typically combine with animate reference points in syntactic similes, i.e. categories like PERSON or ANIMAL account for most cases (e.g., *zo bang als een wezel* ‘as scared as a weasel’; *zo blij als een kind* ‘as happy as a child’). In contrast, their morphological counterparts often involve non-animate semantics, such as CONCEPT or QUALITY (e.g., *doodsbang* ‘lit. dead-afraid’; *dolblij* ‘lit. crazy-happy’). Similarly, *dom* ‘dumb’ appears in syntactic similes in most cases with a modifier referring to an ANIMAL (e.g., *zo dom als een ezel* ‘as stupid as a donkey’), whereas morphological similes with the same adjective exclusively combine with the OBJECT category (e.g., *oliedom* ‘lit. oil-dumb’).

There are, however, some exceptions where the reference points included in both patterns share similar semantic profiles. For instance, similes with *trots* ‘proud’ typically refer to ANIMALS. However, within this semantic field, more lexical variation is found in the syntactic type (cf. *zo trots als een pauw/aap/haan/leeuw* ‘as proud as a peacock/monkey/rooster/lion’ vs *apetrots* ‘monkey-proud’).

A final trend that can be drawn from Figure 3 is that the syntactic similes generally exhibit greater semantic variation, while morphological similes typically show more specialization and consistency in the semantics of their reference points.

6.3 Productivity

To compare the productivity of both patterns, we calculated two measures: the type-token ratio (TTR) and the hapax-token ratio, also known as potential productivity (PP) (see Baayen 2009).

In general, both phrasal and morphological similes score low on these measures, indicating limited lexical variation and a relatively low degree of creative potential. However, the syntactic simile patterns overall show higher productivity scores than the morphological ones, as shown in Table 3.

Table 3: Productivity scores of simile constructions (all items merged)

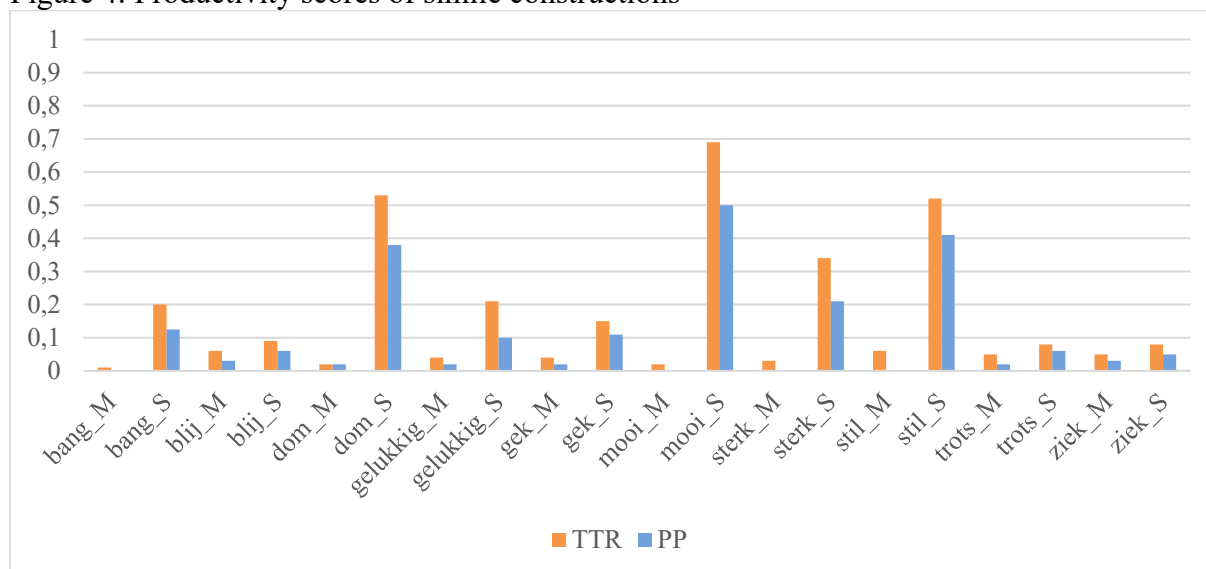
	Syntactic simile	Morphological simile
TTR	0.29	0.04
PP	0.20	0.01

Figure 4 presents the productivity scores for each item. Again, the data reveal not only construction-specific differences but also variation linked to individual adjectives. The highest scores occur in the phrasal similes with *mooi* ‘beautiful’, *dom* ‘dumb’, *stil* ‘quiet’, and *sterk* ‘strong’. These cases show relatively broad lexical variation in reference points and a certain degree of creative potential, as evidenced by hapax formations (see examples 27–28).

- (27) *Niets zo mooi als een lege pagina die gevuld mag worden* ‘Nothing as beautiful as a blank page waiting to be filled’
- (28) *In mijn ogen is dit net zo dom als een schaakpartij waarbij men de tegenstander de volgende 10 zetten meedeelt.* ‘In my view, this is just as foolish as a chess game in which one tells the opponent the next 10 moves.’

The lowest productivity scores occur in simile compounds with *bang* ‘afraid’, *blij* ‘happy’, and *dom* ‘dumb’, which tend to form (semi-)lexicalized compounds: *doodsbang* ‘lit. dead-afraid’ ($n = 100/100$), *dolblij* ‘lit. crazy-happy’ ($n = 94/100$), and *oliedom* ‘lit. oil-dumb’ ($n = 92/100$).

Figure 4: Productivity scores of simile constructions



7 Discussion

To summarize our corpus findings, syntactic and morphological similes differ significantly in structural properties, semantics, and productivity. In terms of form, syntactic similes exhibit greater structural flexibility, allowing the inclusion of complex noun phrases, and are predominantly used in a predicative function. Conversely, morphological similes allow more variation across the parts of speech they combine with (i.e., nouns, adjectives, verbs, and prepositions) and display broader syntactic versatility, functioning not only predicatively but also attributively and adverbially. From a semantic perspective, syntactic similes primarily draw on animate categories for the point of comparison and often demonstrate considerable semantic variation. Morphological similes, however, tend to involve non-animate semantic domains and are characterized by a higher degree of semantic specialization. Finally, regarding productivity, syntactic similes show medium productivity, whereas morphological similes are markedly less productive, often restricted to fixed or conventionalized combinations. These contrasts highlight the distinct linguistic behavior and functional scope of the two simile types.

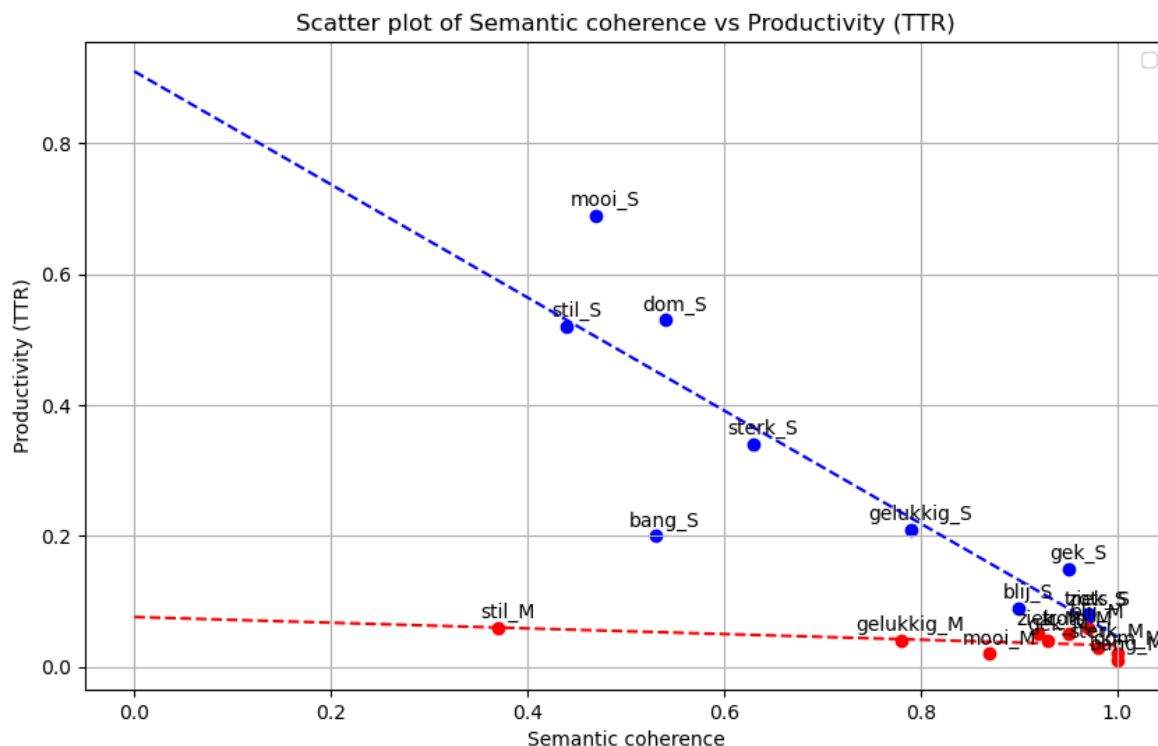
Interestingly, in the earlier case study on the Evaluative Binominal Construction (Section 3), we observed parallel differences between the syntactic and morphological variants. The syntactic variant also permitted more complex fillers, exhibited a similar preference for animate fillers, and demonstrated higher productivity compared to its morphological counterpart.

Moreover, similar to what we observed in the EBC patterns, both types of simile constructions differ not only at the schematic level but also at the level of semi-filled patterns. For instance, in syntactic similes such as *zo sterk als een beer/leeuw/paard...* ‘as strong as a bear/lion/horse...’ *sterk* ‘strong’ is typically used in a literal sense to describe human strength, whereas the morphological simile *ijzersterk* ‘iron-strong’, although based on the same adjective, more often carries a figurative meaning, as in *een ijzersterk argument* ‘an iron-strong argument’.

To further examine the distinction between syntactic and morphological simile-based constructions across the different patterns, we will now focus on the relationship between their productivity and semantic behavior. Barðdal (2008) proposes an inverse correlation between semantic coherence, i.e., the internal consistency among the components of a construction, and productivity. Specifically, high type frequency would reduce the need for semantic coherence, whereas low type frequency would require greater coherence for a construction to remain extendable (Barðdal 2008: 34). To investigate this relationship in the case of simile constructions, we plotted type-token ratio (TTR) scores against a measure of semantic coherence, defined here as the proportion of slot-filler lemmas belonging to the largest semantic category within each pattern. For example, the adjective *mooi* is most frequently associated with the semantic category OBJECT in both syntactic and morphological similes, but to different degrees. In syntactic similes, *mooi* combines with objects in 47% of cases (e.g., *zo mooi als een plaatje* ‘as beautiful as a picture’), whereas in morphological similes, this proportion rises to 87% (e.g., *bloedmooi* ‘blood-beautiful’). These proportions yield semantic coherence scores of 0.47 for *mooi_S* and 0.87 for *mooi_M*.

Figure 5 represents in blue all syntactic simile patterns investigated, and in red all morphological items. The blue and red dashed lines are the regression lines that connect the syntactic and morphological patterns.⁵

Figure 5: Scatter plot of semantic coherence vs productivity (TTR)



The results shown in Figure 5 confirm the contrasting patterns for syntactic and morphological similes. For syntactic similes, the Spearman correlation coefficient indicates a strong negative association between the two variables ($r_s = -0.86$, $p = .001$), which is statistically significant. This suggests that as productivity increases, semantic coherence decreases, and vice versa, in line with Barðdal's (2008) prediction that high type frequency reduces the need for semantic consistency. In contrast, morphological similes display a moderate negative correlation ($r_s = -0.50$), but this relationship is not statistically significant ($p = .145$).⁶ These findings imply that while syntactic similes conform to the expected inverse correlation between productivity and semantic coherence, morphological similes do not exhibit a clear pattern. As the plot clearly shows, this is due to their overall lower type-token ratio and high degree of semantic coherence.

As regards syntactic expressions, there is a strong negative correlation between productivity and semantic coherence in simile formation. Highly productive similes, such as those following the schema [*zo mooi/dom/stil als een N*] ‘as beautiful/dumb/quiet as a N’, exhibit low semantic specialization and allow for considerable creative expansion. Example (29) demonstrates the lexical and semantic flexibility of this schema.

⁵ Data visualization in Figure 5 was generated using Microsoft Copilot (2025), which facilitated automated chart creation and formatting, based on our TTR and semantic coherence scores calculated in Excel.

⁶ We report the Spearman correlation coefficient because this non-parametric method does not require a linear relationship or normally distributed data. The following online tool was used to compute the coefficient and the p values: <https://www.gigacalculator.com/calculators/correlation-coefficient-calculator.php>.

- (29) *zo dom als (het achtereind van) een koe/varken/ezel/rund/kwartel/gans/garnaal/deur/steen/schaakpartij/(...)* ‘as dumb as (the backend of) a cow/pig/donkey/ox/quail/goose/ shrimp/door/stone/chess game/(...)’

In contrast, syntactic similes with low productivity, such as [*zo trots/blij/gek als een N*] ‘as proud/happy/crazy as a N’, tend to be semantically specialized and often semi-lexicalized, as in *zo blij als een kind* ‘as happy as a child’. While these expressions are less open to lexical variation, creative expansion is still possible within the same semantic category, for instance *zo blij als een kind in een snoepwinkel/in een speelgoedwinkel/op Pakjesavond* ‘as happy as a child in a candy store/in a toy store/on Sinterklaas Eve’.

Unlike phrasal similes, morphological similes show no significant correlation between productivity and semantic coherence, and overall, their productivity is low. Most combinations exhibit high semantic coherence and tend to be lexicalized, as in *doodsbang* ‘lit. dead-afraid’, *dolblij* ‘lit. crazy-happy’, *ijzersterk* ‘lit. iron-strong’, and *apetrots* ‘lit. monkey-proud’. An exception is the schema [*X-stil*], which displays the greatest semantic variation, and hence the lowest degree of semantic coherence: *muisstil* ‘mouse-quiet’ (ANIMAL), *stokstil* ‘stick-quiet’ (OBJECT), *doodstil* ‘dead-quiet’ (CONCEPT), and *fluisterstil* ‘whisper-quiet’ (ACTION).

Over time, many of these formations have undergone bleaching of the original comparative interpretation, leading to prefixoidization into intensifiers (Booij 2010; Hoeksema 2012; Van Goethem 2020) (see Section 2). This process has resulted in semantically opaque combinations from a synchronic point of view, such as *bloedmooi* ‘lit. blood-beautiful’, *oliedom* ‘lit. oil-dumb’, *stapelgek* ‘lit. pile-crazy’, and *reuzeblij* ‘lit. giant-happy’, including taboo-based intensifiers, such as *strontgelukkig* ‘lit. shit-happy’ and *schijtziek* ‘lit. shit-sick’.⁷

8 Conclusion

Our study has revealed that comparable tendencies arise in both Evaluative Binominal Constructions and Simile-based Constructions. Phrasal evaluative constructions are generally more flexible in terms of complexity and semantic variation, which results in higher productivity and greater creativity. By contrast, evaluative compounds, although frequently attested in corpora, tend to involve (semi-)lexicalized combinations and grammaticalized affixoids, limiting their variability. Furthermore, phrasal patterns typically attract animate slot fillers, whereas compound patterns favor inanimate fillers, pointing to a functional division of labour between syntactic and morphological evaluative constructions. Hence, at an abstract level, these patterns exhibit complementary distribution: while both types of constructions can express similar evaluative meanings, they differ significantly in their structural properties and functional tendencies.

⁷ While some of these intensifying prefixoids remain productive (e.g., *reuze-*; Van Goethem & Hiligsmann 2014), others occur only in a single or very limited set of adjectival compounds and have become fully opaque. One such example is *oliedom* ‘lit. oil-dumb’, whose etymology is not entirely clear. According to *Etymologiebank*, *oliedom* likely originates from the now obscure expression *olijk dom*. The adjective *olijk*, still known today in the sense of ‘playful’, formerly conveyed the meaning ‘bad in its kind’ (<https://etymologiebank.nl/trefwoord/oliedom>). Another example that has become opaque is *stapelgek* ‘lit. pile-crazy’, where *stapel* originally referred to a cricket or a grasshopper (<https://etymologiebank.nl/trefwoord/stapelgek>).

Nevertheless, our corpus analysis also indicates that, at the level of the individual constructions, the relationship is much more nuanced. There is sometimes overlap and even competition: for example, both types of similes based on the adjective *trots* ‘proud’ typically pair with the ANIMAL category. More frequently, however, we observe a consistent pattern of divergence, as observed in the mostly attributive and metaphorical use of *ijzersterk* ‘iron-strong’, compared to the predominantly predicative and literal use of the simile *zo sterk als een beer* ‘as strong as a bear’.

This interplay of convergence and divergence between phrasal and morphological expressions highlights the dynamic nature of evaluative language, where creative expression may be counterbalanced by structural constraints and lexicalization.

References

- Amiot, Dany & Dugas, Edwige. 2020. Combining forms and affixoids in morphology. In Aronoff, Mark (ed.), *The Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.
- Baayen, R. Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Lüdeling, Anke & Kytö, Merja (eds.), *Corpus linguistics: An international handbook*, 899–919. Berlin: Walter de Gruyter.
- Barðdal, Jóhanna. 2008. *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.
- Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.
- Booij, Geert. 2015. Adjectival compounds. *Taalportaal*.
<https://taalportaal.org/taalportaal/topic/pid/topic-13998813295511067>.
- Booij, Geert. 2019. Compounds and multi-word expressions in Dutch. In Schlücker, Barbara (ed.), *Complex lexical units: Compounds and multi-word expressions*, 95–126. Berlin: De Gruyter.
- Broekhuis, Hans. 2015. Modification by a complex intensifying phrase. *Taalportaal*.
<https://taalportaal.org/taalportaal/topic/pid/topic-14286624810936411>.
- Broekhuis, Hans. 2020. The *N van een N* ‘N of a N’ construction. *Taalportaal*.
https://www.taalportaal.org/taalportaal/topic/link/syntax___Dutch__np_n4__nouns4_binominal.4.2.1.xml.
- Dressler, Wolfgang U. & Merlini Barbaresi, Lavinia. 1994. *Morphopragmatics: Diminutives and intensifiers in Italian, German, and other languages*. Berlin: Mouton de Gruyter.
- Everaert, Martin. 1992. Nogmaals: een schat van een kind. In Bennis, Hans & De Vries, Jan W. (eds.), *De binnenbouw van het Nederlands: Een bundel artikelen voor Piet Paardekooper*, 45–54. Dordrecht: ICG Publications.
- Finkbeiner, Rita & Schlücker, Barbara. 2019. Compounds and multi-word expressions in the languages of Europe. In Schlücker, Barbara (ed.), *Complex lexical units: Compounds and multi-word expressions*, 1–44. Berlin: De Gruyter.
- Foolen, Ad. 2004. Expressive binominal NPs in Germanic and Romance languages. In Radden, Günter & Panther, Klaus-Uwe (eds.), *Studies in linguistic motivation*, 75–100. Berlin: Mouton de Gruyter.

- Haeseryn, Walter & Romijn, Kirsten & Geerts, Guido & de Rooij, Jaap & van den Toorn, Maarten C. 2019. Het type ‘een schat van een kind’. *Algemene Nederlandse Spraakkunst*. <https://e-ans.ivdnt.org/topics/pid/ans1405030602lingtopic>.
- Haspelmath, Martin & Buchholz, Oda. 1998. Equative and similative constructions in the languages of Europe. In Van der Auwera, Johan (ed.), *Adverbial constructions in the languages of Europe*, 277–334. Berlin: De Gruyter.
- Hendriks, Isa & Van Goethem, Kristel & Meunier, Fanny & Hiligsmann, Philippe. 2017. Language-specific tendencies towards morphological or syntactic constructions: A corpus study on adjective intensification in L1 Dutch, L1 French and L2 Dutch. *Nederlandse Taalkunde/Dutch Linguistics* 22(3). 389–420.
- Hoeksema, Jack. 2012. Elative compounds in Dutch: Properties and developments. In Oebel, Guido (ed.), *Intensivierungskonzepte bei Adjektiven und Adverbien im Sprachenvergleich*, 97–142. Hamburg: Kováč Verlag.
- Hyvärinen, Irma Kaarina. 2019. Compounds and multi-word expressions in Finnish. In Schlücker, Barbara (ed.), *Complex lexical units: Compounds and multi-word expressions*, 307–336. Berlin: De Gruyter.
- Israel, Michael & Harding, Jennifer Riddle & Tobin, Vera. 2004. On simile. *Language, Culture, and Mind* 100. 123–135.
- Kilgariff, Adam & Baisa, Vít & Bušta, Jan & Jakubíček, Miloš & Kovář, Vojtěch & Michelfeit, Jan & Rychlý, Pavel & Suchomel, Vít. 2014. *The Sketch Engine: Ten years on*. *Lexicography* 1. 7–36. <http://www.sketchengine.eu/>.
- Masini, Francesca. 2019. Compounds and multi-word expressions in Italian. In Schlücker, Barbara (ed.), *Complex lexical units: Compounds and multi-word expressions*, 153–188. Berlin: De Gruyter.
- Norde, Muriel & Van Goethem, Kristel. 2014. Bleaching, productivity and debonding of prefixoids: A corpus-based analysis of ‘giant’ in German and Swedish. *Linguisticae Investigationes* 37(2). 256–274.
- Paardekooper, Piet C. 1956. Een schat van een kind. *De Nieuwe Taalgids* 49. 93–99.
- Rainer, Franz. 2015. Intensification. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *Word-formation: An international handbook of the languages of Europe*, 1339–1351. Berlin: De Gruyter Mouton.
- Schlücker, Barbara. 2019. Compounds and multi-word expressions in German. In Schlücker, Barbara (ed.), *Complex lexical units: Compounds and multi-word expressions*, 69–94. Berlin: De Gruyter.
- Van Goethem, Kristel. 2020. Affixation in morphology. In Lieber, Rochelle (ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.
- Van Goethem, Kristel. 2023. Een pracht van een constructie vs een prachtconstructie: Een vergelijkende corpusstudie van de syntactische en morfologische expressieve binominale constructie. *Nederlandse Taalkunde/Dutch Linguistics* 28(3). 325–349.
- Van Goethem, Kristel & Hiligsmann, Philippe. 2014. When two paths converge: Debonding and clipping of Dutch reuze ‘lit. giant; great’. *Journal of Germanic Linguistics* 26(1). 31–64.

Verhagen, Arie. 2005. Constructiegrammatica en ‘usage-based’ taalkunde. *Nederlandse Taalkunde* 3(4). 197–222.

Wouden, Ton van der & Foolen, Ad. 2017. A most serious and extraordinary problem: Intensification of adjectives in Dutch, German, and English. *Leuvense Bijdragen* 101. 82–100.

Kristel Van Goethem
F.R.S.-FNRS & UCLouvain
Institute of Language and Communication
1348 Louvain-la-Neuve, Belgium
e-mail: kristel.vangoethem@uclouvain.be