

# Afrikaans epithetical, suffixoidal, metonymic, exocentric attributive compounds

---

**Gerhard B. van Huyssteen** (North-West University, Potchefstroom, South Africa)  
*gerhard.vanhuysssteen@nwu.ac.za*

**Warning:** This article contains examples of language that might be offensive to some users.

**Disclaimer:** Our annotations (e.g., glossing and translation) of racist, homophobic, and other offensive constructions are based on either dictionary data, or our subjective understanding and analysis of their context and intended meaning, and does not imply our endorsement or agreement with such views.

## Abstract

Many languages in the world, including Afrikaans, use exocentric compounds to form personal names, like French *garde-malade* (watches-sick.person 'nurse'). These kinds of compounds are also sometimes in the literature known as bahuvrihi or possessive compounds (e.g., Booij 2020), and viewed by some scholars (e.g., Schlücker 2023) as not truly exocentric. Despite Scalise, Fábregas, and Forza (2009, 52) finding that "Germanic languages are not significantly characterised by exocentric compounding", numerous studies of individual Germanic languages have been published in the past. For example, Van Niekerk (2002, 2006) presented a comprehensive, corpus-based, functional description of exocentric compounds in Afrikaans.

A subcategory of these compounds can be used evaluatively as personal names – either pejoratively as epithets (for example *geld+wolf*), or melioratively as hypocoronyms, laudatives, or approbatives (for example *werk+esel*) (see Van Niekerk 2006, 132–166). In this paper, the focus is on two metonymic subcategories, namely: (1) BODY PART FOR PERSON, such as *slap÷gat* (soft÷arse 'quitter, slacker'), *groot÷bek* (large÷beak 'braggart'), or *klip÷kop* (rock÷head 'person with limited cognitive abilities'); and (2) GARMENT FOR PERSON, such as *vet÷sak* (fat÷bag 'fatso'), *dronk÷lap* (drunk÷cloth 'drunkard'), or *klets÷kous* (chat÷sock 'chatterbox'). The aim of this paper is to characterise this construction based on an annotated dataset of EECs. The paper concludes with ideas to refine and re-use this dataset in future research.

**Keywords:** Afrikaans, Dutch, English, Germanic language, approbative, bahuvrihi, comparative linguistics, construction morphology, epithet, evaluative language, exocentric compound, hypocoronym, laudative, meliorative, metonymy, pejorative, personal names, possessive compound

---

## How to cite this article:

Van Huyssteen, Gerhard B. 2023. "Afrikaans epithetical, suffixoidal, metonymic, exocentric attributive compounds". In *Proceedings of the 4<sup>th</sup> International Afrikaans Grammar Workshop*, edited by Adri Breed. Potchefstroom: North-West University. pp. 17-38. DOI:10.25388/nwu.25045796



**Copyright:**  
© 2023 Van Huyssteen.  
Licensed via [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)

## Opsomming

### *Afrikaanse epitetiese eksosentriese komposita*

Vele tale in die wêreld, insluitend Afrikaans, gebruik eksosentriese komposita om persoonsname te vorm, soos Franse *garde-malade* (oppas-sieke 'verpleër'). Hierdie soort komposita staan ook soms in die literatuur bekend as bahuvrihi- of besitlike komposita (Booij 2020) en word deur sommige geleerdes (bv. Schlücker 2023) nie as werklik eksosentriese beskou nie. Ten spyte van Scalise, Fábregas en Forza (2009:52) wat bevind het dat “Germanic languages are not significantly characterised by exocentric compounding”, is talle studies van individuele Germaanse tale in die verlede gepubliseer. Van Niekerk (2002, 2006) het byvoorbeeld 'n omvattende, korpusgebaseerde, funksionele beskrywing van eksosentriese komposita in Afrikaans aangebied.

'n Subkategorie van hierdie verbindings kan evaluerend as persoonsname gebruik word – óf pejoratief as epitetons (byvoorbeeld *geld+wolf*), óf melioratief as hipokoronieme, lofprysings, of goedkeurings (byvoorbeeld *werk+esel*) (kyk Van Niekerk 2006:132–66). In hierdie referaat word op twee metonimiese subkategorieë gefokus, naamlik: (1) LIGGAAMSDEEL VIR PERSOON, soos *slap÷gat*, *groot÷bek*, en *klip÷kop*; en (2) KLEDINGSTUK VIR PERSOON, soos *vet÷sak*, *dronk÷lap* of *klets÷kous*. Die doel van hierdie artikel is om hierdie konstruksie te beskryf aan die hand van 'n geannoteerde datastel. Die artikel sluit af met idees om die datastel te verfyn en in toekomstige navorsing te hergebruik.

**Sleutelwoorde:** Afrikaans, Engels, Germaanse taal, Nederlands, bahuvrihi, besitlike samestelling, eksosentriese kompositum, evaluatiewe taal, goedkeuring, hipokoroniem, konstruksiemorfologie, lofprying, melioratief, metonimie, pejoratief, persoonsname, skelnaam, spotnaam, vergelykende taalkunde

## 1 Introduction

Many languages in the world, including Afrikaans, use exocentric compounds to form personal names. Compare some of the examples from Štekauer, Valera, and Kórtvélyessy (2012, 79-88):<sup>1</sup>

- |   |                 |   |
|---|-----------------|---|
| 1 | <b>French:</b>  | <i>garde-malade</i><br>watches-sick.person<br>'nurse'                   |
| 2 | <b>Chinese:</b> | <i>lǐnglù</i><br>lead-road<br>'guide'                                   |
| 3 | <b>Greek:</b>   | <i>κοκκιν·ο+μάλλης</i><br>red·LK+hair<br>'person with red hair, ginger' |

---

<sup>1</sup> The middle dot (·) is used in morphological analyses to indicate affix boundaries, the division sign (÷) to indicate affixoid boundaries, and the plus sign (+) to indicate word/stem boundaries in compounds. A list of abbreviations is available at <http://bit.ly/taalportaal-afr>.

These kinds of compounds are also sometimes in the literature known as *bahuvrihi* or possessive compounds (Booij 2020), and viewed by some scholars (e.g., Schlücker 2023) as not truly exocentric. This “confusion” can also be seen in the contradiction that some scholars claim that “Germanic languages are not significantly characterised by exocentric compounding” (Scalise, Fábregas, and Forza 2009, 52) while numerous studies of exocentric compounds in individual Germanic languages have been published in the past. For example, Van Niekerk (2002, 2006) presented a comprehensive, corpus-based, functional description of exocentric compounds in Afrikaans.

To resolve this conflict of views, we firstly define and demarcate the construction under discussion in this paper, viz. epithetical, suffixoidal, metonymic, exocentric attributive compounds (*EECs* for short) (§ 2). Thereafter, we describe a dataset of EECs that were extracted from Van Niekerk (2002, 2006), and verified using recent corpus data (§ 3). This dataset is freely available in the open-data domain, and could be used for future research.<sup>2</sup>

The main aim of the paper is to provide a comprehensive, all-inclusive characterisation and formalisation of the core set of EECs in this dataset, which could serve as basis to formulate hypotheses for future, corpus-based research. For these descriptive purposes, we use conventions from various flavours of construction morphology, as briefly explained in § 0. The phonological form (§ 5), morphosyntactic form (§ 6), morphosyntactic function (§ 7), and stratal restrictions (§ 8) of EECs are then described, before concluding with an extensive agenda for future research (§ 9). Note that all aspects related to the meaning and usage of the construction fall outside the scope of this paper and will therefore only be included in said research agenda.

## 2 Definition and demarcation

The perceived confusion mentioned above, can be resolved by and large by a carefully articulated taxonomy of compounding. For example, within Bauer, Lieber, and Plag’s (2013) descriptive framework and taxonomy for compounds, the main distinction is between argumental<sup>3</sup> and non-argumental compounds, with the latter further dividing into coordinative and **attributive compounds**; see Figure 1. All these main types of compounds can be exocentric; however, exocentric argumental compounding (e.g., *pickpocket*, *turncoat*, and *rotgut*) is not productive in English, and exocentric coordinative compounding (e.g., *father-daughter [relationship]*, and *London-New York [flight]*) mainly functions to create premodifiers of other nouns. It is only **exocentric** attributive compounding (e.g., *blockhead*, *cheesehead*, and *redhead*) that is highly productive in English (Bauer, Lieber, and Plag 2013, 473–474, 478–479). They also point out that it is only exocentric attributive compounds that are called **bahuvrihi** or **possessive compounds**.

---

<sup>2</sup> Available at <https://gerhard.pro/files/DataVanHuyssteen2023a-1.xlsx>.

<sup>3</sup> Also called **subordinate compounds** in other frameworks.

One of the key conceptual motivations of such possessive compounds, is that they are all based on metonymies or metaphors (Bauer, Lieber, and Plag 2013, 478). In addition to the extensional range of these compounds (i.e., the category of entities designated by each compound, such as people, physical objects, plants, animals, etc.), they could be categorised according to the different degrees of **metonymic-metaphorical** interactions (Barcelona 2011, 157–165), for example:

- [-metonymic], [+metaphoric], e.g., *agter+os* (behind+ox 'hind ox (lit.); slogger, tail-ender (fig.)')
- [±metonymic], [-metaphoric], e.g., *blou+blas+ie* (blue+bubble·DIM 'bluebottle');
- [+metonymic], [±metaphoric], e.g., *eend÷stert* (duck÷tail 'tailfeathers of a duck (lit.) > hairstyle in the 1950s resembling the tailfeathers of a duck (metaphor) > person with a hairstyle in the 1950s resembling the tailfeathers of a duck (metonym) > riffraff, lowlife);
- [+metonymic], [-metaphoric], e.g., *vloek÷bek* (swear+beak 'person who habitually swears').

The focus in this article is only on the latter two types, i.e., **metonymic, exocentric attributive compounds**. Since it is often a matter of argumentation or imagination to distinguish between cases that are non-metaphoric (like *vloek÷bek*), and cases where some metaphoric interaction might be possible (like *eend÷stert*), we include here all instances where the metonymic conceptualisation is predominant and undeniable.

As **personal names** such metonymic, exocentric attributive compounds can be used either melioratively as hypocorisms, laudatives, or approbatives (for example *petrolhead*), or pejoratively as **epithets** (for example *dick head*), or **epithetical modifiers** (e.g., *dick-head attitude*). The focus here, however, is only on those metonymic, exocentric attributive compounds that are used as epithets (hence **EECs**, for **epithetical exocentric compounds**). Simplified, EECs are thus defined as morphological constructions (i.e., compounds) whose semantic head is not included in the word (i.e., exocentric), and which are used metonymically as pejorative personal names (i.e., epithets) to refer to either an antagonist, or outgroup (Brewer 2016).<sup>4</sup>

One last demarcation is required. The righthand constituent (i.e., the morphological head) of EECs can be either ordinary stems that are directly related to the attribute (e.g., a *redhead* is a 'person with a red head / red hair'), or suffixoids. A suffixoid is a form that has developed

---

<sup>4</sup> The difference between **protagonist/antagonist** and **ingrouper/outgroup** has mostly to do with agency. An antagonist is commonly seen as an opponent who opposes or attacks you (i.e., with agency), while an outgroup is merely a member of another social category, an outgroup (i.e., without agency necessarily as part of its construal). Oftentimes epithets can be used to describe both antagonists and outgroupers – simply depending on the context.

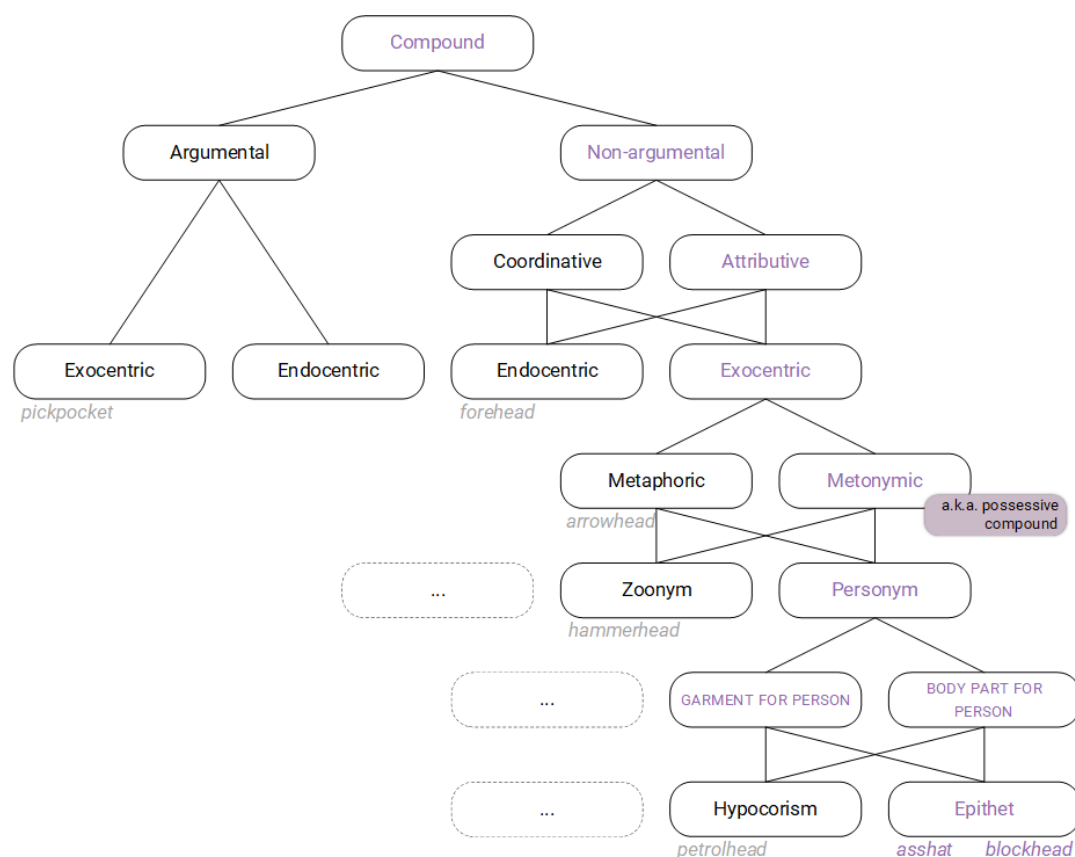


Figure 1: Partial taxonomy to illustrate demarcation (in purple)

bound uses in addition to unbound uses (Amiot and Dugas 2020; Booij and Hüning 2014), and is therefore halfway between a suffix (like *-ing* in *annoy-ing* that only has a bound use) and a word (like *annoy* that only has an unbound use) (Bauer 2018). A typical example of an English suffixoid is *-free* in *sugar-free* and *caffeine-free*, where *-free* has developed the new meaning ‘without X’, in addition to its multitude of original, unbound meanings (e.g., ‘able to move in an unrestricted manner’, or ‘chemically uncombined’). From all the English examples above, we can deduce that the word *head* (‘uppermost part of a vertebrate’) has constructionalised into the suffixoid *-head* (‘antagonist; outgroup’), which is only found with this specialised meaning in compounds like *block-head*, *cheese-head*, and *dick-head*.

In this paper, the focus is on two metonymic subcategories of such suffixoids, namely BODY PART FOR PERSON (see 4), and GARMENT FOR PERSON (see 5).

<b>4 BODY PART FOR PERSON</b>		
<i>pampoen+kop</i>	<i>dronk+gat</i>	<i>poes+neus</i>
pumpkin+head	drunk+arse	cunt+nose
‘simpleton’	‘drunkard’	‘fuck nut’
<b>5 GARMENT FOR PERSON</b>		
<i>vet+sak</i>	<i>dronk+lap</i>	<i>klets+kous</i>
fat+bag	drunk+cloth	chat+sock
‘fatso’	‘drunkard’	‘chatterbox’

Figure 1 provides a visual representation of the relevant parts (in purple) of this taxonomy. All examples are from Bauer, Lieber, and Plag (2013).

Given the above characterisation of EECs, the following are related but contrastive types of words, and are hence not considered in this article:

- epithetical **non-compounds**, such as simplexes (6), and affixed words (7);
- epithetical **endocentric** compounds (whose semantic head is included in the word) (8);
- epithetical **endocentric** compounds, used **metaphorically**,<sup>5</sup> e.g., compounds with personal names (9), or animal names (10) as their heads;
- **neutral endocentric** compounds that are also occasionally (secondarily) used as EECs (11);
- **neutral** metonymic, exocentric compounds, used **rarely** as EECs (12);<sup>6</sup> and
- **hypocoristic** metonymic, exocentric compounds, used melioratively (13).

6	<i>ghwar</i> 'person with limited cognitive abilities'	<i>mamparra</i> 'person with limited cognitive abilities'
7	<i>vett·ie</i> fat·DIM 'corpulent person'	<i>lamm·eling</i> weak·NMLZ 'weakling, coward'
8	<i>bed+piss·er</i> bed+piss·NMLZ 'bed-wetter'	<i>hoer+kind</i> whore+child 'child of a whore'
9	<i>geld+piet</i> money+Piet (typical Afrikaans male name) 'fat cat'	<i>plaas+japie</i> farm+Japie (typical Afrikaans male name) 'country bumpkin, yokel'
10	<i>nag+wolf</i> night+wolf 'night owl > person who stays up late at night'	<i>nes+kuiken</i> nest+chicken 'nestling > youngest child'
11	<i>vark+nek</i> pig+neck 'pig's neck, pork neck; asshole'	<i>wyn+sak</i> wine+bag 'wine bag; drunkard'
12	<i>kaal+kop</i> naked+head '(person with) bald head'	<i>lang+been</i> long+leg '(person with) long leg(s)'
13	<i>pop+lap</i> pop+cloth 'sweetheart'	<i>waag+hals</i> dare+neck 'dare devil'

<sup>5</sup> Bauer, Lieber, and Plag (2013, 478) indicate that these types of compounds can easily be analysed as either endocentric, or exocentric; we choose the former analysis. Also see Bauer (2017, 68–71), De Haas and Trommelen (1993, 413), and Van Niekerk (2002, 26–27) for discussions of these two possible analyses.

<sup>6</sup> In these types of compounds, paraphrases starting with "person with" or "person has" are the default, primary interpretation, e.g., a *baldhead* is a 'person with a bald head'. That it can be used pejoratively as an epithet, is not inherent to the compound. This is in contrast with a "true" EECs like *cheesehead* that cannot be paraphrased as \*'person with a cheese head', or \*'person who has a head made of / filled with cheese'.

### 3 EEC dataset

The main sources of examples are Van Niekerk's MA dissertation (2002), and PhD thesis (2006). The former is based on dictionary and corpus data, while the latter on dictionary data only. All examples were extracted automatically from the PDFs (n = 801 and n = 3,731 respectively, resulting in a total of n = 4,532 unique lemmas).<sup>7</sup> Where possible, the dictionary definitions and/or meanings and usage categories assigned by Van Niekerk were also extracted with the relevant lemma. From this large set of exocentric compounds all potential epithets were extracted (n = 922). In addition to information available from the two sources (e.g., page references, dictionary definitions, and a few example sentences), the dataset was manually enriched by broadly categorising each word in terms of:

- **promiscuity**: whether it is used exclusively (14, 15), primarily (16), or only secondarily (17) as an epithet; and if it is promiscuous, whether it has multiple functions (16), meanings (17), or connotations (18);
- **connotation**: whether it is pejorative (14), racist (15), meliorative (19), neutral (20), or combinations of these (18);
- **name type**: whether it is an epithet (14, 15), a hypocoronym (19), a descriptor (20), or a combination of these (18);
- **nature of the righthand (RH) component**: whether it is a suffixoid (14 to 20), stem (21), derived stem (22), phrase (23), etc.;
- **constructional types**: e.g., exocentric compounding (14) vs. endocentric compounding (20), parasynthetic compounding (22), folk etymology (24), etc.; and metonymisation (14 to 20) vs. metaphorisation (21); and
- **corpus attestation**: whether any corpus results for the word could be found (14 to 20), or not (23).

14 *dronk+gat*

drunk+hole  
'drunkard'

15 *rooi+huid*

red+skin  
'indigenous American'

16 *wind+gat*

wind+arse  
'braggart (N); boastful(ly) (ADJ, ADV)', and rarely 'ventilation opening (N)'

17 *aap+kop*

monkey+head  
'monkey's head; unpleasant person'

18 *sproet+gesig*

freckle+face

---

<sup>7</sup> We use **n** for types, and **N** for tokens.

	'person with freckles; face with freckles' (used pejoratively or melioratively)
19	<i>kwyl+=bekk·ie</i> drool+=beak·DIM 'baby (with drooling mouth)'
20	<i>donker+=kop</i> dark+=head 'person with dark hair; head with dark hair'
21	<i>pad+=buffel</i> road+=buffalo 'road hog'
22	<i>onder+=kruip·er</i> under+=crawl·NMLZ 'interloper'
23	<i>Daantjie-lang+verwag</i> Daantjie.N.PR-long+expected 'guest who arrives unannounced'
24	<i>hans+wors</i> < German <i>Hans Wurst</i> (literary character) hand-fed+sausage 'clown'

Additionally, **test paraphrases** for different categories were formulated. For example, a true EEC (14) can only be used in the sentence 'S/he is a horrible/annoying/no-good [x]<sub>N</sub>', while a promiscuous EEC with multiple functions (16) could in addition be paraphrased with 'S/he stomps around [x]·ly<sub>ADV</sub>', and/or 'The horrible/annoying/no-good [x]<sub>ADJ</sub> child'.

Based on these initial annotations, the dataset was divided into several subsets, shown in Table 1. Subsets U–Z are not directly relevant to this paper, since these cases are either metaphoric (U–W; n = 365), or not compounds (X–Y; n = 158).

Table 1: Subsets of data after first round of annotation

	Description	Example	Translation	n
–	Population: Potential EECs	<i>dikgat</i>	fatty	395
<b>U</b>	Exclusively metaphoric epithet	<i>motorbuffel</i>	road hog	286
<b>V</b>	Secondarily metaphoric epithet	<i>tierwyfie</i>	tigress, shrew	49
<b>W</b>	Exclusively metaphoric hypocoronym	<i>hartlam</i>	darling	30
<b>X</b>	Epithet with other constructionalisation	<i>pennelekker</i>	penpusher	155
<b>Y</b>	Hypocoronym with other constructionalisation	<i>hartkruipertjie</i>	heart-throb	3
<b>Z</b>	Looks and behaves like an EEC, but is metaphoric	<i>poephol</i>	asshole	4
			<b>TOTAL</b>	<b>922</b>

Four words that look and behave like EECs, but that should not be considered EECs, have also been separated (subset Z) from the rest of the data. These four words (25 to 28) have RH constituents meaning [BODY PART], but none of these constituents occur as RH constituents in other EECs. Moreover, all four words fit a metaphoric interpretation perfectly, unlike EECs.



25 *poep+hol*  
fart+hole  
'ass-hole' > 'S/he is **like** an ass-hole'

26 *moer·s+kont*  
mother-LK+cunt  
'cunt' > 'S/he is **like** his/her mother's cunt'

27 *hond·s+vot*  
dog-LK+vagina  
'scoundrel' > 'S/he is **like** a bitch's genitals'

28 *blou+bloed*  
blue+blood  
'royalty' > 'S/he behaves **like** s/he has blue blood'

The remaining lemmas (n = 395) form our **population set** of potential EECs, since they can all be used evaluatively to refer to people, while their RH constituents have the literal senses [BODY PART] (76%), or [GARMENT] (24%). The population set was then enriched by:

- providing **plural forms** for all lemmas;
- giving broad **semantic characterisations** of the composite constructions, using nominal synonyms (e.g., 'habitual criminal' for *blou+baadjie* blue+jacket), prepositional phrases (e.g., 'with split palate' for *drie+lip* three+lip 'person with split palate'), adjectives (e.g., 'slow' for *draai+gat* slow+ass 'slow person'), or relative clauses (e.g., 'who day-dreams' for *droom+kop* dream+head 'person who day-dreams');
- doing **morphological analyses** of all EECs, including:
  - the lemma, part-of-speech, semantic, and stratum information for each constituent;
  - syllable length and stress, plus whether the word is sanctioned by the schwa schema in [2][b]; and
  - whether an interfix is present or not, and if so, what the allomorph of the lefthand (LH) constituent is (e.g., in 29 the *·e·* is an interfix, and when it combines with *nuk* –the LH constituent – it results in the allomorph *nukke*);<sup>8</sup>
- identifying **temporal and usage labels** for the RH constituents (e.g., if a constituent is obsolescent (temporal), or derogatory (usage)); and
- adding combined **frequency counts** of lemmas and plural forms from the following corpora of edited text: Radio Sonder Grense and CText (2018); Lapa uitgewers and CText (2018); Maroela Media and CText (2018); Protea Boekhuis and CText (2018); Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns (2011); and unedited text: Centre for Text Technology (CTeX) (2023).

---

<sup>8</sup> See Trollip and Van Huyssteen (2018) for motivation of this analysis.

29 *nukk·e÷bol*  
whim·LK+head  
'curmudgeon'

Based on these annotations, the population set ( $n = 395$ ) was divided into more subsets (subsets A–G in Table 2). Because the words in D don't occur in contemporary corpora, and because the words in E–G are not primarily used as epithets, descriptions in the remaining sections of this paper are mainly based on subset A–C, our **core set** ( $n = 170$ ), unless mentioned otherwise. Shortcomings and possible extensions or refinements of the dataset are discussed in § 9. The full dataset (including the other extracted data, i.e., non-EECs) is freely available for research purposes; see footnote 2.

Table 2: Subsets of population after second round of annotation

	Description	Example	Translation	n
<b>A</b>	EEC	<i>dikgat</i>	fatty	134
<b>B</b>	EEC, but with multiple functions	<i>hardegat</i>	hard-ass (N; ADJ)	18
<b>C</b>	EEC, but with multiple secondary meanings	<i>skeeloog</i>	boss-eyed (N; ADJ); strabismus	18
<b>D</b>	EEC, but $n = 0$	<i>dreungat</i>	swagger	140
<b>E</b>	EEC, but primarily neutral	<i>asgat</i>	rubbish dump; asshole	50
<b>F</b>	EEC, but almost exclusively neutral	<i>swartkop</i>	black-haired person	15
<b>G</b>	EEC, but primarily hypocoristic	<i>lekkerbek</i>	gourmet (i.a.)	20
<b>TOTAL</b>				<b>395</b>

#### 4 Descriptive formalism

For descriptive purposes, we use conventions from various flavours of construction morphology, as detailed in Van Huyssteen (accepted)<sup>9</sup>. Suffice to demonstrate how a prototypical example of a schema should be read and interpreted, before providing the full schema for Afrikaans EECs; this schema will be discussed and motivated in the remainder of the paper.

The word *pap÷broek* (soft÷pants 'softy, coward') has the highest frequency of all EECs in subset A and can hence be considered a prototypical example. Based on this prototypical example, an EEC can be characterised and formalised as an exocentric, attributive ground compound – see the schema in [1]. This formalisation simply reads: A adjectival stem **x** (e.g., *pap*) combines with a nominal suffixoid **÷y** (e.g., *÷broek*) in an exocentric, attributive ground

<sup>9</sup> Several introductory texts on construction morphology are widely available, e.g., Audring, 2022; Booij, 2010, 2012, 2017; Booij & Audring, 2017; Jackendoff & Audring, 2020; Masini & Audring, 2019.

compound (*pap÷broek*) that is also a noun, to characterise a person in terms of the semantics of the adjectival stem (i.e., ‘person who is soft, weak, fearful, cowardly’).

[1] **Schema: type *pap÷broek*** (soft÷pants ‘softy, coward’)

$$[ [x]_{\text{ADJ}}^{\text{stem}} [\div y]_{\text{N}}^{\text{fxoid}} ]_{\text{N}}^{\text{cmpd.attr.grnd.exo}} \Leftrightarrow [\text{PERSON}_{\text{N}}^{\text{fxoid}} \text{ WHO IS SEM}_{\text{ADJ}}^{\text{stem}} ]_{\text{N}}^{\text{cmpd}}$$

In a more elaborate format – following and adapting some of the conventions of Jackendoff and Audring (2020) (see Van Huyssteen accepted) – the same information can be presented as in [2] below. The subscript indices are used as links between the different descriptive levels, e.g., constituent <sub>1</sub> is an adjectival stem that is construed as attribute of the [PERSON] that the compound refers to. This more elaborate representation, which offers the opportunity to add many more levels of information, will be used in the remainder of the article.

The formalised schema in [2] is a comprehensive, all-inclusive description of all instances of EECs in the core set (A–C). Since the semantics of EECs fall outside the scope of this paper, we only incorporate high-level representations of the conceptualisation of EECs ([2][k] to [2][m]). All the other levels are discussed and motivated in the following sections.

[2] **Schema: EEC**

Phonological realisation:

- [a] Prototype: /['[σ...]<sub>1</sub> [σ...]<sub>2/3</sub>/
- [b] Elaboration 1: /['[σ...] [σ...ə...]<sub>1</sub> [σ...]<sub>2/3</sub>/
- [c] Elaboration 2: /['[σ...]<sub>1</sub> [σ...ə]<sub>4</sub> ]<sub>5</sub> [σ...]<sub>2/3</sub>/
- [d] Elaboration 3: /'[σ...]<sub>1</sub> [[σ...][σ...]<sub>2/3</sub>/
- [e] Phonological extensions: *baklei÷gat*, *kerrie÷bek*, *peperkorrel÷kop*, *petrol÷kop*, *suur÷gesig*, *twee÷gesig*

MorphSynForm:

- [f] Prototype: [<sup>cmpd.attr.grnd.exo</sup> [stem]<sub>1</sub> [suffixoid]<sub>2</sub> ]<sub>3</sub>
- [g] Elaboration: [<sup>cmpd.attr.grnd.exo</sup> [<sup>ADJ</sup> [<sup>ADJ</sup> stem]<sub>1</sub> [<sup>CN.ATTR</sup> ·e]<sub>4</sub> ]<sub>5</sub> [suffixoid]<sub>2</sub> ]<sub>3</sub>
- [h] Morphological extensions: *bull·e÷bakk·ie·s*, *nukk·e÷bol*

MorphSynFunction:

- [i] Prototype: [<sub>N</sub> [ADJ|N|V]<sub>1</sub> [N]<sub>2</sub> ]<sub>3</sub>
- [j] Functional extensions: *een÷oog*, *klaas÷neus*, *twee÷gesig*

Conceptualisation:

- [k] Superordinate: [PERSON<sub>2</sub> IN RELATION R TO SEM<sub>1</sub>]<sub>3</sub>
- [l] Subordinate (prototype): [ [BODY PART FOR PERSON]<sub>2</sub> IN RELATION R TO SEM<sub>1</sub> ]<sub>3</sub>
- [m] Subordinate: [ [GARMENT FOR PERSON]<sub>2</sub> IN RELATION R TO SEM<sub>1</sub> ]<sub>3</sub>

Stratum:

[n] [ <<S<sub>gem</sub>>><sub>1</sub> <<S<sub>gem</sub>>><sub>2</sub> ]<sub>3</sub>

[o] Stratal extensions: *petrol*÷*kop*, *kerrie*÷*bek*

## 5 Phonological realisation

With regard to Dutch EECs, De Haas and Trommelen (1993, 412) note that they are characterised by conciseness, which makes them particularly suitable as epithets. Jansen (1985, 93) confirms: “The wordform [of the Dutch constituents] must apparently be preferably short, simple, and native.” [my translation – GBVH]. He reports that from the 158 Dutch words discussed in his book, 81% of the constituents have only one syllable, 18% two syllables, and 1% (i.e., two words) three syllables (Jansen 1985, 93). The Afrikaans data, summarised in Table 3 confirm these phonological patterns for Afrikaans:

- 75% of the 170 EECs consist of one monosyllabic LH constituent, and 94% of one monosyllabic suffixoid. In all these cases, stress falls on the first syllable (like in ordinary compounds), e.g., /'kas.kɔp/ (cheese÷head ‘Dutch person’), or /'lœi.xat/ (lazy÷ass ‘lazy person’); see the schema in [2][a].
- If the LH constituent is multisyllabic (n = 43), the second, unstressed syllable contains a schwa in 91% (n = 39) of the cases, e.g., /'kri.vəl.kœus/ (twiddle÷sock ‘fidget’), or /'flɛn.tər.gat/ (rag÷ass ‘untidy person’); see [2][b] and [2][c]. The four exceptions listed in [2][e] are:
  - /bɑ'klœi.xat/ (fight÷ass ‘provoker, quarreller’);
  - /'kæ.ri.bæk/ (curry÷beak ‘person who swears a lot’)
  - /'pɛ.pər.kɔ.rəl.kɔp/ (peppercorn÷head ‘person who is black’); and
  - /'pɛ.trəl.kɔp/ (petrol÷head ‘id.’).
- If the suffixoid is multisyllabic (n = 11), stress falls on the first syllable in nine of these cases (which means that secondary stress of the composite construction also falls on this syllable); see [2][d]. These suffixoids are:
  - /'a.səm/ (‘breath’);
  - /'bai.ci/ (‘jacket’);
  - /'bɑ.cis/ (‘face’);
  - /'ba.li/ (‘stomach’);
  - /'dæ.rəm.pi/ (‘intestine·DIM’); and
  - /'fɛ.ŋərs/ (‘finger·PL’).

The only extension of this schema is EECs with the suffixoid /xə'səx/ (‘face’) (n = 2), as in *suur*÷*gesig* (sour÷face ‘id.’), and *twee*÷*gesig* (two÷face, ‘two-faced (person)’); see [2][e]).

Table 3: Number of syllables in constituents (n = 170)

	LH constituent	RH constituent
1 syllable	127 (74,7%)	159 (93,5%)
2 syllables	42 (24,7%)	10 (5,9%)
3 syllables	0 (0%)	1 (0,6%)
4 syllables	1 (0,6%)	0 (0%)

## 6 Morphosyntactic form

With respect to morphosyntactic form, constructionalisation occurs through the process of exocentric compounding. In all cases, their morphosyntactic head (RH component) is not their semantic head, and they are therefore (per definition) semantically opaque. For example, in *wys÷neus* (wise÷nose 'know-it-all') *÷neus* is the morphosyntactic head that determines the part-of-speech (N) and inflectional morphology of the composite construction (the plural of *wys÷neus* is the same as *neus*: *wys÷neus·e* and *neus·e*). However, since *÷neus* does not mean [BODY PART USED FOR SMELLING] (as per its conventional interpretation) in the composite construction, it should rather be analysed as a suffixoid with the specialised meaning [PERSON]. This is captured in the schema in [2][f], repeated here for convenience of reference.

### [2] Schema: EEC

MorphSynForm:

[f] Prototype: [<sub>cmpd.attr.grnd.exo</sub> [stem]<sub>1</sub> [suffixoid]<sub>2</sub>]<sub>3</sub>

[g] Elaboration: [<sub>cmpd.attr.grnd.exo</sub> [<sub>ADJ</sub> [<sub>ADJ</sub> stem]<sub>1</sub> [<sub>CN.ATTR</sub> ·e]<sub>4</sub>]<sub>5</sub> [suffixoid]<sub>2</sub>]<sub>3</sub>

[h] Morphological extensions: *bull·e÷bakk·ie·s*, *nukk·e÷bol*

On a morphological level, Afrikaans EECs rarely contain interfixes, unlike the case in Dutch (De Haas and Trommelen 1993, 413; Trollip 2022, 56–57). In instances like *hard·e÷kop* (hard·ATTR÷head 'stubborn person'), and *styw·e÷nek* (stiff·ATTR÷neck 'grumpy person') (n = 7), the *·e* is not an interfix, but rather an attributive suffix that is part of the adjective. This information is captured in [2][g], where index number <sub>4</sub> marks the category-neutral attributive suffix, and <sub>5</sub> the affixed adjective as a whole; the phonological realisation is captured in [2][c].

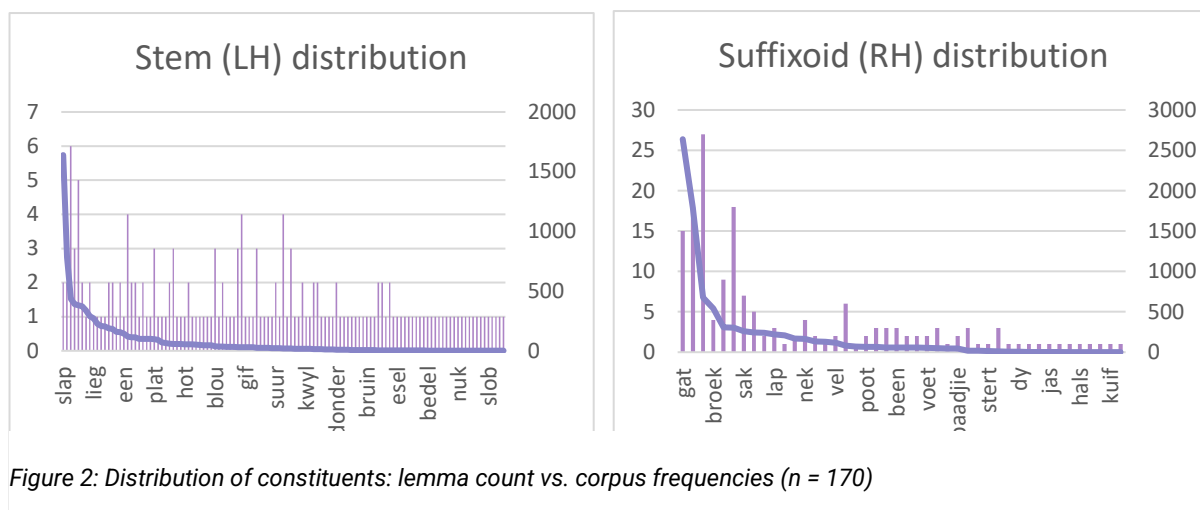
The *·e·* interfix occurs only in two archaic (i.e., inherited) EECs: *nukk·e÷bol* (whim·LK÷head 'curmudgeon' < Dutch id.); and *bull·e÷bakk·ie·s* (< Dutch *bull·e÷bak* < *bulder·en+bakkes* roar·LK+face 'bully'). They are both listed as extensions in [2][h]. In the larger population set, four additional archaic/inherited EECs with an *·e·* interfix is found (see 30).

30 <i>ap·e÷bakkies</i> monkey·LK÷face 'unattractive person'	<i>brek·e÷been</i> break·LK÷leg 'clumsy person'	<i>siel·e÷poot</i> soul·LK÷paw clumsy person'	<i>skatt·e÷bol</i> darling·LK÷head 'sweetheart'
---	---	---	---

The *·s·* interfix does not occur in the core set, and only four times in the population set, viz., *kalf·s÷kop* (31), *korrel·tjie·s÷kop* (32), *rok·s÷band·jie* (33), and *skref·ie·s÷oog* (34). Except for *korrel·tjie·s÷kop*, the others have first and foremost literal meanings, i.e., ‘calf’s head’, ‘spaghetti straps’, and ‘squinted, slitty eyes’ respectively. Since they fit the paradigm (i.e., pattern) of comparable endocentric compounds (see Trollip & Van Huyssteen 2018), they could therefore rather be seen as ordinary endocentric compounds that are being used metaphorically. Since they are not part of the core set, they are excluded from the schema in [2]. *Korrel·tjie·s÷kop* is excluded from the core set, because no corpus attestations could be found; it is therefore in subset D (see Table 2).

<b>31</b>	<i>kalf·s÷kop</i>
	calf·LK+head
	‘calf’s head; person with limited cognitive abilities’
<b>32</b>	<i>korrel·tjie·s÷kop</i>
	bead·DIM·LK+head
	‘person who is black’
<b>33</b>	<i>rok·s÷band·jie</i>
	skirt·LK+strap·DIM
	‘spaghetti straps (on a dress); favoured child, teacher’s pet’
<b>34</b>	<i>skref·ie·s÷oog</i>
	slit·DIM·LK+eye
	‘squinted, slitty eyes; person from the Orient’

Looking at the distribution of constituents, we unsurprisingly see a Zipfian-like distribution<sup>10</sup> for both LH and RH constituents (see the violet lines in Figure 2). The “less than perfect” Zipfian distribution can be attributed to the small number of types in the core set (Lavi-Rotbain and Arnon 2023), rather than to some kind of uniqueness of distributions in the construction. For example, in the case of the RH constituents, *÷kop* appears in the most lemmas ( $n = 27$ ), but only contributes the third most wordforms in our corpus data ( $N = 677$ ). In contrast, *÷gat* (‘arse’) only occurs in 15 EECs, while appearing in almost a third of all the wordforms ( $N = 2,638$ ). Since there is a wider variety of LH constituents ( $n = 117$ ) than RH constituents ( $n =$



<sup>10</sup> For a brief introduction to Zipf’s Law and Zipfian distribution, see <https://nlp.stanford.edu/IR-book/html/htmledition/zipfs-law-modeling-the-distribution-of-terms-1.html>.

44), we also see more of these seemingly exceptional cases that could be attributed to data sparsity.

## 7 Morphosyntactic function

Given our demarcation of EECs here, all RH constituents are nouns denoting outside EECs either a type of [BODY PART], or [GARMENT]. Since the RH constituent is the morphological head of the construction (see § 6), all composite constructions are nouns as well. In three cases, the RH constituent is a plurale tantum: *botter÷vinger-s* (butter÷finger·PL ‘id., person who frequently drops items’), *donder÷dy-e* (thunder÷thigh·PL ‘person with oversized legs and buttocks’), and *slap÷hand·jie-s* (limp÷hand·DIM·PL ‘person who is lazy’). Another two RH constituents are diminutiva tantum: *melk÷derm·pie* (milk÷intestine·DIM ‘person who is easily scared; nancy boy’), and *riet÷been·tjie* (reed÷leg·DIM ‘person with long, thin legs; matchstick’).

Similar to Dutch (De Haas and Trommelen 1993, 412), the LH constituent can be an adjective (ADJ), noun (N), verb (V), and rarely a proper noun (N.PR) or numeral (NUM). Distributions are reported in Table 4, where a more or less equal distribution between adjectival (40%) and nominal (34%) LH components can be seen, while the former makes up almost two-thirds (64%) of the total number of corpus cases. In the population set, however, there are more nominal (40%) LH components than adjectival (35%) ones, with the compounds with adjectival LH components still making up almost two-thirds (64%) of the corpus cases (table included in the dataset). This confirms that the data sparsity of the core set should be kept in mind when drawing any conclusions about EECs.

Table 4: Distribution of LH constituents

	n = 170	N = 8,935
<b>ADJ</b>	67 (39,4%)	5713 (63,9%)
<b>N</b>	58 (34,1%)	1713 (19,2%)
<b>V</b>	42 (24,7%)	1192 (13,3%)
<b>NUM</b>	2 (1,2%)	300 (3,4%)
<b>N.PR</b>	1 (0,6%)	17 (0,2%)

These emerging patterns are reflected in the schema (pasted below for convenience of reference) in [2][i]. Note that low-frequency cases with NUM and N.PR (and N.DIM in the population set) are listed in [2][j] as extensions of the schema.<sup>11</sup>

<sup>11</sup> This functional dimension of the schema seems to hold also for other EECs without constituents referring to [BODY PART] or [GARMENT]. For example, Van Huyssteen and Koekemoer (2023) mention three additional schemas: PREP÷V (*by÷slaap* with÷sleep ‘sexual partner; the sex deed’), V÷PREP (*flap÷uit* flap÷out ‘somebody who easily discloses confidential information’), and V÷NUM (*bemoei÷al* meddle÷all ‘person who interferes in other people’s business’). None of these, however, are productive schemas.

[2] **Schema: EEC**

MorphSynFunction:

[i] Prototype: [<sub>N</sub> [ADJ|N|V]<sub>1</sub> [N]<sub>2</sub>]<sub>3</sub>

[j] Functional extensions: *een÷oog, klaas÷neus, twee÷gesig*

## 8 Stratal restrictions

While these types of compounds have already been present in the prehistory of Dutch (Van der Sijs 2019, 36), it is especially in Early New Dutch when compounding with a noun denoting [BODY PART] (35, 36) or [GARMENT] (37) became a productive morphological process.

<p>35 <i>brek·e+been</i> break·LK+leg 'clumsy person'</p>	<p><i>huil·e+balk</i> sad+ear 'person/child who is gnawingly sad'</p>
<p>36 <i>dom+kop</i> stupid+head 'person with limited cognitive abilities'</p>	<p><i>lach·e+bek</i> laugh·LK+beak 'person who laughs easily and a lot'</p>
<p>37 <i>klets+kous</i> chat+sock 'person who can talk, gossip a lot'</p>	<p><i>dik+zak</i> thick+bag 'person who is corpulent'</p>

Given this diachronic development, and given that we know that Dutch EECs have a very strong preference for Germanic constituents only (Jansen 1985, 93), it is not surprising that the two constituents of Afrikaans EECs are almost exclusively from the primary native stratum (Germanic). 100% of the righthand (RH) constituents in the dataset are native, while only two lefthand (LH) constituents are not from the Germanic stratum, viz. *petrol* (38), and *kerrie* (39).<sup>12</sup>

<p>38 <i>petrol÷kop</i> petrol+head 'mechanic'</p>
<p>39 <i>kerrie÷bek</i> curry+beak 'person who swears a lot'</p>

These facts of the core dataset are captured in the schema as [2][n] and [2][o], pasted below for ease of reference. Stratal information (S) is presented between double arrow parentheses (<< >>), with the ISO language abbreviation ("gem" for "Germanic") in subscript.

[3] **Schema: EEC**

Stratum:

[k] [ <<S<sub>gem</sub>>><sub>1</sub> <<S<sub>gem</sub>>><sub>2</sub> ]<sub>3</sub>

<sup>12</sup> If we look at the population dataset, 97% of the lefthand (LH) constituents are from the Germanic stratum. Of the remaining 3%, two are non-native proper names (*Janus, Judas*), one a verb from the Classic stratum (*lakseer* 'to defecate'), and the rest highly conventionalised loan words: *ghries* 'grease'; *kaboe* 'unbroken, boiled grain'; *kalbas* 'calabash'; *lemoen* 'orange (fruit)'; *makaster* '(obsolete) cap for baby'; plus the abovementioned *petrol* 'id.', and *kerrie* 'curry'.



[1] Stratal extensions: *petrol=kop*, *kerrie=bek*

## 9 Conclusion

To conclude this paper, we present an agenda for future descriptive and comparative research on Afrikaans EECs. We also include themes related to previous research.

- **Dataset**

- Only datasets A–G have been annotated fully; these annotations could be extended to the full set of potential epithets (A–Z), especially if one wants to gain insight in other types of semantic motivations for epithets.
- For example, one should take note of the categories of Barcelona (2011) regarding metonymy-metaphor interaction in the conceptual motivation of these types of compounds. In addition to the extensional range of these compounds (i.e., the category of entities designated by each compound, such as people, physical objects, plants, animals, etc.), the conceptualisation of the characteristic property mapped by the preceding metonymy could be categorised as one of the following types (Barcelona 2011, 157):
  - nonmetonymically and nonmetaphorically (i.e. “literally”), e.g., *humpback* (‘person with a hump’);
  - metonymically and non-metaphorically, like *acidhead* (‘one who habitually takes the drug LSD’); and
  - metaphorico-metonymic interaction, e.g., *fathead* (‘stupid person’) (Barcelona 2011, 162), and *featherweight* (‘very light combatant’) (Barcelona 2011, 165).
- One of the most important shortcomings of the current dataset, is that it has not yet been cross-validated by external annotators. Nonetheless, most of the current annotations have at least been “indirectly checked” three times: first by the editors of the dictionaries on which Van Niekerk (2002, 2006) based her data; secondly interpreted by Van Niekerk herself for these studies; and thirdly by me for the current study. While this is not necessarily sufficient for the strictest scientific scrutiny, it at least provides a baseline dataset for future work.

- **Morphological productivity**

- Regarding exocentric compounding in Afrikaans in general, Van Niekerk (2002, 26–27) noted: “Endocentric compounding as a word formation process is much more productive than exocentric compounding ...” [my translation – GBVH]. However, she does not provide any empirical proof, and while this could

be investigated in future research, it would perhaps be more interesting to focus on the productivity of EECs and EEC-like compounds.

- Regarding suffixoids in EECs, she remarked: “In certain cases a certain stem may indeed be more productive in forming a series of exocentric compounds, on the same pattern as endocentric compounds...” [my translation –GBVH]. Since she also did not provide any empirical evidence in this regard, Trollip (2022) investigated the matter in more detail. He concluded:

The person-forming suffixoids most often found in Afrikaans appear to be *÷gat* [*÷arse*], *÷kop* [*÷head*], and *÷kous* [sock]. The other suffixoids, such as *÷balie* [*÷tummy*] and *÷broek* [*÷pants*], occur only sporadically... Based on usage data, these types of suffixoids, and specifically *÷head* and *÷nose*, offer users a productive means to create nicknames. (Trollip 2022, 121) [my translation – GBVH]

- Five of the six suffixoids that Trollip mentioned appear also in our top six most frequently occurring (not necessarily productive) suffixoids, with his *÷balie* only occurring in position 19 in our data. While *÷bek* is in position two in our data, Trollip does not mention it in his summary.
- Since it is an interesting construction to investigate in terms of productivity – also cross-linguistically – it might be worthwhile to investigate the matter further, using other corpora than Trollip. One could also take the productivity of certain semantic groups into account in future.

- **Semantics**

- While Barcelona (2011, 2008) provided extensive semantic analyses of bahuvrihi compounds in English and Spanish within a Langackerian framework (Langacker 1999, ch. 6), specific matters related to personal names (and more specifically epithets) have not yet been addressed comprehensively.
- In a future cognitive semantic characterisation of EECs, it could be interesting to incorporate the insights and methods of the Natural Semantic Metalanguage framework, especially since the usage of EECs might vary between different closely-related cultures (Goddard 2015). While a comparison between, for instance, (West-)Germanic languages regarding form and function might not reveal too many differences between these languages, the meaning and usage of EECs in these languages might vary.
- Such comparative research could especially be insightful if one uses techniques from data-based semantic analysis and clustering. For example, following the methodology of Van Goethem and Norde (2020), one could

combine a corpus study with statistical methods in distributional semantics and collexeme analysis.

- **Socio-pragmatics**

- In a recent study, Van Huyssteen and Koekemoer (2023) investigated the potential offensiveness of EECs, using online questionnaires. They found, on the one hand, that there was not enough convincing evidence to be able to say with the necessary certainty that non-EECs are less offensive than EECs. On the other hand, they were able to order the four investigated suffixoids on a spectrum of potential offensiveness, namely from least offensive to most offensive: *÷neus* (÷nose)  $\approx$  *÷broek* (÷pants)  $<$  *÷sak* (÷bag)  $<$  *÷gat* (÷arse).
- Since none of their conclusions pointed to clear-cut categories, we might assume for the time being that these kinds of epithets are not more or less offensive than other types of epithets.
- We would rather suggest that a better avenue might be to do an investigation into the role of metonymy in impolite, hateful, aggressive, and/or dehumanising discourse – similar to what Knoblock and Sazonova (2022) have done regarding “it” as a dehumanising metaphor. This is, of course, a theme that should encompass more than EECs only, but EECs could potentially serve as practical starting point.

## 10 Acknowledgements

I would like to thank Martin Puttkammer for assisting with automatic data extraction from the PDFs, as well as for obtaining corpus frequency counts for large lists of words. Thank you also to Adri Breed and Suléne Pilon who always serve as critical soundboards, while Anneke Butler and Monique Rabé also provided feedback on some earlier classifications. None of the results and/or opinions in this paper can be ascribed to any of them.

Ethical clearance for the research project was obtained through the Language Matters Ethics Committee of the NWU (ethics number: NWU-00632-19-A7).

## 11 References

- Amiot, Dany, and Edwige Dugas. 2020. *Combining Forms and Affixoids in Morphology*. Oxford University Press.
- Barcelona, Antonio. 2008. "The interaction of metonymy and metaphor in the meaning and form of 'bahuvrihi' compounds." *Annual Review of Cognitive Linguistics* 6: 208-281. <https://doi.org/10.1075/arcl.6.10bar>.
- . 2011. "The conceptual motivation of bahuvrihi compounds in English and Spanish." In *Cognitive Linguistics : Convergence and Expansion*, edited by Mario Brdar, Stefan Th. Gries and Milena Žic Fuchs, 151–178. Amsterdam: John Benjamins.
- Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge: Cambridge University Press.
- . 2018. *Morphological Entities: Overview and General Issues*. Oxford: Oxford University Press.
- Bauer, Laurie, Rochelle Lieber, and Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Booij, Geert. 2020. "Exocentric compounds." *Taalportaal*. Accessed 18 June.
- Booij, Geert, and Matthias Hüning. 2014. "Affixoids and constructional idioms." In *Extending the scope of Construction Grammar*, edited by Ronny Boogaart, Timothy Coleman and Gijbert Rutten, 77–106. Berlin: Mouton de Gruyter.
- Brewer, Marilyn B. 2016. "Intergroup Discrimination: Ingroup Love or Outgroup Hate?" In *The Cambridge Handbook of the Psychology of Prejudice*, edited by Chris G. Sibley and Fiona Kate Barlow, 90-110. Cambridge: Cambridge University Press.
- Centre for Text Technology (CTeXt). 2023. *NWU Kommentaarkorpus 2.2 [NWU Commentary Corpus 2.2]*. edited by North-West University. Potchefstroom: Queried on Sketch Engine: [app.sketchengine.eu](http://app.sketchengine.eu).
- De Haas, Wim, and Mieke Trommelen. 1993. *Morfologisch handboek van het Nederlands. Een overzicht van de woordvorming*. 's-Gravenhage: SDU Uitgeverij.
- Goddard, Cliff. 2015. "'Swear words" and "curse words" in Australian (and American) English. At the crossroads of pragmatics, semantics and sociolinguistics." *Intercultural Pragmatics* 12 (2). <https://doi.org/10.1515/ip-2015-0010>.
- Jackendoff, Ray, and Jenny Audring. 2020. *The Texture of the Lexicon*. Kindle ed. Oxford: Oxford University Press.
- Jansen, Frank. 1985. *Van Apekop tot Zielepoot : hoe mensen elkaar typeren*. 's-Gravenhage: Nijgh en Van Ditmar.
- Knoblock, Natalia, and Yaroslava Sazonova. 2022. "Neutering unpopular politicians: The neuter gender and 'it' as a dehumanizing grammatical metaphor." In *The grammar of*

- hate: morphosyntactic features of hateful, aggressive, and dehumanizing discourse*, edited by Natalia Knoblock, 140–160. Cambridge: Cambridge University Press.
- Langacker, Ronald W. 1999. *Grammar and Conceptualization*. Vol. 14. Berlin: Mouton de Gruyter.
- Lapa uitgewers, and CText. 2018. NWU/Lapa-korpus 1.2. Potchefstroom: CText, Noordwes-Universiteit.
- Lavi-Rotbain, Ori, and Inbal Arnon. 2023. "Zipfian Distributions in Child-Directed Speech." *Open Mind* 7: 1-30. [https://doi.org/10.1162/opmi\\_a\\_00070](https://doi.org/10.1162/opmi_a_00070).
- Maroela Media, and CText. 2018. NWU-Maroela Media-korpus 1.3. Potchefstroom: CText, Noordwes-Universiteit.
- Protea Boekhuis, and CText. 2018. PUK/Protea Boekhuis-korpus 2.2. Potchefstroom: CText, Noordwes-Universiteit.
- Radio Sonder Grense, and CText. 2018. RSG Afrikaanse korpus 2.2. Potchefstroom: CText, Noordwes-Universiteit.
- Scalise, Sergio, and Antonietta Bisetto. 2009. "The classification of compounds." In *The Oxford Handbook of Compounding*, edited by Rochelle Lieber and Pavol Štekauer, 34–53. Oxford: Oxford University Press.
- Scalise, Sergio, Antonio Fábregas, and Francesca Forza. 2009. "Exocentricity in compounding." *Gengo Kenkyu/Journal of the Linguistic Society of Japan* 135: 49–83.
- Schlücker, Barbara. 2023. Compounding and Linking Elements in Germanic. In *The Oxford Encyclopedia of Germanic Linguistics*, edited by Sebastian Kürschner and Antje Dammel. Oxford: Oxford University Press.
- Štekauer, Pavol, Salvador Valera, and Livia Kórtvélyessy. 2012. *Word-Formation in the World's Languages: A Typological Survey*. Cambridge: Cambridge University Press.
- Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2011. Taalkommissiekorpus 1.1. edited by North-West University. Potchefstroom.
- Trollip, Eddie Benito. 2022. "Morfologiese evalueringskonstruksies in Afrikaans [Morphological evaluation constructions in Afrikaans]." PhD, Department of Afrikaans and Dutch, North-West University, Local.
- Trollip, Eddie Benito, and Gerhard B. Van Huyssteen. 2018. "The linking morpheme in Afrikaans: a Cognitive Grammar description." *SKASE Journal of Theoretical Linguistics [online]* 15 (3): 37–70. [http://www.skase.sk/Volumes/JTL38/pdf\\_doc/03.pdf](http://www.skase.sk/Volumes/JTL38/pdf_doc/03.pdf).
- Van der Sijs, Noline. 2019. *15 eeuwen Nederlandse taal*. Gorredijk: Sterck & De Vreese.
- Van Goethem, Kristel, and Muriel Norde. 2020. "Extravagant "fake" morphemes in Dutch. Morphological productivity, semantic profiles and categorical flexibility." *Corpus Linguistics and Linguistic Theory* 16 (3): 425–458. <https://doi.org/10.1515/cllt-2020-0024>.

Van Huyssteen, Gerhard B. accepted. "The morphology of Afrikaans." In *Contemporary Afrikaans Linguistics*, edited by Wannie A. M. Carstens and Nerina Bosman. Pretoria: Van Schaik.

Van Huyssteen, Gerhard B., and Simoné Koekemoer. 2023. "'n Empiriese vergelyking van die potensiële aanstootlikheid van enkele skelnaampare in Afrikaans [An empirical comparison of the potential offensiveness of some epithet pairs in Afrikaans]." *Tydskrif vir Geesteswetenskappe* 63 (3): 560–584. <https://doi.org/10.17159/2224-7912/2023/v63n3a10>.

Van Niekerk, Lariza. 2002. "'n Korpusanalise van Afrikaanse eksosentriese komposita." M.A.-verhandeling, Universiteit van Suid-Afrika.

—. 2006. "Funksionele aspekte van Afrikaanse eksosentriese komposita." DLitt et Phil, Afrikaans, Universiteit van Suid-Afrika.