

## **Hunchbacks and Palefaces**

Baumgartner, Kirsten

In: Student Journal of the Department of Anglophone Studies / Volume 1 (2018)

This text is provided by DuEPublico, the central repository of the University Duisburg-Essen.

This version of the e-publication may differ from a potential published print or online version.

DOI: <https://doi.org/10.17185/duepublico/47605>

URN: <urn:nbn:de:hbz:464-20181123-142644-4>

Link: <https://duepublico.uni-duisburg-essen.de:443/servlets/DocumentServlet?id=47605>

License:

As long as not stated otherwise within the content, all rights are reserved by the authors / publishers of the work. Usage only with permission, except applicable rules of german copyright law.

Source: This essay was initially submitted as a term paper in the BA seminar “Morphology” and supervised by Julia Salzinger. Published in Student Journal of the Department of Anglophone Studies (Vol. 1, 2018)

# Hunchbacks and Palefaces

## Exocentric Compound Nouns and their Productivity

---

KIRSTEN BAUMGARTNER<sup>1</sup>

What do words like *skinhead*, *blackboard*, *pickpocket*, *underdog*, and *lazybones* have in common? First of all, these are words that we use in everyday English, but that would be too simplistic in a linguistic context. What they really have in common is that they are compounds. These words consist of two constituent parts: *skinhead* is a compound consisting of *skin* and *head*, *pickpocket* consists of *pick* and *pocket*, *underdog* of *under* and *dog*, and *lazybones* consists of *lazy* and *bones*. We might notice that the constituent parts have, in a literal sense, nothing to do with the meaning of the whole compound. A *skinhead* is not a kind of *head* or *skin*, it describes a *person*. A *blackboard* differs from the meaning in the way that it is not a kind of *board* that is black, which would be embodied by *black board*. And *underdog*? This has nothing to do with a kind of *dog* as well as *lazybones* has nothing to do with kinds of *bones*. In linguistics, compounds whose meaning is not derived from the constituent parts are known as *exocentric compound nouns*.

Exocentric compound nouns are said to show low productivity according to Andrew Carstairs-McCarthy (2002) and Laurie Bauer (2004).<sup>2</sup> Why is that and can this be said for all types of exocentric compound nouns? I will investigate this question first using a theoretical base that defines characteristics of exocentric compound nouns, including pre-understandings and their influence. Secondly, I will compare selected types of exocentric compound nouns regarding the question of productivity. Research so far just pointed out difference in productivity by comparing endocentric and copulative to exocentric compound nouns.<sup>3</sup> Therefore, it is important to look at internal differences in productivity, limited to exocentric compound nouns, to prevent generalisations like saying exocentric compound

---

1 This essay was initially submitted as a term paper in the BA seminar “Morphology” and supervised by Julia Salzinger.

2 *Low productivity* means here that there are not many exocentric compound nouns newly invented (see Carstairs-McCarthy 110 and Plag 52-54f.).

3 See for example Bauer (2004) and Plag (2003).

nouns are less productive than endocentric and copulative compound nouns. By comparing different types, we can find out if all types of exocentric compound nouns are productive on the same level or if there are internal differences. This will be done by including an analysis that compares examples of different categories according to relevant aspects and their token frequency, using the *Corpus of Contemporary American English* (COCA), the *Oxford Advanced Learner's Dictionary of Current English* and *The Oxford English Dictionary*.

My hypothesis is that not all exocentric compound nouns are equally productive or non-productive, meaning that there is a continuum that sets the range for productivity of exocentric compound nouns.

## HEADEDNESS AND ITS APPLICABILITY IN EXOCENTRICITY

If we want to tackle the subject of exocentric compound nouns, it is important to first look at characteristics that turn them into a distinct category compared to endocentric and copulative compounding. One of the most popular characteristics with which we distinguish exocentric compound nouns is *headedness*. Headedness can be viewed from different perspectives that do not only have a syntactic implication, but also a morphological one.

In general, headedness means that one of the constituent parts of a compound determines not only the word class syntactically, but also the meaning, semantically speaking (see Carstairs-McCarthy 64). If we transfer this definition to exocentric compounding, they are compounds without a head, “having a ‘centre’ outside themselves, figuratively speaking” (65). I include this definition because it is not only the most common but does imply some complications in classifying compounds as exocentric. The actual meaning of *pickpocket*, “a person who steals money, etc. from other people’s pockets, especially in crowded places” (*Oxford Advanced Learner's Dictionary* s.v. *pickpocket*), for example, derives not from its constituent parts, but refers to another centre that is external, so not obvious. However, the invisible part is important because it carries the core meaning of *pickpocket*.

There is another complication in this definition: here are indeed two distinct concepts, semantic and morphological headedness, fused (see Bauer “English Exocentric Compounds” 7). This indicates that although we call *lazybones*

exocentric from a semantic perspective because of its unpredictable meaning, we should not ignore the circumstance that it contains head (*bones*) and modifier (*lazy*) from a morphological perspective (see Bauer “English Exocentric Compounds” 7).

Another aspect neglected in the general definition is that, syntactically speaking, the head of a compound “determines the category, plurality, and other general features of the word” (Di Sciullo, Williams 24). Although this definition is highly functional, as it describes not only where the head is but also which functions it takes in a compound, this notion leaves some questions unanswered. These questions also include how people can know the word class of a compound if the centre is neither included syntactically in the entire compound nor in its constituent parts. The meaning of most of the newly introduced exocentric compounds is opaque until they are looked up in a dictionary. A person who has never heard of a *pickpocket* will not be totally sure if he or she interprets it correctly. That is why Edwin Williams and Anna Maria Di Sciullo propose that “instead of being rejected or restricted, the notion ‘head’ should be relativized” (25–26). This qualification is important because it means that we have to include contextual aspects when defining the head from a morphological perspective (see 26).

Although the head in an exocentric compound noun is not visible, it certainly has a function, which is revealed when taking the context into account. This is the only remaining way to identify the head of an exocentric compound noun. Therefore, the consequence is that syntactically, we cannot find the head of an exocentric compound. In *pickpocket*, it is not the head, but more the usage context that determines its word class. The head in this example is *person*, but only because we understand and use it this way, taking into account our available knowledge to overcome its unpredictability (see Carstairs-McCarthy 93–94).

Without this “relativization” (Di Sciullo, Williams 26), another problem occurs: if an exocentric compound noun is determined by an outlying centre, the function of the constituent parts is likely to be neglected. To prevent this, we have to look closely at the constituent parts and their meaning. *Pickpocket* includes *pick*, a verb, and *pocket*, a noun. If we know the meaning of *pick* and tie this to the meaning of *pocket*, we can have a slight, although not specified, clue what it could mean. Therefore, I agree with the view that constituent parts of an exocentric compound should be seen as those parts which subtly determine the compound because they, albeit not the centre of reference, add to the meaning (see Di Sciullo,

Williams 26). A *pickpocket* is, on a closer look, someone who picks stuff illegally ‘from other people’s pockets’ (see *Oxford Advanced Learner’s Dictionary*), so both parts add here to the overall meaning of *pickpocket*.

Consequently, the head of an exocentric compound noun can only be determined by taking a morphological view, thereby considering the importance of both constituent parts.

## PRODUCTIVITY IN EXOCENTRICITY

Another term that has to be specified with reference to exocentricity is *productivity*. It is important to note that productivity and non-productivity “only mark the end-points of a scale” (Plag 52). So there has to be a middle-position that can be determined by taking into account *formal generality*, *formal regularity*, *semantic regularity*, and *token frequency* (see 51).

Generally speaking, productivity enables us to form an infinite amount of new words and sentences and therefore allows us to extend our vocabulary (see Bauer *English Word-Formation* 63). However, this definition just highlights one aspect of productivity, namely that it is predominated by how many new words are coined in a category, the so-called “type frequency” (Plag 52). There are more relevant factors that influence productivity. This is especially underlined when talking about productivity in relation to exocentric compound nouns. The relation between them is characterised by one major problem, which has its origin in the metaphorical connotation inherent in almost all exocentric compound nouns like *lazybones*. As they are unique and therefore rather opaque, it is a widespread assumption that compounds of this category are not productive (see Carstairs-McCarthy 110). This assumption leads to the question how productivity is measured, which aspects influence the productivity of exocentric compound nouns, and which make them apparently unproductive.

Part of the answer is to distinguish between *formal generality* and *formal regularity* on the one hand, and *semantic regularity* on the other hand (see Carstairs-McCarthy 85-90). In short, a compound is *formally general* when its form reveals the meaning and it is hence transparent (see 86). In contrast to this, *formal regularity* is concerned with two things: whether a word is a possible base with which we can form a compound and whether we can conclude from its base that

the compound is a noun (see 86). However, I doubt that the notion of formal generality can be applied to the concept of exocentric compound nouns in exactly this way, considering the circumstance that these compounds have an opaque meaning due to their phonological state (see Miller 53).<sup>4</sup> Nevertheless, both terms can be considered factors that exert influence on the word-formation process in exocentric compound nouns. Simultaneously, it has to be questioned whether Carstairs-McCarthy's notion that "although not by definition, formal generality presupposes formal regularity, but not vice versa" (88) remains valid in this case, when we hold on to the assumption that exocentric compound nouns are hardly transparent in meaning. In my view, we have to consider the idea that there are patterns in the formation of exocentric compound nouns that show which word classes are often or not often put together, and those word classes that are often put together, especially nouns, are usually the types of compounds that are considered productive (see Carstairs-McCarthy 93).

Another criterion which has to be taken into account is *semantic regularity*, which means investigating the question whether the meaning of a compound can be derived from its constituent parts in a "uniform and consistent" way (Carstairs-McCarthy 88). This concept may be restricted in this context as it was already stated that the meaning of exocentric compound nouns is opaque and cannot be easily derived from their constituents. However, when we see an exocentric compound noun like *skinhead*, which consists of two nouns, I would assume that we interpret it more easily as a person because *head* and *skin* both refer to human body parts and have the same word class. In a compound like *underdog*, on the other hand, where both words have different word classes and the meaning is not transparent, the idea of semantic regularity may be inadequate (see Carstairs-McCarthy 94).

The second part of the answer to how productivity can be measured in exocentricity involves quantifiable measurements that can be applied when working with a corpus. However, this can be a major problem when analysing exocentric compounds, as the amount of newly formed words in a category commonly seems to provide more information than how often an individual word is used. Unfortunately, exocentric compound nouns are numerous. Hence, this

---

4 Phonological state means, there are exocentric compound nouns like *sit-in*, where there is no "phonologically realized element" (Miller 53) that gives a hint that this is a noun because the compound consists of a verb and a preposition (see Scalise, Fábregas, and Forza qtd. in Miller 53).

method is difficult to apply; this even exacerbates the aspect that it is in general very difficult to search for exocentric compound nouns in a corpus. Therefore, a competing view not only questions the significance of type frequency, but also argues that it does not have to be the best measurement when analysing word-formations (see Plag 53-54): Token frequency includes measuring how often a word is used by taking a set of different examples of exocentric compounds (see 54-55).

According to Rolf Harald Baayen (1993), an advocate of token frequency, the knowledge people have determines how they understand and interpret certain formations (see 181). The link to productivity is achieved by assuming that people use a morphological rule in perceiving and producing words when they are not acquainted with a certain pattern of word-formation (see 181). If a person meets a new word such as *hunchback*, the first thing the person will do is to divide the word into its constituent parts according to a “word-formation rule in the mental lexicon” (Plag 54). Consequently, the parts will be probably remembered by the person and this offers the possibility to form a new word with one of the constituent parts such as *greenback* (54). This is a very important point because it stresses the multidimensional character of productivity (see Baayen 181). Thus, I also agree with Baayen that unproductive processes are marked by a large number of high-frequency words and a low number of low-frequency words (see 181).

The opposite case would be if a person meets a familiar noun-noun combination such as *chairman*. The person will possibly not divide this compound into its different parts if there are many words in the familiar category of noun-noun combinations (see Plag 54). The word is therefore remembered as a whole and not as a whole consisting of parts (see Baayen 181). Hence, Baayen’s point of view perfectly corresponds with the concepts of formal generality, formal regularity, and semantic regularity. Formal regularity, on the one hand, is determined by the fact that the person is able to acknowledge a part of a compound as a possible base for another coinage. This influences, on the other hand, formal generality, as the compound becomes transparent because of the words used. Semantic regularity is achieved when the words contribute to a uniform meaning. This relation can be seen when considering examples such as *braveheart* or *faintheart*. If *faintheart* refers to a person, *braveheart*, having the same base, will most likely do, too.

Eventually, productivity in exocentricity is determined by formal generality, formal regularity, and semantic regularity, and adequately measured by token frequency.

## CONCEPTS OF EXOCENTRIC COMPOUND NOUNS: MARCHAND'S TYPE DISTINCTION

The topic of exocentric compound nouns does not only concern today's linguistic research. It is a topic that has a long tradition throughout history, especially in relation to descriptivist approaches (see Benczes 16). One of those descriptivist approaches is the one by Hans Marchand; his analyses predominantly exert influence on today's understanding of how to analyse and understand exocentric compound nouns.<sup>5</sup>

Marchand introduces the general term 'compound' as "made up of a determining and a determined part" (11). Referring to exocentric compound nouns, this means that the determining part is non-existent considering formal aspects and the compounds in this category are therefore zero morphemic, which means that none of the constituent parts determines the meaning of the compound: "A *pickpocket* is neither a *pick* nor a *pocket*..." (11). However, the underlying concept is understood by people, so the non-existent part only refers to formal aspects and not semantic aspects (see 11). This notion has exerted great influence on today's understanding of exocentric compound nouns, especially on the term *head*, which can be seen synonymous to Marchand's *determinatum*.<sup>6</sup> Applying this view on exocentric compound nouns, Marchand proceeds his analysis by introducing major categories of exocentric compound nouns (see 37-46).

He begins with type *pickpocket* which is described as a formal connection of a predicate and object that indicates what is performed by the agent, including nicknames, animal names and "impersonal material agents" (37-38). He proceeds by finding connections to type *runabout* which he sees as a development from type *pickpocket* (see 39). What in his view makes these two types related is the underlying concept, namely that both seemingly have a negative impact, although type *runabout* consists of an adverbial complement in a verbal phrase, which specify together what the agent is doing (see 38; see Benczes 19). However, Marchand makes an exception for animal names such as *wagtail* because they lack this derogative undertone (see 38). Furthermore, Marchand specifically refers to type

5 This chapter is based on Marchand's monograph *The Categories and Types of Present-Day English Word-Formation* (1960).

6 This influence reveals itself by taking into account the investigations by Williams and Di Sciullo (1987).



*pickpocket* as a spoken phenomenon because under historical aspects, words in this category were derived from Roman languages that did not have a literary tradition (see 38). Type *pickpocket* is seen as “exceedingly productive” (39), meaning in this case that there are many examples in this category, which he also enumerates (see 38). Therefore, it can be said that he sees type frequency as the factor that divides productive from non-productive types. This is especially underlined by the circumstance that he does not go into deeper detail with type *dugout*, because this category does not consist of many examples, so in his view there is no necessity to examine them (see 45).

The last types he introduces are type *blackout* and a category consisting of the subcategories *hunchback*, *paleface*, *five-finger*, and *scatterbrain* (see 41-45). Type *blackout* is here explained as an activity pointed out in the “verbal phrase” (41). *Blackout*, in this case, describes an action of forgetting (see *Oxford Advanced Learner's Dictionary* s.vv. *blackout*, n.5). While he examines this type, he takes apart the semantic and morphological aspect. This type is semantically close to type *pickpocket* on the one hand and shows morphological similarity to type *blackout* (see Marchand 41; see Benczes 19). The semantic similarity is justified by the underlying concept that the agent is doing something that is expressed in the compound (see Marchand 41). However, *pickpocket* and *blackout* are morphologically different, as indicated by the word classes involved (see 41). Therefore, *blackout* implies morphological similarity to type *runabout*, considering that both consist of a verbal phrase together with an adverbial complement (39-41). Taking the similarities into account, it can be argued whether or not *blackout* is different enough to build an own group (see Benczes 19).

Type *hunchback* is concerned with a characterisation namely that a person or thing has an outstanding characteristic that describes them perfectly (see Marchand 42). In my view, this explanation slightly touches the topic of synecdoche, meaning that one single unit describes one whole, the so-called *pars pro toto* (see Bauer “English Exocentric Compounds” 7). A person not only consists of one characteristic, but one special character trait suffices to identify the person (see Marchand 42). What Marchand also points out is that type *hunchback* seems to be very near to endocentricity because it is very difficult to define if a *hunchback* describes a person having a *hunched back* or if a *hunchback* is a certain kind of *back*, namely a hunched one (see 42). Considering this aspect, it can be argued whether this indicates the beginnings of semantic and syntactic exocentricity,

which is an on-going discussion (see Bauer “English Exocentric Compounds” 1). Indeed, from a semantic perspective, a *paleface* is a nickname for a person, but syntactically, it describes a face that is pale (see Marchand 42). Why he especially refers to *paleface*, *five-finger*, and *scatterbrain* as distinct subcategories does not become clear in this context. He refers to them as “bahuvrihi” (42),<sup>7</sup> but examines their productivity separately without giving much explanation. However, he generally describes them as being characterised by what the compound expresses (see 42). Nevertheless, he makes an exception for *five-finger* because it does not fit in his perception of bahuvrihi (see 45).

What becomes apparent in Marchand’s approach of how to separate exocentric compound nouns in distinct categories is the diachronic perspective, namely that he not only describes distinctions between the types, but also draws connections between them while regarding their historical origin and development through time. This is especially emphasized by type *pickpocket*. Generally speaking, there are some aspects to be discovered that remain valid until now, indicated by his distinction of *determinant* and *determinatum*, and the term *zero morpheme* (see Marchand 11).

## BLOOMFIELD’S LINES OF CLASSIFICATION

Another descriptivist approach, although slightly different, is the one Leonard Bloomfield introduced in his monograph *Language* (1933). He categorises compounds based on *two lines* (see 233-237).

First, he looks at the relation between the parts constituting the compound (see 233). Two main types of compounds result from this perspective, namely *syntactic* and *asyntactic* compounds. If a compound is syntactic considering its constituents, these two parts resemble the identical grammatical order “as a word in a phrase” (233). To name one example, we can say that *greenhouse* as a compound consisting of an adjective and a noun shows the same structure as the comparable phrase *green house* (see 233). On the other hand, asyntactic compounds show the exact opposite (see 233). *Chairman* would therefore be asyntactic because we can

---

7 Bahuvrihi compounds are exocentric and give a characterisation of someone or something by referring to a striking characteristic; a *paleface* can, for example, describe a person who has an extraordinarily white face (see Marchand 42).

find no phrase such as *chair man* (see 233; see Benczes 16). However, Bloomfield also mentions compounds that take a middle position between these types, namely “semi-syntactic compounds” (234). They are seen as “intermediate” (234), because there are compounds like *undertaker* which have a syntactical order but have a different order in a phrase. The adequate phrase for *undertaker* would be *take under*, so the order is totally different than the one in the compound (see 234). This line of classification is, slightly hinting at future results, more important than assumed at first sight.

He goes on by referring to his second line of classification, namely the relationship of the compound to its constituent parts (see 235), which I think is more important in the consideration of exocentricity. Whereas before he remained mainly general in his explanation of syntactic and asyntactic compounds, as he did not fully specify it to the case of exocentricity, this line of categorisation specifically focuses on exocentricity. This is proven by Bloomfield's assumption that if the head member of the compound and the whole compound itself resemble differing functions considering their word class, the compound will likely be described as exocentric or “bahuvrihi” (235). The main problem appears when the compound as a whole has the same grammatical category as the head member because in his view the compound can no longer be classified as exocentric from a formal perspective. As the only solution to this problem, he proposes the context as the remaining way to figure out the meaning of an exocentric compound (see 236).

There are several points to discover in this explanation. Bloomfield tries to avoid the problem of classifying a compound as exocentric by including different perspectives, a syntactic as well as a semantic view. As I discussed at the beginning of this paper, the head member is not visible formally in exocentric compound nouns, so semantic criteria play a major role in identifying a compound as exocentric. His analysis, although remaining descriptive, as he does not prescribe distinctive rules of how to classify, gains a formal aspect because he takes into account syntactic criteria besides semantic criteria. Therefore, his approach exerts a main influence on today's understanding concerning this topic, as discovered in the previously introduced terms *headedness*, *formal generality*, *formal regularity*, as well as *semantic regularity*.

## COMPARISON AND CRITICISM

As we have seen in the previous subsections, both Marchand (1960) and Bloomfield (1933) exerted main influence on the topic of exocentric compound nouns and how to classify them. Albeit descriptivist, both approaches have internal differences not only in their general understanding, but also in their logical conclusions. Some of the differences and criticism concerning both approaches will be pointed out in the following.

The most obvious difference to be discovered is the base of analysis both choose. The main essence of Marchand's approach is to present several types of exocentric compound nouns and to refer to their development throughout the years, hence adopting a diachronic approach (see Benczes 18). Unlike Bloomfield, who neglects the diachronic part, Marchand seemingly tries to compensate his lack of systematic analysis by counting examples for each type. The problem is that he does not set clear boundaries in his type-distinction, proven by the obviously insufficient way he examines the types *hunchback*, *paleface*, *five-finger*, and *scatterbrain* (see Benczes 18). He tries to draw connections between the types, shown by the fact that he states that type *runabout* came from type *pickpocket* (see Marchand 39). Exactly at this point there are logical inconsistencies to be discovered. First, type *pickpocket* is here described as consisting of an agent performing the action of the "nexus" (37). Marchand tries to find the link to imperatives, which he justifies with the negative impact compounds in this category show (see 37). However, imperative does not seem to be an adequate term, as *pickpocket* for example is not an order to people to steal things from other people (see Bauer "English Exocentric Compounds" 2). However, his general explanation of type *pickpocket* seems adequate considering the example of *pickpocket*, where the agent is a person performing the action of stealing, indicated by the part *pick*. Although I have to qualify that the performance is visible as the adequate verb *to pickpocket* exists (see *OED* s.vv. *pickpocket*, n.2), problems can occur if the person confronted with the noun does not know the adequate verb. Then the performance is, as well as underlying agent, not transparent. Again, the context in which the word is used has to be taken into account.

However, this semantic aspect is pointed out by Bloomfield (see 236). Therefore, Bloomfield's approach opposes Marchand's rather unfounded explanation, reflected in the lack of explanation of type *dugout*. He, in contrast,

proposes two ways to find out if a compound is exocentric by taking into account the relation of the parts constituting the compound and how the constituents relate to the whole compound (see 233-237). Bloomfield repeatedly uses the term *bahuvrihi* as a synonym for exocentric whereas Marchand uses it when talking about type *hunchback*. Therefore, Bloomfield's use of this term raises criticism as bahuvrihis are mostly adjectives and do not occur that often among exocentric compounds to use them as a synonym for exocentric (see Bauer "English Exocentric Compounds" 1).

Hence, it can be said that Bloomfield has an entirely different starting point to classify in contrast to Marchand, who proposes basic types of exocentric compound nouns. I have to qualify that Marchand refers more to productivity than Bloomfield does, indicated by the fact that he enumerates examples for each type to justify his line of argumentation. Nevertheless, there is no mentioning of the word class of the head in comparison to the word class of the whole exocentric compound, which is replaced by the underlying idea with which Marchand is mainly concerned, underlined by words such as "pejorative tinge" (37). It may be easier to track Bloomfield's line of argumentation because it appears more funded and thought-through, whereas considering Marchand's approach, the reader is more or less left with the question of how exactly he classified the compounds on an objective, explicit base (see Benczes 18). Productivity predominates in his approach, so it is self-evident that he neglects certain aspects. In contrast, Bloomfield does not take into account that processes develop in their productivity over years and therefore ignores productivity to some extent.

Eventually, both approaches must not be seen as separate, but complementary units. Marchand's examples have exerted great influence on linguistic research. This is mainly indicated by Thomas Biermeier's list of exocentric compound nouns which consists of a lot of examples from Marchand's investigations (see Biermeier 2008) and the constant criticism linguists utter referring to his distinction (see Bauer 2004). In contrast, Bloomfield's criteria have the function to determine whether a compound is exocentric or not, so they serve as a proof-read and general idea, although the synonymous use of *bahuvrihi* is highly criticised.

## ANALYSIS OF SELECTED EXAMPLES

After having considered different influential factors and pre-understandings, I will now come to the comparison of selected examples. Before that, I would like to make some general remarks. This comparison is based on the aforementioned word list of exocentric compound nouns by Biermeier (2008), which I separated into noun-noun, adjective-noun, verb-noun, and preposition-noun categories. I also added examples by searching the corpus and considering some of Marchand's (1960) examples. That means that the lists consist of subjective picks and therefore have no right to be called representative for the whole categories. The aim is to examine whether we can find and prove the factors that were discussed as relevant, such as formal generality, formal regularity, and semantic regularity, while considering token frequency of these categories. This analysis also includes how token frequency is in general distributed among the categories *spoken* and *fiction*. I chose them as spoken and fiction can be seen as juxtaposed genres and the distribution provides information on whether exocentric compound nouns are in general more important in spoken or in fictional language. I have considered those exocentric compound nouns that are actually used as nouns on grounds of specification and limitation. Therefore, I made use of the POS-function in the corpus. If you search a word, for example *blackout*, in a corpus and you use the POS-function by selecting 'noun.ALL', the corpus will automatically restrict the search to those cases, where *blackout* appears as a noun.

Due to the fact that my analysis is grounded on COCA, I will quickly present the corpus. The Corpus of Contemporary American English, mostly founded by Mark Davies, Professor of Linguistics at Brigham Young University in Provo, Utah, is freely available. It covers a large range with its collection of over 520 million words of text, which are made up of 20 million words of each year from 1990 to 2015. It also contains a subdivision into spoken, fiction, popular magazines, newspapers and academic texts (for more information see: <http://corpus.byu.edu/coca/>).

## DIFFERENT EXAMPLES IN COMPARISON

First, I will start by presenting my findings and open the discussion of the relevant factors afterwards. On the surface level, we can say that the noun-noun combinations listed in table (2) (see page 15) can be considered the most frequent type, followed by verb-noun constructions in table (3) (see page 15), preposition-noun constructions in table (4) (see page 15), and at last adjective-noun combinations in table (1):

		Token Frequency	Spoken	Fiction
1	redneck	444	78	162
2	hunchback	269	21	95
3	busybody	118	4	70
4	greenback	97	13	6
5	loudmouth	91	11	33
6	dreadnought	59	1	19
7	braveheart	52	15	0
8	redbreast	15	0	6
9	paleface	13	1	7
10	lazybones	11	0	10
11	scatterbrain	10	0	7
12	greybeard	6	0	4
13	faintheart	1	0	0

Table 1: adjective-noun

		Token Frequency	Spoken	Fiction
1	chairman	34013	9807	970
2	paperback	2421	385	486
3	goosebumps	423	75	206
4	skinhead	289	86	90
5	heartthrob	237	66	18
6	brainstorm	193	19	50
7	bookworm	120	19	41
8	egghead	78	9	30
9	killjoy	46	6	20
10	butterfingers	8	1	3

Table 2: noun-noun

		Token Frequency	Spoken	Fiction
1	playboy	1183	372	159
2	scapegoat	755	202	84
3	scarecrow	575	30	411
4	copycat	374	83	67
5	breakwater	184	3	60
6	pickpocket	150	18	86
7	crybaby	127	18	60
8	spoilsport	24	0	12
9	cut-throat	9	0	4

Table 3: verb-noun

		Token Frequency	Spoken	Fiction
1	underworld	1209	117	359
2	underdog	1032	229	56
3	undertaker	343	30	201
4	undertone	311	19	164
5	underboss	26	11	6
6	underdark	8	0	8
7	undershaft	1	0	0
8	undereating	1	0	0
9	underoath	1	0	0

Table 4: preposition-noun

I have to qualify that the noun-noun category is mainly influenced by the extreme high token frequency of *chairman* with a token frequency of 34013. Considering the distribution of all the individual compounds in each category, a very consistent pattern can be discovered, namely that most of the words in each category are more used in fiction than in spoken, and even when there are some exceptions, they do not have a high token frequency. This may lead to the conclusion that exocentric compound nouns are mainly produced in a creative writing process where it is the aim to achieve a certain effect, whereas spoken language has the aim to be efficient and easy. As this is just an assumption, this would have to be proved by further investigations.

Albeit this obvious finding that most words are used in fiction, there are internal differences between the categories reflected in the different token frequency. Especially the problem with headedness seems to exert influence here.



If I compare for example *brainstorm* with *loudmouth*, it can be seen that *brainstorm* is more metaphorically connotated than *loudmouth*, so the head will be very hard to identify, even with the help of context. In fact, the compounds of table (1) follow a certain pattern, namely the pattern of specifying a certain personal characteristic (see Bauer “English Exocentric Compounds” 4). There may be a similarity to Marchand’s approach to be discovered, but it has to be mentioned that I draw the line between the categories according to the word classes involved and not according to their underlying sense (see Bauer “English Exocentric Compounds” 4-7).

In contrast to the adjective-noun combinations, most of the noun-noun combinations (see table (2)) are not only highly metaphorically connotated, they also barely follow any pattern. Therefore, we can say that according to their meaning, these examples are, in contrast to the listed adjective-noun combinations, less formally general because in lacking a consistent pattern, we can hardly understand what this compound could mean. Taking the relation of formal generality and formal regularity in account, this aspect seems to prove valid in both categories. It was simpler to find adequate examples in the adjective-noun category than in the noun-noun category, so people rather acknowledge an adjective together with a noun as a new coinage than two nouns tied together. Additionally, the restricted set of examples in the exocentric noun-noun category is already noticed in further academic investigations (see Plag 146). However, this does not have to mean that this category is not productive (see Bauer “English Exocentric Compounds” 6).

This aspect is also reflected in considering the token frequency of both categories. As I have mentioned, a high token frequency does not have to result in high productivity. It is more the other way around that many examples with low token frequency reveal productive patterns. This assumption may be proved valid in the adjective-noun category. This category namely shows a lower token frequency than those of the noun-noun-category in table (2). A reason for this occurrence could be that a lot of exocentric compound nouns in this category are rather old-fashioned, and thus have a low token frequency because of the lack of use (see Bauer “English Exocentric Compounds” 2). However, if we look at the single units of the compounds in the adjective-noun category, there are many parts that are redundant in different words, such as *heart* in *faintheart* or *braveheart*, or *back* as in *hunchback* or *greenback*. This could have many reasons. As *heart* and

*back* are associated with humans and this category reveals a rather consistent pattern, people coming across those words take these parts of the compound and connect this with another characteristic to specify it. A *hunchback* can be therefore transformed into a *greenback*. Although the whole compound may be not transparent, the parts are and consequently both are remembered separately and used for coining new words (see Baayen 181). Another possible explanation of this could be that all compounds in table (1) are syntactic, so they exist in the same order as complements in a sentence. A *redneck* also exists as a *red neck*, so people are indeed syntactically familiar with this combination. Bloomfield's first line of classification can therefore be considered very useful in examining the categories (see 233). In contrast, this cannot be said for the noun-noun category as the examples are of high frequency, but lack redundant parts. They are asyntactic, as there is no possible combination like *goose bumps*. This may be the reason why these examples lack transparency. I assume that a *chairman* is remembered as one full expression, and not an expression consisting of *chair* and *man*. Considering semantic regularity, both categories do not provide any consistency. This may be because uniformity and consistency do not correspond with flexibility, which I think is the key to finding new coinages.

The next category that I will now focus on is the one consisting of verb-noun combinations, pointed out in table (3). This category is mainly characterised by the pattern of *a person or thing that does something* (see Marchand 37). A *spoilsport* is indeed someone who 'spoils' fun (see *Oxford Advanced Learner's Dictionary* s.v. *spoilsport*). Albeit this seemingly thorough pattern considering the examples in table (3), those are not as formally general as the adjective-noun compounds, but more formally general than the noun-noun compounds in table (2). Some of the examples could be considered syntactic like *spoilsport*, but most of them like *playboy* are asyntactic. This is emphasized by the fact that not all follow the more or less same pattern as in the adjective-noun category. A *copycat* is not a person who copies a cat (see *Oxford Advanced Learner's Dictionary* s.vv. *copycat*, n.1), so the meaning remains opaque from its form in this case. Therefore, I assume that these verb-noun constructions are also less semantically regular because they do not always refer to someone, who does something, especially considering a compound like *scapegoat*. Although the verb-noun examples are numerous, they do not fully provide a transparent meaning as in *cut-throat*. This also corresponds with the former assumption that type frequency does not always have to say

something about the productivity of words (see Plag 53f.). Consequently, I consider the examples in this category as less formally general due to their unpredictability. They are also less formally regular because most of the words do not resemble redundant parts, compared to the adjective-noun category. The token frequency in this category is higher than in the adjective-noun category, but the verb-noun category resembles a similar frequency to the noun-noun category. This could imply that the verb-noun category is also the one consisting of compounds that are seen as a whole and not as separate parts which could be a possible basis for another coinage. This assumption would correspond with the low formal regularity I mentioned.

The last category is the preposition-noun category in table (4). This category was probably the one where it was hardest to find any adequate examples because prepositions like *off* or *by* are mostly connected with other word classes, for example in *offshore*, or are used as an adjective when they are formed with a noun. In addition, the examples in this category can be considered asyntactic, as *underdog* does not exist as *under dog* or *dog under*. Therefore, I can say that this category seems rather restricted, also taking token frequency into account, which differs widely among the examples. However, table (4) nevertheless provides useful information on the productivity of this category. They are indeed the perfect demonstration of the fact that the meaning in exocentric compound nouns can vary, although the same word is used. All of the examples are formed with *under*, but *under* does not always mean the same when they are put together with another word. In *underworld*, *under* is commonly understood as 'hidden' or 'not visible' (see *Oxford Advanced Learner's Dictionary* s.v. *underworld*, n.2); in *underdog* it means 'inferior' (see *Oxford Advanced Learner's Dictionary* s.v. *underdog*). Leaving out the fact that these examples always include *under*, they do not have any common pattern like the other compounds. The reason for this can be that prepositions may not be considered as a good base for exocentric compound nouns, so they are not that formally regular and people are not acquainted with the underlying pattern. The relation between formal generality and formal regularity seems valid here, as this category lacks formal regularity and the presupposition for formal generality is therefore restricted, reflected by the few examples in combination with high token frequency. Semantic regularity could also not be found, due to the lack of pattern. Examining the list of preposition-noun combinations, we might notice that *undershaft*, *undereating*, and *underoath* only occur once as a noun. However, this

does not mean that these are *hapax legomena*,<sup>8</sup> as these words occur only as nouns once. In other categories, when used as adjectives for example, they occur more often. The use of exocentric compound nouns, especially in this category, seems therefore flexible. Another reason to call these set of examples unproductive in contrast to the other ones is that this category is not even mentioned in the literature I chose, whether the recent ones like Carstairs-McCarthy (2002) and Bauer (2004) or older ones like Marchand (1960) and Bloomfield (1933). This category seems rather unexplored, so it is no surprise that there are not many examples.

## CONCLUSION

Concluding from these findings, I can say that the adjective-noun category, presented with its set of examples, is the most productive, followed by the noun-noun category, the verb-noun category, and at last the preposition-noun category. This observation is also supported by other linguistic investigations (see Plag 146). I have pointed out the link between the level of pattern consistency of a category and the appearance of the characteristics formal regularity, formal generality, and semantic regularity. A reason for transparency and intransparency in the examined categories was found when taking Bloomfield's first line of classification into account. Therefore, categories consisting of exocentric compound nouns which are mostly syntactic (table (1)) are considered more productive than those that are mostly asyntactic (tables (2), (3), (4)). This may also lie in the fact that the lists consist of a subjective pick, so this result can only be related to these examples. However, semantic regularity has been proven rather restricted in this context, due to the flexibility involved in coining new words in the category of exocentric compound nouns. This link was made according to the token frequency where I found out that the categories with a respectable low token frequency are the ones considered very productive. The distribution among the genres did not provide as much information as I hoped, except the information that most of the compounds presented are more used in fictional language than in spoken language. This emphasized my assumption that most of the compounds are created in a process of

---

8 *Hapax legomena* are words that just occur once in a corpus (see Carstairs-McCarthy 96); the word is of Greek origin and means "said only once" (96).

constructiveness and that consequently, the forming of exocentric compound nouns is motivated in some sense.

As I have tried to make clear in this paper, the study of productivity according to the different categories of exocentric compound nouns affords multidimensional perspectives considering their productivity and classification, especially outlined by the different pre-understandings and aspects that were discussed. It is therefore not sufficient to just look at certain characteristics and study them apart from each other. The link between formal generality and formal regularity could be partly verified because the given examples hardly represent whole categories consisting of thousands of words. My two hypotheses - there is a continuum that sets the range for productivity and not all of the exocentric compound nouns are productive on the same level - could be verified in the context of the paper. However, not all of the categories provide useful information, especially underlined by the preposition-noun category. Therefore, further studies could gain more knowledge in this category and investigate whether this category is really marked by unproductivity.

## WORKS CITED

- Baayen, Rolf Harald. "On Frequency, Transparency, and Productivity." *Yearbook of Morphology 1992*, edited by Geert Booij and Jaap van Marle, Springer, 1993, pp.181-208.
- Bauer, Laurie. *English-Word-Formation*. Cambridge University Press, 1983.
- . "English Exocentric Compounds." Victoria University of Wellington, 2004, [www.victoria.ac.nz/lals/about/staff/publications/Bauer-EnglishExocentricCompounds.pdf](http://www.victoria.ac.nz/lals/about/staff/publications/Bauer-EnglishExocentricCompounds.pdf). Accessed 7 July 2017.
- Benzces, Réka. *Creative Compounding in English*. John Benjamins Publishing Company, 2006.
- Biermeier, Thomas. *Word-Formation in New Englishes: A Corpus-based Analysis*. LIT Verlag Hopf, 2008.
- Bloomfield, Leonard. *Language*. The University of Chicago Press, 1933.
- Carstairs-McCarthy, Andrew. *An Introduction to English Morphology: Words and their Structure*. Edinburgh University Press Ltd., 2002.
- Davies, Mark. "FAQs." [corpus.byu.edu/faq.asp](http://corpus.byu.edu/faq.asp). Accessed 5 August 2017.
- . *The Corpus of Contemporary American English*, [corpus.byu.edu/coca](http://corpus.byu.edu/coca). Accessed

- 5 August 2017.
- Di Sciullo, Anna Maria, and Edwin Williams. *On the Definition of Word*. MIT Press, 1987.
- Marchand, Hans. *The Categories and Types of Present-Day English Word-Formation*. Otto Harrassowitz, 1960.
- Miller, D. Gary. *English Lexicogenesis*. Oxford University Press, 2014.
- Oxford Advanced Learner's Dictionary of Current English*. 6th ed., Oxford University Press, 2000.
- Plag, Ingo. *Word-Formation in English*. Cambridge University Press, 2003.
- The Oxford English Dictionary*, oed.com. Accessed 8 August 2017.