

Expanding the boundaries of word-formation – What German phrasal compounds tell us about the lexicon

Katrin Hein-Antonioli
Leibniz Institute for the German Language (IDS)
hein@ids-mannheim.de

1. Introduction

Phrasal compounds like *Leck-mich-am-Arsch-Tag* ‘kiss-my-ass day’ or *Beckham-im-Gedächtnis-Grätsche* ‘Beckham-in-mind straddle’ are characterized by the integration of a syntactical unit (an entire sentence or syntactic phrase) into a complex word. Even if they deviate from the prototypical pattern of compounding and (among others) violate the Principle of Lexical Integrity (cf. Spencer 1991: 42), there is strong empirical evidence that we are dealing with a productive process of German word-formation which is by no means irregular or marginal (cf. Hein 2015). The sketched-out divergence from canonical compounding can come with expressive effects (cf. Meibauer 2007) and justifies the labeling of phrasal compounds as a phenomenon of extravagant morphology in the sense of Haspelmath (1999).

By integrating a phrasal component into a word, phrasal compounds expand the boundaries of word-formation. Thus, we are dealing with an interface phenomenon between morphology and syntax, and also with a perfect example for the interplay of word-formation and multi-word expressions (cf. Steyer & Hein 2018). In the linguistic literature, phrasal compounds are part of a general debate (which is not limited to German) about “the role of phrases in the word formation process” (Radimský 2015: 33) and also provide interesting insights into the relation between creativity and routine.

Due to their status as an interface phenomenon, phrasal compounds are challenging for morphological theorizing and for an adequate conception of the lexicon. This observation is the starting point for this article. With reference to corpus data from written and spoken German¹, I will pursue the question of what phrasal compounds tell us about the lexicon.

Following Gaeta (2016), phrasal compounding is a phenomenon that can be understood by invoking “two different senses of the term lexicon”: the Bloomfieldian sense which “generally refers to the set containing any sort of entrenched or idiomatic expressions” and a second sense which “refers to the ‘infinite set of potential [...] lexemes for any given language’” (Gaeta 2016: 109; Aronoff 1994: 10). As I will argue in this article, this observation clearly holds for phrasal compounds whose non-head can be analyzed as a fixed multi-word unit, e.g. *Alles oder nichts-Devise* ‘all-or-nothing motto’, but is less convincing for formations whose non-head is formed ‘on the fly’ (e.g. *Guck-ich-hab-es-wieder-aus-dem-Bett-geschafft-Rap* ‘Look-I-made-it-out-of-bed-again rap’. The latter are clearly attested in corpus data from written and spoken German and play a significant role for the pattern of phrasal compounding, even if lexicalized non-heads dominate quantitatively (cf. Hein & Antonioli 2025; Gaeta 2016: 136). Then the

¹ More precisely, I will refer to newspaper texts from the German Reference Corpus (DeReKo; <https://www.ids-mannheim.de/en/digspra/pb-s1/projects/corpus-development/>), to song lyrics provided in the song corpus (<https://songkorpus.de/>) as well as to spoken interaction from the FOLK corpus (<https://dgd.ids-mannheim.de/DGD2Web/jsp/Welcome.jsp>).

question arises where those non-heads “made out of the blue” (Gaeta 2016: 136) are built and what consequences they bring for an adequate conception of the lexicon.

Before I turn to this question in more detail in the main part of this article (section 3), I will first provide an overview of phrasal compounding in German (section 2) sketching out the main properties of the phenomenon as well as its treatment in the literature. In the conclusion (section 4), I will also make a connection to Construction Grammar respectively Construction Morphology (Booij 2010). This theoretical framework seems to be best suited to describing an interface phenomenon like phrasal compounding.

2. Phrasal compounding in German

The aim of this section is to carve out the characteristics of phrasal compounding in German, with particular regard to the notions of creativity and extravagance.

2.1 Central properties and ignorance in the research literature

With the integration of a syntactic unit into a complex word, phrasal compounds form a special type of determinative compound; more prototypical instances are words like *Baumhaus* ‘tree house’. Nonetheless, their status as genuine compounds is uncontroversial (cf. Schlücker 2012; Hein 2011), as they are characterized by a binary structure and a classical determinans-determinatum relation between the immediate constituents. Thus, the head (*Rap* in Fig. 1) determines the main semantic class as well as the formal properties (part of speech, gender, inflection) of the corresponding complex word, while the non-head (*Ich finde du bist weak* in Fig. 1) specifies the meaning of the head semantically:

[“Ich-finde-du-bist-weak”]_{nonhead} - [Rap]_{head}
(‘I-think-you-are-weak rap’)

Figure 1. Binary structure of phrasal compounds

Nonetheless, phrasal compounds are less frequent than prototypical N+N-compounds, and they also go against some word-specific properties like Lexical Integrity (cf. Lieber & Scalise 2006; cf. Hein 2011: 355–356 for an overview).

According to the research literature, phrasal compounding is mainly used to form nouns (cf. Lawrenz 2006), i.e., nominal phrasal compounds are the most productive subtype of phrasal compounding. Also, the only empirical study available for German to date focuses on nominal phrasal compounds (cf. Hein 2015). However, it is claimed in the research literature that adjectival phrasal compounds can also be formed (cf. Günther, Kotowski & Plag 2020). This aligns with the attestation, though very rarely, of adjectival phrasal compounds in the data of my recent investigation of word-formation products in German song lyrics (cf. Hein 2023), e.g. *nach Oma-Schlüpfer-riechendes (Miststück)* ‘smelling-like-granny-panties (bitch)’ (Corpus of Song Lyrics: Die Ärzte/MISC).

Please note that the phenomenon is not limited to expressive text types like song lyrics. Rather, it is also widespread in less marked text types like newspaper texts. Moreover, I want to highlight that not all phrasal compounds are as expressive as the previous example. Rather, with the help of corpus data one can observe that the expressivity (and also the extent of creativity) of phrasal compounding varies gradually. On the one hand, there are completely lexicalized phrasal compounds like *Armeleuteessen* ‘poor-people meal’, which have their own lexical entry and thus are not very remarkable. The same holds for compounds containing a modifier with numerals (e.g. *1000-Stunden-Jahr* ‘1000 hours year’, a very productive subtype of phrasal compounding. On the other hand, there is a variety of phrasal compound whose modifier is a complex sentence formed on the fly. Corresponding formations like *Guck-ich-*

hab-es-wieder-aus-dem-Bett-geschafft-Rap ‘Look-I-made-it-out-of-bed-again rap’ are very expressive and display a high degree of creativity. As those examples indicate, expressivity does not only depend on lexicalization, but is also influenced by the syntactic complexity of the first constituent.

Accordingly, the construction grammatical modelling carved out in Hein (2015) claims that phrasal compounding can be traced back to two main sub-patterns: One relatively unmarked construction which has (almost) no expressive effects and one marked construction whose instantiations are quite expressive and can fulfill specific communicative functions (cf. also Hein & Antonioli 2025: 68–69).

In the past, the distinctive features of phrasal compounding were often equated with irregularity. Accordingly, phrasal compounds were long considered an unproductive, marginal phenomenon that violates central principles of word-formation. As a consequence, they have been marginalized in the research literature (cf. Meibauer 2003). Even worse, their existence was denied in some models of generative grammar (cf. Hein 2011).

However, even if phrasal compounds violate the Principle of Lexical Integrity and deviate from the prototypical pattern of compounding, the first broad corpus-based study of nominal phrasal compounding in newspaper texts (Hein 2015, cf. 2017 for a shorter version in English) has shown that we are dealing with a productive process of German word-formation which is by no means irregular or marginal. Follow-up studies on song lyrics (Hein 2023) and spoken German (Hein & Antonioli 2025) confirm those findings.

Recently, there is a growing (and theoretically motivated) interest in phrasal compounding which is not limited to German, cf. Goldberg & Shirtz (2025), Hein & Antonioli (2025), Gutzmann & Turgay (2024), Günther, Kotowski & Plag (2020), Hein (2017, 2015, 2011), Gaeta (2016), Meibauer (2015, 2007, 2003), Trips & Kornfilt (2017), Lieber & Scalise (2006), etc.

Due to their status as an interface phenomenon between morphology and syntax, phrasal compounds are challenging for morphological theorizing. For example, they pose a severe problem for models of generative grammar that assume a linear interaction between morphology and syntax. As I have shown in previous work (Hein 2015; cf. also Goldberg & Shirtz 2025: 307), a construction-grammar approach is fruitful and suitable for capturing a phenomenon like phrasal compounds (cf. section 4).

2.2 Creativity & extravagance

As stated above, phrasal compounds display a classical determinans-determinatum relation between their immediate constituents which suggests that they are indeed compounds. Nonetheless, due to the integration of a syntactical unit into a complex word we are dealing with a phenomenon that somehow diverges from the ‘routine in compounding’, i.e. from canonical compounds like *Baumhaus*. As already mentioned above, this can cause expressive effects (cf. Meibauer 2007) and justifies the labeling of phrasal compounds as a phenomenon of ‘extravagant morphology’ (cf. Haspelmath 1999; Eitelmann & Haumann 2022a). Accordingly, the use of phrasal compounds may be in part explained by Haspelmath’s (1999: 1063) ‘maxim of extravagance’ (“Talk in such a way that you are noticed”).

„The maxim of extravagance seeks to grasp the observation that communicative behaviour does not boil down to achieving the goal of being understood [...] In order to be noticed, language users would also overstep boundaries and ‘choose new ways of saying old things’.” (Eitelmann & Haumann 2022b: 2)

In a recent publication by Goldberg and Shirtz (2025: 313), the witty and sarcastic potential of phrasal compounds is confirmed by various large-scale surveys:

“Sentences with PALs [Phrase-As-Lemma; K.H.] are also judged to be more witty [...] and more sarcastic [...]. Remarkably, these factors are evident even in extremely frequent PAL tokens when compared to paraphrases [...]. To be sure, we are not claiming that all or most PAL tokens imply a close intimacy with the comprehender, are laugh-out-loud funny, or are witheringly sarcastic. But naive English speakers recognize that the construction adds a dash of these rhetorical flourishes.”

This is in line with the claim mentioned above that the degree of expressivity in phrasal compounds is gradual (cf. section 2.1).

As the notion of creativity used above is quite fuzzy, I would like to clarify it with the help of Schmid (2025) who addresses the question of “What is Creative to Whom and Why?” and makes a crucial reference to Sampson’s (2016) distinction between F-creativity and E-creativity. While more general definitions use adjectives like ‘original’ or ‘unconventional’ to capture creativity, Sampson brings the criteria of regularity into play: While *fixed creativity* (F-creativity) captures word-formation phenomena “that are based on existing word-formation patterns”, *extending creativity* (E-creativity) “denotes original formations that go beyond regular applications of productive patterns [...] and tend to be salient enough to be recognized as being motivated by the intention to be creative” (Schmid 2025: 98).²

Turning back to the phenomenon in focus, I would argue that in one sense we are dealing with a case of E-creativity. This applies if we relate phrasal compounding to the more productive pattern of ‘normal’ compounding (e.g. *Baumhaus*). Under this perspective, phrasal compounds “go beyond regular applications of productive patterns” (Schmid 2025: 98). This is due to the fact that the non-head position of phrasal compounds is filled by a syntactic unit, while the non-head position in a general schema for compounding normally is reserved for lexical categories. On the other hand, one could also argue that phrasal compounding itself is a productive pattern of German word-formation and thus can be seen as a manifestation of F-creativity.

The observation that phrasal compounds are salient in the sense that the intention to be creative is apparent is in favor of their treatment as a case of E-creativity (cf. Schmid 2025: 98). This at least applies for one of the two sub-patterns which have been introduced above, more precisely for the more expressive variant of phrasal compounds (e.g. *Guck-ich-hab-es-wieder-aus-dem-Bett-geschafft-Rap*). For sure, the intention of being creative is less apparent in formations like *1000-Stunden-Jahr*.

To conclude this excursion to creativity, it has to be highlighted that phrasal compounds – regardless of whether you label them E-creative or F-creative, display a “balance of originality and effectiveness”, i.e. their creativity does not jeopardize “effortless mutual understanding” between speaker and hearer (Schmid 2025: 98). Among others, their use in advertising underlines that phrasal compounds are a very effective way of communicating.

3. Phrasal compounds and the lexicon

What do phrasal compounds tell us about the lexicon? And what implications do they bring for more general questions like “What is a lexeme?” or “How lexical is morphology?” (Gaeta 2016)?

² For a critical discussion of the dichotomy between E-creativity and F-creativity and extensions of Sampson’s (2016) distinction cf. Stumpf (under review).

The theoretical ‘attractiveness’ which I have sketched out in section 2.1 will be used below to discuss how an adequate conception of the lexicon has to look like from the perspective of phrasal compounding. For that purpose, I will build on the approach developed by Gaeta (2016). This approach has emerged in a broader context, with the aim to clarify “how the lexicon has to be concretely conceived” (Gaeta 2016: 110). Gaeta states that the relation between morphology and lexicon has received relatively little attention to date. In what follows, he argues for a “consistently lexical morphology” (Gaeta 2016: 110). In this context, among other phenomena, he also explicitly refers to phrasal compounding.

3.1 Gaeta’s “Lexicality of the Input” principle

Crucial for Gaeta’s (2016) approach is that it combines two traditionally competing views by assuming an interplay of two different conceptions of the lexicon. This idea is captured in the so-called „Lexicality of the Input Principle” (LIP): On the one hand, the term lexicon has to be understood in a Bloomfieldian sense. Bloomfield defines the lexicon (= Lex₁) as a collection of any sort of entrenched or idiomatic expressions. Consequently, the lexicon in a Bloomfieldian sense „consists of a list of form/meaning arbitrary pairs coming from different procedures and sources as it contains all expressions presenting idiomatic traits, including those larger than one word“ (Gaeta 2016: 116). On the other hand, the term lexicon denotes an infinite set of potential lexemes for any given language (= Lex₂). In this understanding, the lexicon consists of the (infinite) set of all potential lexemes in a given language and therefore results exclusively from lexeme-formation procedures“ (Gaeta 2016: 116).

The ‘double nature’ of the lexicon sketched above is comparable to the idea of the so-called ‘construction’ in Construction Grammar (e.g. Goldberg 2006), where the strict line between lexicon and syntax is blurred and Lex₁ lexemes and Lex₂ lexemes (in the sense of Gaeta 2016) are equally part of the Constructicon, the inventory of constructions.

A consequence of combining two senses of the term ‘lexicon’ is that Gaeta’s approach “accounts for complex words formed on the basis of possible but unattested words (i.e. pure Lex₂-Lexemes)” and at the same time “is not biased against complex units being stabilized in the Lex₁“ (Gaeta 2016: 131–132).

Gaeta (2016: 134) lists different phenomena which can be explained by the interaction between Lex₁ and Lex₂ mentioned above: „Even if lexicalization might at first sight appear as an inertial force [...] we will see [...] that its interaction with the Lex₂ provides us the key to understanding a number of phenomena which have been traditionally considered marginal [...]”. One of these phenomena is phrasal compounding.

In this paper, I will argue that Gaeta’s observation clearly holds for phrasal compounds whose non-head can be analyzed as a fixed multi-word unit, but is less clear for non-heads which seem to be formed ‘on the fly’, i.e. which are not part of the Lex₁.³ In what follows, this hypothesis is discussed on the basis of corpus data from written and spoken German.

3.2 The ‘lexical status’ of phrasal modifiers from a corpus-based perspective

As several corpus-based investigations clearly indicate (Hein & Antonioli 2025; Hein 2023, 2015), the modifier of a phrasal compound is not necessarily a fixed multi-word unit. Rather, the modifiers vary in terms of their fixedness:

³ To my mind, modifiers formed ‘on the fly’ are not part of the Lex₂ either because they are syntactical units.

- (1) Fixed multi-word expression
Weniger-ist-mehr-Motto
‘Less-is-more credo’
- (2) Multi-word expression formed ‘on the fly’
Okay-wir-sind-nur-ne-Woche-da-Abschnitt
‘Okay-we’re-only-there-for-a-week section’
- (3) Completely lexicalized phrasal compounds (with own dictionary entry)
Armeleuteessen
‘Poor-people meal’

On the one hand, the modifier can be formed by a fixed multi-word expression (i), often in the form of a proverb or a slogan (cf. Steyer & Hein 2018), e.g. *Weniger-ist-mehr-Motto* ‘less-is-more credo’. While cases like that are clearly covered by Gaeta’s LIP (cf. section 3.1), this does not apply to cases like (ii) *Okay-wir-sind-nur-ne-Woche-da-Abschnitt* ‘Okay-we’re-only-there-for-a-week section’, an example from spoken German (cf. Hein & Antonioli 2025) where the modifier quite obviously is formed ‘on the fly’. The third case (iii) that has to be mentioned in this context is illustrated by phrasal compounds such as *Armeleuteessen*. The latter is lexicalized from a synchronic point of view and has a separate lexical entry (<https://www.duden.de/rechtschreibung/Armeleuteessen>). Consequently, formations of that type are not meaningful for the question of what phrasal compounds tell us about the lexicon. From the perspective of the lexicon, they have the same status as completely lexicalized N+N compounds such as *Haustür*, even if internally they display a marked structure.

Considering cases like *Okay-wir-sind-nur-ne-Woche-da-Abschnitt*, my hypothesis is that the inventory of phrasal compounds can only be captured in part through the conception of the lexicon outlined by Gaeta (cf. section 3.1).

The systematic use of corpus data allows us to have a look at the quantitative dimension of the lexical status of phrasal non-heads (cf. Table 1) – instead of claiming without any empirical evidence that phrasal compounds only occur with lexicalized non-heads (cf. Hein 2011: 332–334) or categorically assigning the theoretical status of a quotation to phrasal non-heads (cf. Wiese 1996; and more recently Gutzmann & Turgay 2024).⁴ See also Goldberg & Shirtz, who emphasize that the non-heads “very often have grammatical features associated with quotes”, while not all PALs are even hypothetical quotes” (Goldberg & Shirtz 2025: 298–299). A look at corpus data from written German (DeReKo) (cf. Hein 2015) and spoken German (FOLK corpus) (cf. Hein & Antonioli 2025)⁵ reveals that formations with lexicalized modifier are the most frequent subtype of phrasal compounding in both modalities. Lexicalized non-heads occur with a proportion of ~ 51 % respectively 64 %, e.g. *Leck-mich-am-Arsch-Tag* ‘Kiss-my-ass day’.

⁴ In my opinion, it is insufficient and unsatisfactory to regard the syntactic units in phrasal compounds categorically as quotations and not to examine their degree of fixedness in greater empirical depth. Nonetheless, it is true that quotes play a role in phrasal compounding, for example for the realisation of stereotypes described in section 3.3).

⁵ We considered 1.576 types from written German and 269 types from spoken German. Even if those inventories differ clearly in size (as well as the underlying corpora DeReKo and FOLK, cf. Hein & Antonioli 2025 : 71), the results obtained from the qualitative analysis of the two datasets can be compared and related to one another.

Table 1. Lexical status of phrasal modifiers: DeReKo vs. FOLK (cf. Hein & Antonioli 2025: 73)

Lexical status of the modifier	Examples	Proportion in FOLK (%)	Proportion in DeReKo (%)
Fixed	<i>Leck-mich-am-Arsch-Tag</i>	51,3	63,83
Built ‘on the fly’	<i>Wir-spielen-Spiele-Tag</i>	18,6	35,91
Lexicalized phrasal compound	<i>Acht-Stunden-Tag</i>	29,73	0,25

However, it has to be highlighted that non-heads formed ‘on the fly’ like in *Wir-spielen-Spiele-Tag* ‘We-play-games day’ occur with a non-marginal proportion of ~ 19 % (spoken German) respectively of ~ 36 % (written German). The fact that such non-heads are more frequent in written than in spoken German might be due to the fact that their production as well as their understanding involves more effort, which probably makes them more difficult to use in direct interaction.

Lastly, also the aforementioned type of lexicalized phrasal compound is attested in the data, e.g. *Acht-Stunden-Tag*. Interestingly, the proportion of such formations is particularly high in spoken German (~ 30 %), while it is particularly low (~ 0,3 %) in written German. This finding is intuitively expectable: It can be assumed that phrasal compounds which have a separate lexical entry can be produced (and understood) more easily and faster compared to their non-lexicalized counterparts. This for sure favours the use of completely lexicalized phrasal compounds in spoken language (cf. Hein & Antonioli 2025: 73–74).

If one takes empirical data seriously, one must conclude that phrasal compounds with a non-lexicalized modifier are not marginal at all. Even if the operationalisation of a concept such as lexicalization is not trivial⁶, determining the lexical status of the modifier is an important part of the empirical description of phrasal compounding. To my mind, a proportion of ~19 % respectively ~36 % of modifiers built on the fly cannot be ignored in the debate about the theoretical implications of phrasal compounding. Thus, the question arises how we can account for this phenomenon from the point of view of the lexicon.

The hypothesis that I will push forward is that the LIP (Gaeta 2016) might be too strict to explain the entire range of phrasal compounds. Moreover, the situation is even more complex, as there are two kinds of ‘special cases’ that also have to be discussed in the context of lexicalization, namely the role of abstract patterns and stereotypes for the coinage of phrasal non-heads.

3.3 Abstract patterns and stereotypes as blueprints for phrasal modifiers

From the perspective of multi-word units, some modifiers of phrasal compounds can also be classified as instantiations of more abstract underlying linguistic patterns. An example of such pattern is [X *oder* ‘or’ Y] (cf. Steyer 2013):

⁶ The following criteria were used to classify the lexical status of the modifier : (1) ‘fixed’: the modifier is listed in the dictionary (www.duden.de) or occurs recurrently in DeReKo ; (2) ‘built on the fly’: the modifier is not listed in the dictionary and does not occur recurrently in DeReKo ; (3) ‘lexicalized phrasal compound’: the entire compound is listed in the dictionary with an own entry (cf. Hein 2015: 205-209).

[X oder ‘or’ Y]*Zu-mir-oder-zu-dir-Gequatsche* ‘my-place-or-yours chatter’*Alles oder nichts-Devise* ‘all-or-nothing motto’*Friß-oder-stirb-Rede* ‘eat-or-die speech’

The idea illustrated here is that phrasal non-heads like *Zu-mir-oder-zu-dir*, *Alles oder nichts* or *Friß-oder-stirb* can be seen as pair formulas leading back to the more abstract pattern [X oder Y] (cf. Steyer & Hein 2018). At the same time, they are fixed multi-word units in their own right, and thus are supposed to be captured by the Lex₁ component in Gaeta's model (cf. section 3.1). The status of [X oder Y] as pattern, in turn, is externally motivated, i.e. it does not depend on the frequency of corresponding non-heads in phrasal compounds.

The specific quality of modifiers such as *Zu-mir-oder-zu-dir* (and many others) results from their double nature: On the one hand, they are lexicalized multi-word expressions, while on the other hand they are concrete instantiations of more abstract underlying linguistic patterns. Again, the idea of a Constructicon which allows for the representation of both (semi)abstract patterns and concrete lexical structures and also captures the relation between different kinds of linguistic information seems to be most suitable to account for this property of phrasal non-heads.

As a second type of ‘special case’, I would like to outline the role of stereotypes for phrasal compounding. There are many non-heads which are characterized by the use of stereotypes, e.g. *“Ich-kann-Golf-Ski-und-Wandern-und-bin-schöner-als-die-andern”-Franz* ‘I-can-golf-ski-and-hike-and-I-am-more-beautiful-than-the-others Franz’. By using a fictional quotation⁷, i.e. a potential utterance, this phrasal compound makes use of the cliché of Franz Beckenbauer being arrogant.⁸

In general, stereotypes play an important role in phrasal compounding. They are used to characterize a person, an attitude, etc. They can be easily realized by using „original or fictional quotation“ (Finkbeiner & Meibauer 2016: 39; cf. also Pascual, Królak & Janssen 2013) and are very effective:

Quotes can be powerful. They provide an effective way of identifying or characterizing the content of political movements (e.g. *fight racism*), familiar situations (e.g. *it's not you, it's me*), attitudes (e.g. *life is good*), or types of people (e.g. *don't mess with me*). The fact that a well-chosen quote can evoke so much is why quotes are used as bumper stickers, worn on T-shirts, included in email. (Goldberg & Shirtz 2025: 299)

But quotations are not the only way to express stereotypes. This is illustrated in *Dumme-Mädchen-Grill* ‘stupid-girls grill’, an example from spoken German (cf. Hein & Antonioli 2025: 75-76). Here, the stereotype of girls being stupid is played with without the use of a direct quote.

Complex words like *“Ich-kann-Golf-Ski-und-Wandern-und-bin-schöner-als-die-andern”-Franz* are not understandable without the help of world knowledge. Those who do not know Franz Beckenbauer, the famous football player and coach, will not understand the meaning of this complex word. Such extra-linguistic knowledge is not necessarily part of the lexicon. In

⁷ The fact that stereotypes in phrasal compounds are sometimes realized through quotations does not mean, in my view, that the modifier always has the status of quotations (cf. section 3.2).

⁸ Interestingly, in the data of Hein (2015) there are several phrasal compounds with the head *Franz*, all referring to Franz Beckenbauer (<http://uwv.ids-mannheim.de/phraskomp>).

this sense, Goldberg & Shiritz (2025: 292) argue that “speakers who produce a PAL construction are perceived to share more common knowledge with their intended audience than speakers who produce a close paraphrase without a PAL”.

Turning back to our initial question about the nature of the lexicon, I would like to emphasize that *Ich-kann-Golf-Ski-und-Wandern-und-bin-schöner-als-die-andern* is ‘fixed’ with regard to the underlying conceptual pattern, but it is not a typical multi-word unit. Thus, it is not the kind of lexicalized structure which is supposed to be part of the Lex₁ in Gaeta’s (2016) model of the lexicon.

All in all, this section has illustrated that not only lexicalization, but also (semi)abstract patterns and non-linguistic concepts such as stereotypes play a role for phrasal compounding. When addressing the question of what phrasal compounds tell us about the lexicon, such blueprints must be taken into account.

4. Conclusion

In this article, I tried to work out the consequences that phrasal compounds stipulate for an adequate conception of the lexicon.

Phrasal compounding is a productive process of German word-formation and by no means irregular or marginal. Aside from the fact that the empirical investigation of the phenomenon in various text types – newspaper texts (Hein 2015), spoken German (Hein & Antonioli 2025) and song lyrics (Hein 2023) – is very fruitful and brings to light many highly creative formations, phrasal compounding is also particularly suitable for theoretical questions due to its status as an interface phenomenon between morphology and syntax (cf. section 2). This theoretical potential is explored in this article.

In order to show which conception of the lexicon is needed to deal with phrasal compounds, I took Gaeta’s approach (2016) as a starting point (cf. section 3.1). In a next step, I tried to demonstrate on the basis of several corpus-based studies that the LIP is too strict to explain the whole range of phrasal compounds. The reason is that non-lexicalized modifiers are neither part of the Lex₁ nor part of the Lex₂ in the sense of Gaeta (2016) (cf. section 3.1)

Given the remarkable number of modifiers which are built ‘on the fly’, phrasal compounding cannot be fully explained by assuming an interplay of two different conceptions of the lexicon (cf. sections 3.2 and 3.3).⁹ Rather, the formation of phrasal compounds requires, at least in certain cases, an interaction between morphology and syntax. To my mind, phrasal compounds with modifiers built ‘on the fly’ are the result of integrating a syntactically formed unit into a product of word-formation. Consequently, the phenomenon does not provide clear evidence of a “consistently lexical morphology” (Gaeta 2016: 110) (cf. section 3).

Then the question arises how this interaction between morphology and syntax can be embedded in a theoretical framework (cf. Hein 2011). A solution might be a construction grammar approach in which the strict line between morphology and syntax is blurred and where all kind of grammatical knowledge is stored in a so-called Constructicon – whether they are (semi)abstract patterns or concrete linguistic structures at word-level or above it. I already demonstrated that phrasal compounding can be fruitfully modelled within the framework of Construction Grammar, considering a large empirically gained inventory of 1.576 phrasal compounds from the German Reference Corpus (DeReKo) (cf. Hein 2015). Also Gaeta (2016:

⁹ This is unless one assumes that ‘intended creativity’ does not fall within the scope of word-formation patterns and consequently does not need to be covered by them either (cf. Schultink 1961) – a point of view that I do not find satisfactory, given the high proportion of freely formed modifiers (cf. section 3.2).

109) claims that a phenomenon like phrasal compounding can only be explained “if the multifaceted perspective of the Constructicon is adopted as the interface of different modules of grammar”.

All in all, like many other phenomena, phrasal compounds underline the need for a dynamic, “richly textured” concept of the lexicon which is “not just an unstructured list of exceptions” (Jackendoff & Audring 2020: 83), but that contains abstract patterns or schemata in addition to concrete lexical realizations (cf. also Bauer 2017). Moreover, the crucial role of patterns is confirmed: „Patterns are the key to why speakers understand each other in everyday communication in spite of ever-changing lexical material, syntactic variance and strong idiosyncrasies“ (Steyer 2020: 275; cf. Schmid 2014).

With regard to the general role of phrases in word-formation, further investigation is definitely needed. For example, neither adjectival phrasal compounds nor other types of word-formation which involve syntactical units (cf. Lawrenz 2006) have been researched with empirical depth. Since such phenomena provide important insights into the nature of the lexicon and its relation to morphology and syntax, their further investigation is an important issue for future work.

References

- Aronoff, Mark. 1994. *Morphology by itself*. Cambridge: MIT Press.
- Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge: Cambridge University Press.
- Booij, Geert E. 2010. *Construction morphology*. Oxford: Oxford University Press.
- Eitelmann, Matthias & Dagmar Haumann. (eds.) 2022a. *Extravagant morphology*. Amsterdam/Philadelphia: John Benjamins.
- Eitelmann, Matthias & Dagmar Haumann. 2022b. Extravagance in morphology. Introduction. In Matthias Eitelmann & Dagmar Haumann (eds.), *Extravagant morphology. Studies in rule-bending, pattern-extending and theory-challenging morphology*, 1-18. Amsterdam/Philadelphia: John Benjamins.
- Finkbeiner, Rita & Jörg Meibauer. 2016. Boris “Ich bin drin“ Becker’ (‘Boris “I am in” Becker’). Syntax, semantics and pragmatics of a special naming construction. *Lingua* 181. 36-57.
- Gaeta, Livio. 2016. How lexical is morphology? The constructicon and the quadripartite architecture of grammar. In Pavel Štekauer & Livia Körtevélyessy (eds.), *Word-Formation across Languages*, 109–146. Cambridge Scholars Publishing.
- Goldberg, Adele E. 2006. *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, Adele E. & Shahrir Shitka. 2025. The English Phrase-As-Lemma Construction: When a phrase masquerades as a word, people play along. *Language* 101. 291-320.
- Günther, Christine, Sven Kotowski & Ingo Plag. 2020. Phrasal compounds can have adjectival heads. Evidence from English. *English Language & Linguistics* 24. 75-95.
- Gutzmann, Daniel & Katharina Turgay. 2024. Phrasal compounds are quotational compounds. In Anke Himmelreich, Daniel Hole & Johannes Mursell (eds.), *To the left, to the right, and much in between. A Festschrift for Katharina Hartmann*, 319–331. Frankfurt: Goethe University Frankfurt.
- Haspelmath, Martin. 1999. Why is grammaticalization irreversible? *Linguistics* 37(6). 1043-1068.
- Hein, Katrin. 2011. Phrasenkomposita - ein wortbildungs-fremdes Randphänomen zwischen Morphologie und Syntax? *Deutsche Sprache. Zeitschrift für Theorie, Praxis, Dokumentation* 39(4). 331-361.
- Hein, Katrin. 2015. *Phrasenkomposita im Deutschen. Empirische Untersuchung und konstruktionsgrammatische Modellierung*. Tübingen: Narr.
- Hein, Katrin. 2017. Modeling the properties of German phrasal compounds within a usage-based constructional approach. In Carola Trips und Jaklin Kornfilt (eds.), *Further investigations into the nature of phrasal compounding*, 119-148. Berlin: Language Science Press.
- Hein, Katrin. 2023. Beinahe-ums-Leben-kommen-in-Regenpfützen und Chauvi-Macho-Macker-Stuss – kreative Wortbildungen in Songtexten. *Journal for Language Technology and Computational*

- Linguistics* 36(1). Special Issue: *Challenges in Computational Linguistics, Empiric Research & Multidisciplinary Potential of German Song Lyrics*. 73-92.
- Hein, Katrin & Giorgio Antonioli. 2025. Phrasenkomposita im gesprochenen Deutsch an der Schnittstelle von Wortbildungs- und Gesprächsforschung. In Adriano Murelli (ed.), *Gesprochenes Deutsch heute. Variation, Wandel, Sprachkontakt.*, 65–90. Berlin & Boston: De Gruyter.
- Jackendoff, Ray & Jenny Audring. 2020. *The texture of the lexicon. Relational Morphology and the Parallel Architecture*. Oxford: Oxford University Press.
- Lawrenz, Birgit. 2006. *Moderne deutsche Wortbildung. Phrasale Wortbildung im Deutschen: linguistische Untersuchung und sprachdidaktische Behandlung*. Hamburg: Dr. Kovač.
- Lieber, Rochelle & Sergio Scalise. 2006. The Lexical Integrity Hypothesis in a new theoretical universe. *Lingue e linguaggio* 1. 7–32.
- Meibauer, Jörg. 2003. Phrasenkomposita zwischen Wortsyntax und Lexikon. *Zeitschrift für Sprachwissenschaft* 22, 153–188.
- Meibauer, Jörg. 2007. How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction. *Morphology* 17(2). 233–259. doi:10.1007/s11525-008-9118-1.
- Meibauer, Jörg. 2015. On “R” in phrasal compounds – a contextualist approach. *Language typology and universals* 68(3). Special issue: *Phrasal compounds from a typological and theoretical perspective*. 241–261.
- Pascual, Esther, Emilia Królak & Theo Janssen. Direct speech compounds. Evoking socio-cultural scenarios through fictive interaction. *Cognitive Linguistics* 24(2). 345-366.
- Radimský, Jan. 2015. *Noun+Noun Compounds in Italian*. České Budějovice: University of South Bohemia in České Budějovice.
- Sampson, Geoffrey. 2016. Two ideas about creativity. Martin Hinton (ed.), *Evidence. Experiment and argument in linguistics and philosophy of language*, 15–26. Bern: Peter Lang.
- Schlückner, Barbara. 2012. Die deutsche Kompositionsfreudigkeit. Übersicht und Einführung. In Livio Gaeta & Barbara Schlückner (eds.), *Das Deutsche als kompositionsfreudige Sprache. Strukturelle Eigenschaften und systembezogene Aspekte*, 1–25. Berlin: de Gruyter.
- Schmid, Hans-Jörg. 2025. What is creative to whom and why? Creativity in word-formation and phraseology against the backdrop of shared Conventions and individual routines. In Sabine Arndt-Lappe & Natalia Filatkina (eds.): *Dynamics at the lexicon-syntax interface. Creativity and routine in word-formation and multi-word expressions*, 97–130. Berlin & Boston: de Gruyter.
- Schultink, Henk. 1961. Produktiviteit als morfologisch fenomeen. *Forum der Letteren* 2. 100–125.
- Spencer, Andrew (1991): *Morphological theory. An introduction to word structure in generative grammar*. Oxford: Blackwell.
- Steyer, Kathrin. 2013. *Usuelle Wortverbindungen. Zentrale Muster des Sprachgebrauchs aus korpusanalytischer Sicht*. Tübingen: Narr.
- Steyer, Kathrin & Katrin Hein. 2018. Satzwertige usuelle Wortverbindungen und gebrauchsbasierte Muster. In Stefan Engelberg, Henning Lobin, Kathrin Steyer & Sascha Wolfer: *Wortschätze. Dynamik, Muster, Komplexität*, 107-130. Berlin: de Gruyter.
- Stumpf, Sören. under review. Analogy, creativity, and Diachronic Construction Morphology: Tracing the emergence of a semi-schematic compound in German. In Carmen Mellado Blanco, Pedro Ivorra Ordines & Maricel Esteban Fonollosa (eds.), *Creativity in the Intersection between Phraseology and Construction Grammar*. Amsterdam & Philadelphia: Benjamins.
- Trips, Carola & Jaklin Kornfilt. (eds.) 2017. *Further investigations into the nature of phrasal compounding*. Berlin: Language Science Press.
- Wiese, Richard. 1996. Phrasal compounds and the theory of word syntax. *Linguistic Inquiry* 27 (1). 183–193.