

Alexander Ziem* and Tim Feldmüller

Dimensions of constructional meanings in the German Constructicon: Why collo-profiles matter

<https://doi.org/10.1515/gcla-2023-0010>

Abstract: *A constructicon, i.e., a structured inventory of constructions, essentially aims at documenting functions of lexical and grammatical constructions. Among other parameters, so-called constructional collo-profiles, as introduced by Herbst (2018, 2020), are conclusive for determining constructional meanings. They provide information on how relevant individual words are for construction slots, they hint at usage preferences of constructions and serve as a helpful indicator for semantic peculiarities of constructions. However, even though collo-profiles constitute an indispensable component of constructicon entries, they pose major challengers for constructicographers: For a constructicographic enterprise it is not feasible to conduct collostructional analyses for hundreds or even thousands of constructions. In this article, we introduce a procedure based on the large language model BERT that allows to predict collo-profiles without having to extensively annotate instances of constructions in a given corpus. Specifically, by discussing the constructions X macht Y ADJP ('x makes Y ADJ', e.g. he drives him crazy) and N₁ PREP N₁ (e.g., bumper to bumper, constructions over constructions), we show how the developed automated system generates collo-profiles based on a limited number of annotated instances. Finally, we place collo-profiles alongside other dimensions of constructional meanings included in the German Constructicon.*

Keywords: constructional meaning; reduplication construction; collo-profile; constructicography; Large Language Models; BERT

* **Corresponding author: Alexander Ziem**, Institut für Germanistik, 24.54, U1.83, Universitätsstr. 1, 40221 Düsseldorf, Germany, ziem@hhu.de
Tim Feldmüller, Leibniz-Institut für Deutsche Sprache (IDS), R5 6-13, 68161 Mannheim, Germany, feldmueller@ids-mannheim.de

1 Constructicography and the meanings of constructions

Constructicography is still a very young linguistic subdiscipline which is dedicated to the systematic documentation and ‘archiving’ of lexical as well as grammatical constructions (cf. Boas, Lyngfelt and Torrent 2019; Lyngfelt 2018; Ziem and Flick 2019). Following a constructionist framework, constructicography ultimately aims at developing (digital) ‘dictionaries’ for constructions. Such reference works can serve a variety of purposes; they may, for example, serve as a linguistic documentation of a target language at a specific point in time, or as a rich resource for linguistic studies, including computational endeavors such as building construction parsers, or as pedagogical tools for foreign language teaching and learning.

Constructicography builds on the long traditions of lexicography and grammaticography, both of which have a great deal of (theoretical and practical-empirical) experience and bring with them a wealth of established concepts and tools (cf. Fuß and Wöllstein 2018, Mosel 2006, Osswald 2015). Just like lexicography, constructicography relies on corpus data which provide insights into the linguistic system peculiar to a variety of a speech community. At the same time, it differs from lexicography and grammaticography in that it is guided by the assumption that both lexicon and grammar form poles of a continuum (for an overview cf. Boas 2010, Broccias 2012, Ziem 2022). In this view, grammar and lexicon are not supposed to be different in nature. Instead, grammatical and lexical units are similar in kind and should thus be analyzed within the same framework (cf. in detail Ziem 2022).

Importantly, constructicographers are specifically interested in constructions that fall between lexicon and grammar: semi-schematic constructions, that is, constructions with both open and lexically specified slots (cf. for Russian: Janda et al. 2020, for Swedish: Bäckström et al. 2013, for German Boas and Ziem 2018). For illustration, consider the N_1 an N_1 construction in (1) which we will discuss in more detail later.

- (1) Die Autos stehen **Stoßstange an Stoßstange** auf der Autobahn.
 ‘The cars are standing **bumper to bumper** on the highway’

This construction includes a lexically specified slot (*an* ‘to’) and two open slots. Its function is to quantify the entity addressed in the noun slot (‘bumpers’) and it also expresses continuity (‘bumper to bumper to bumper...’).

Hitherto, neither grammars nor dictionaries feel responsible for such semi-schematic constructions, including constructional idioms.¹ Usually, they are too schematic for a dictionary and too word-like for a grammar. Consequently, semi-schematic constructions fall through the cracks of a lexicon-and-grammar approach. They cannot be integrated into a lexicon; at the same time, they share only some properties with the grammatical units addressed in grammars. Although semi-schematic constructions are commonly considered ‘irregular’, their slots are subject to locally defined regularities, in particular constraints, which determine in a construction-specific way which semantic, pragmatic or formal conditions slot-fillers have to meet (cf., for instance, Fillmore, Kay and O’Connor’s (1988) seminal paper on pragmatic constraints of the *let alone* construction). Despite these peculiarities, semi-schematic constructions belong to the language system just as fully regular expressions and grammatical structures. This is all the more true since they occur in a large variety.

For any constructicographic enterprise it is of pivotal importance to document semi-schematic constructions. But how to capture their meanings? Is it possible to adapt lexicographic methods to also account for constructional meanings? If this is not the case, or only to a limited extent, is there a need to develop new methods to account for meanings of semi-schematic and fully schematic constructions? More generally, which dimensions of meaning could, and should, be addressed in a constructicon?

In this article, we try to provide at least preliminary answers to these questions. Relating to categories implemented in the German FrameNet Constructicon, we first give an overview of the broad range of semantic dimensions included in each construction entry. We argue that due to the conceptual complexity of semi-schematic constructions we need go beyond definition-like meaning description as usually provided in lexicon entries. Instead, it is necessary to offer various points of access to constructional meanings, ranging from semantic specifications of a construction’s slots to information about the construction family of which the target construction is a member and about so-called constructional collo-profiles specific to the construction addressed. In the second part of our paper, we solely focus on the role of collo-profiles as introduced by Herbst (2018, 2020) and Herbst and Hoffmann (2018), showing that collo-profiles are not only

¹ However, for an exception cf., for example, the *Oxford Advanced Learner’s Dictionary* (2020) which does not mention the schematic N_1 PREP N_1 construction but some daughters.

conclusive for determining constructional meanings but also helpful for explicating usage preferences, and thus serve as a rich semantic resource. As a consequence, we consider collo-profiles an indispensable component of construction entries. However, in the context of constructicography it is not feasible to conduct collostructional analyses (CA) for hundreds or even thousands of constructions. Being aware of this limitation, we introduce a procedure based on the large language model (LLM) BERT that allows to predict collo-profiles without having to extensively annotate constructs, i.e. instances of constructions, in a given corpus.

2 Constructicographic approaches to constructional meaning in the German Constructicon

In the last decade, there has been a growing interest in developing and implementing rich language resources based on constructionist assumptions (Lyngfelt et al. 2018). In the spirit of construction grammar as well as (computational) lexicography, the German FrameNet Constructicon project (cf. www.german-constructicon.de), among others, specifically tries to explore and document the continuum between lexicon and grammar on a large scale.

From a constructicographic point of view, a constructicon – a morphological blend of *construction* and *lexicon* coined by Jurafsky (1992: 302) – is the result of examining and documenting both the full range of lexical and grammatical constructions of a target language (including meaning-bearing units of all kinds) and the ways in which they are interconnected and form a complex network of relations. It is the entire network of constructions that makes up the *constructicon* of a language. However, the term *constructicon* is polysemous as it does not only refer to a lexicon-like resource of constructions but also to the linguistic knowledge of speakers in a language community. With respect to such a “mental constructicon”, Goldberg (2003: 219) goes so far as to claim that the totality of our knowledge of language must be captured by a network of constructions. In the remainder, we will use *constructicon* exclusively in the former sense.

Constructicographic analyses of constructions are very demanding and labor-intensive. Since, at least in the case of German as a target language, partially schematic constructions have so far received little attention in grammaticography (Dobrovolskij 2011, Ziem 2018, Finkbeiner 2018), we restricted ourselves to this type of construction in the first phase of the project. The guiding principle was Fillmore’s (1988: 36) claim that it is precisely the ‘irregular’, ‘idiosyncratic’, and

supposedly ‘marginal’ constructions that must be used to measure the viability of an approach. In this area between lexicon and grammar, we find many hundreds of idiosyncratic, semi-schematic, and (partially) idiomatic constructions. They include, for instance, grammatical phrasemes (*let alone, for...sake*), phrasal idioms (cf. Ziem 2018, 2020), support verb constructions (*making a request, making an assertion*, etc.) as well as a multitude of other units, such as headless constructions (e.g., *Was mich ärgert, sind solche Sätze* ‘What annoys me are such sentences’), prepositional phrases with bare nouns (*mit Sorge* ‘with concern’, *in Trauer* ‘in sorrow’), and many more. The range of constructional idioms alone includes examples as diverse as those documented in (2).

- | | |
|--------------------------------|--|
| (2) a. {Seite} an {Seite} | ‘side by side’ |
| b. {Autos} über {Autos} | ‘cars over cars’ |
| c. {Meister} der {Meister} | ‘master of masters’ |
| d. {Ich} und {Golf spielen}? | ‘Me and play golf?’ |
| e. Was {du} immer {meckerst}! | ‘What you are always complaining about!’ |
| f. Es ist {zum Mäusemelken}. | ‘It’s enough to drive up the wall’ |
| g. Wie dem auch sei {...} | ‘Be that as it may’ |
| h. Weg mit {...} | ‘Away with’ |
| i. {...} geschweige denn {...} | ‘let alone’ |
| j. {ein Idiot} von {Lehrer} | ‘{an idiot} of a {teacher} |
| k. Je mehr {...}, desto {...} | ‘the more {...}, the {...}’ |

Generally, semi-schematic constructions with varying productivity are particularly interesting and methodologically challenging since they do not only occur with great variety but also exhibit properties that are usually described lexicographically, while other properties, in particular those addressing construction elements (slots) coming with specific constraints, require special attention for a comprehensive description of a construction’s function. In order to capture functions of constructions as comprehensively as possible, each entry contains a range of semantic information, including the following.

(a) Meaning of construction elements (slots). Constructicographic investigations start with a careful analysis of the slots of the target construction (construction elements). To illustrate, consider for example the N_1 an N_1 construction as illustrated in (1). This construction is both in formal and functional terms a daughter of the more abstract prepositional reduplication construction. At the same time, it is a sister of the N_1 über N_1 construction as exemplified in

(2b), and it also relates to the non-schematic construction *Seite an Seite*, cf. (2a). Despite its formal similarities to this construction, it differs from it in that its construction elements are open for a wealth of fillers, namely any kind of count nouns; this is because (2a) is idiomatic even though there are functionally related non-idiomatic instances attested in corpus data. (3a) illustrates the idiomatic, (3b) the non-idiomatic one.

- (3) a. Er kämpfte Seite an Seite mit der FDP (DWDS, Kernkorpus 1900-1999)
 ‘He fought side by side with the FDP’
 b. ...hier hatte er früher Dutzende einander fremde Menschen Seite an Seite beten gesehen (DWDS, Kernkorpus 1900-1999)
 ‘...here he had seen dozens of strangers praying side by side in the past’

(1) also differs from (3a-b) in that its nominal fillers may be modified to some extent (e.g., *blue bumpers to blue bumpers*). This is because *Seite an Seite* is a non-schematic phraseological unit. Thus, it comes with no surprise that it should get its own entry in a construction, just like the semi-schematic N_1 an N_1 construction. The meanings of both constructions do not match, but they do overlap: While both units code spatial contiguity (proximity), only the semi-schematic construction codes continuity and thus quantifies the nominal referent. In addition, there is also only partial overlap between N_1 an N_1 and other preposition reduplication constructions such as (2b). Both fulfill a quantifying function, and in both cases the slots of the constructions only license count nouns; however, only (3b) codes continuity. Hence, it is this semantic dimension that makes the N_1 an N_1 construction functionally unique. Such a slot analysis is an essential starting point of any constructicographic meaning description.

(b) Relations among constructions: constructional networks. Importantly, it is not possible to get a full picture of a construction’s function by only looking at its ‘internal’ properties. Even though analyses of construction elements also rely on comparing constructions, they do not take the habitat of the target construction into account. For this, it is necessary to consider the relations holding between a target construction and its neighbors. A close glance at the direct neighbors promises new insights into the peculiarities of the target construction in several respects. To exemplify, our target construction N_1 an N_1 has one mother and three sisters, some of which have sisters again; Table 1 provides an overview.

Table 1: Local constructional network of the N_1 an N_1 construction

Construction	Example	Relation to N_1 an N_1
Reduplication_continuity_N1_PREP_N1	Autos über Autos (‘cars over cars’) Haus an Haus (‘house next to house’)	mother
Reduplication_competetion_N1_gegen_N1	Freunde gegen Freunde (‘friends versus friends’)	sister
Reduplication_Quantification_N1_über_N1	Autos über Autos (‘cars over cars’)	sister
Reduplication_Proximity_Colour_in_Colour	gelb in gelb (yellow in yellow)	sister

Each construction shares both basic form- and meaning-related characteristics: formally, they all instantiate one and the same reduplication schema provided by the fully schematic N_1 PREP N_1 construction serving as the mother construction; functionally, they all express continuity combined with some sort of quantification. Such local network structures are visualized by the Construction-Grapher (for an overview cf. Ziem and Willich in press). To illustrate, Figure 1 shows the habitat of the N_1 an N_1 construction (for the sake of better visualization restricted to two edges here).

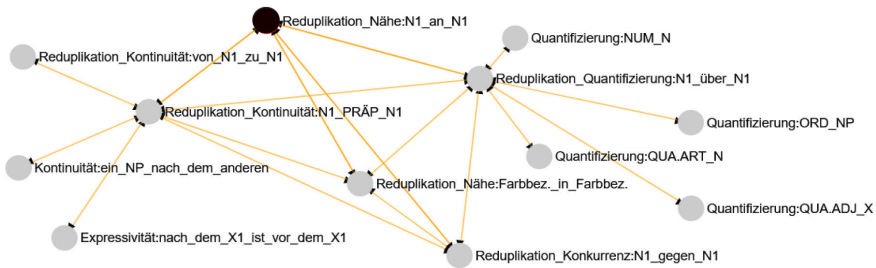


Figure 1: The N_1 an N_1 _continuity construction in its local habitat (cf. <https://gsw.phil.hhu.de/constructicon/constructiondata?id=349>)

(c) Family structures. The habitat of a construction, i.e., its local constructional network, is not only structured by (different types of) relations holding

among constructions. On top of that, they are organized in families (cf. Ziem and Willich, 2023). Such family structures provide additional clues to constructional meanings. In total, there are three types of families, each providing new insights into the nature of the target construction: (a) *Construction families* address constructions that share essential commonalities both in terms of form and meaning; in the case of N_1 an N_1 , the constructions summarized in Table 1 form members of this family. (b) *Form families* relate to structural commonalities among constructions. For instance, the N_1 an N_1 construction belongs to the large family of reduplication constructions comprising a variety of phenomena, ranging from phraseological units, such as *x hin x her* ('x or no x') to intensifying reduplication constructions, such as *Adj_Adj_NP*.² Finally, (c) *meaning families* gather constructions with strong semantic similarities regardless of their form gestalts. Constructions belonging to the same meaning family evoke the same frame or at least closely related frames. Specifically, in the case of the N_1 an N_1 construction it is the frame *Gradierbare_Nähe* ('Gradable_proximity') that shapes the meaning of the construction under investigation. Interestingly, the local network structure provided by the FrameGrapher opens up a broader, more coarse-grained view on the meaning of the N_1 an N_1 construction in that the evoked frame is part of a large group of frames coding spatial relations between entities, including the image schema frame *Nähe* ('proximity') as well as the frame *Lokative_Beziehung* ('local relation') which is essential for any frame in the large space-coding frame family.³ To conclude, family structures allow us to take a different perspective on properties specific to the target construction in general and, more specifically, on constructional meaning. At a higher level of granularity, they capture local network structures in which constructions are organized.

(d) Frame-approach to constructional meaning. As mentioned above, constructional meaning is determined by the frame a construction evokes. One benefit for including frames is to get an accurate sense of what other meaning-bearing entities in the lexicon-grammar continuum (ranging from lexical to phraseological and grammatical units) have in common with the target construction. At the same time, looking at other units evoking the same frame also reveals

² For a detailed description of the first, as exemplified in *Traum hin, Traum her, das ist unrealistisch!* ('Dream, or no dream, this unrealistic!') cf. the construction entry in the German Constructicon: <https://gsw.phil.hhu.de/constructicon/construction?id=1351>. For the latter, as exemplified in *Sie kam vor vielen vielen Jahren* ('She came many many years ago'), cf. <https://gsw.phil.hhu.de/constructicon/construction?id=756>, accessed 28 September 2023.

For a list of constructions constituting the form family cf. <https://gsw.phil.hhu.de/constructicon/constructionfamily?id=5>.

³ A family that entails a total of 19 frames, cf. <https://gsw.phil.hhu.de/framenet/framefamily?id=32>, accessed 28 September 2023.

what distinguishes the construction from others. To this end, mappings between construction elements and elements of the evoked frame are instructive. Figure 2 illustrates these relations.

Since a frame entertains relations to other frames (and may be a member of one or more frame families), every frame evoked by a construction provides access to rich semantics, reaching far beyond the meaning provided by the frame itself. Consequently, taking the evoked frame and its habitat into consideration makes it possible to take a very sophisticated perspective on constructional meaning.

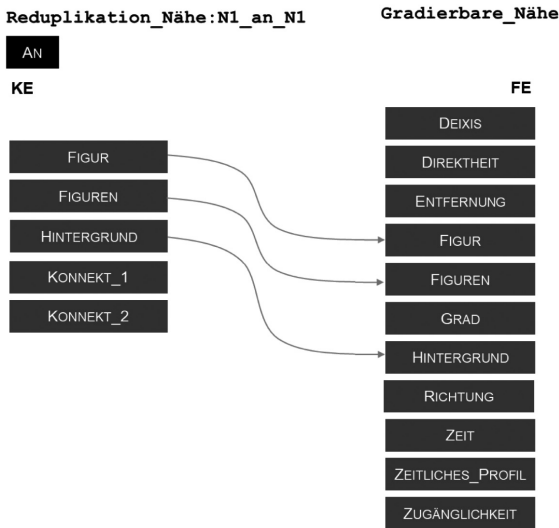


Figure 2: Mappings between construction elements and frame elements (cf. <https://gsw.phil.hhu.de/construction/constructiondata?id=349>)

(e) Collo-profiles. Before constructicography emerged about half a decade ago, the most influential method to analyze constructional meaning relates to so-called collocations (a blend of *collocation* and *construction*). In one variety, collocation analysis (CA) measures the association strength between construction elements (slots), or the entire construction to which they belong, and instances entering the construction elements (Stefanowitsch and Gries 2003). With such a focus, CA combines John R. Firth’s (1951/1957) notion of collocation and Goldberg’s (2003) notion of construction (also adopted here). Crucially, collocations help to identify collo-profiles which in turn provide insights into paradigmatic properties of construction elements by specifying the strength with which a construction element attracts linguistic units (Herbst 2018, 2020). Thus, collo-profiles are powerful clues to a better understanding of constructional meaning: if a collo-profile changes, the construction’s meaning changes.

In the next section, we stick to this dimension of constructional meaning and show in detail how to determine the collo-profile for the schematic N_1 PREP N_1 (continuity) construction with the help of the Large Language Model BERT.

3 Integrating collo-profiles in the German Constructicon

3.1 Collostructional Analysis and its challenges for constructicography

So far, we argued that collo-profiles are - among other dimensions of constructional meaning - an essential part of a construction entry. Usually, so-called count-based methods, most prominently CA (Stefanowitsch and Gries 2003) help to identify collo-profiles; they are designed for providing information on how relevant individual words are for construction slots. Meanwhile, CA has become the gold standard for obvious reasons: With the help of statistical significance tests following the logic of collocation analysis, CA does not only indicate whether and how often a word occurs in a construction slot, but also whether it shows a particular statistical attraction or repulsion to this slot.

Gries and Stefanowitsch developed three variants of CA, each addressing different sets of questions relevant to construction analyses.

- (a) First, they proposed collexeme analysis (Stefanowitsch and Gries, 2003) that allows to calculate for every slot how strongly a certain (complex) linguistic expression is associated with it, for example, *give* with the verbal slot of the ditransitive construction.

Later they developed two major extensions of CA:

- (b) The first one is distinctive collexeme analysis (Gries and Stefanowitsch 2004) that serves for determining the strength of association between a given (complex) linguistic expression and two or more functionally similar constructions, for example, between the verb *give* and the ditransitive construction in contrast to prepositional constructions with a nominal phrase in the dative case.
- (c) The second extension, covarying collexeme analysis (Stefanowitsch and Gries 2005), aims to determine mutual dependencies of (complex) linguistic expressions in a construction's slots. For the ditransitive construction, for example, this would allow to calculate how strongly the verb *send* instantiated in the verb slot is associated with the noun *letter*, i.e. a nominal phrase in the accusative.

While all of these variants of CA provide relevant information for collo-profiles, we have so far focused on collexeme analysis, aiming at identifying relevant fillers of individual construction slots. In the following, we illustrate which kind of data are needed for a collexeme analysis and explain why it is difficult to gather them in construction projects.

A collexeme analysis is based on the following information for all linguistic expressions – usually single lexemes – occurring in the construction slot in a given corpus: (a) the frequency of occurrence of the expression in the construction, (b) the frequency of occurrence of the expression outside the construction, (c) the frequency of occurrence of other expressions in the construction, and (d) the frequency of occurrence of other expressions in other constructions.

For constructicography, this poses several severe challenges. First, for constructions that are not easily searchable by corpus linguistic means, among them most prominently those semi-schematic constructions addressed here, it is a significant effort to identify all of their occurrences in the corpus. Second, although many constructions are retrievable using morphosyntactic automatic annotations, the latter require extensive manual post correction to achieve gold standard quality. For constructicographic purposes, this resulting workload alone is too heavy. Finally, even if the above-mentioned problems were solved, for example by using a small, possibly manually annotated corpus or by limiting ourselves to a selection of constructions, the lexical coverage would be low: Only those types that actually occur in the corpus and in the construction can ever be identified as collexemes.

Hence, in the context of construction development, the conflict between aiming at the broadest possible coverage and following the agenda of a usage-based approach poses considerable challenges that cannot be solved with the help of established tools and methods. For this reason, we are currently testing approaches for both automated extraction of construction candidates (Barteld and Ziem 2020) and for predicting collo-profiles by means of the Large Language Model (LLM) BERT (Devlin et al. 2019). In the remainder, we present the latter in some detail (for an in-depth discussion cf. Feldmüller et al. (in prep.)).

3.2 Using Large Language Models for generating collo-profiles

In the last decade, computational approaches to language have undergone major developments. In particular, the resource-efficient possibility of representing a wide range of linguistic units for very large corpora as so-called

embeddings represents a milestone in this field. This was made possible by the use of neural networks which gained great popularity with Mikolov et al.'s (2013) word2vec algorithm. Such embeddings are high-dimensional representations of words, learned via their distribution in large corpora, and capture semantic as well as morphosyntactic properties of these words. With the advent of the LLMs of the transformer generation such as BERT (Devlin et al. 2019), it has become possible to compute not only type-based but also token-based embeddings for words. Unlike classical *word embeddings* that generate a single numerical representation per word type, these transformer-based models take the surrounding words into account for every token, which addresses issues like homonymy by providing context-aware representations. The typical approach to training such neural network-based language models involves masking portions of sentences and then predicting appropriate fillers for these segments using information from the unaltered portions of the sentence. Via a feedback mechanism the deviation of the prediction from the correct solution is taken into account for further predictions. By repeating this process many times, the neural network is adjusted until the prediction is as good as possible and, as a result, embeddings have been learned.

From a construction grammar perspective, these developments have so far been largely ignored, although some work suggests that LLMs also represent constructions (Tayyar Madabushi et al. 2020; Veenboer and Bloem 2023; cf. for an overview Weissweiler et al. 2023). In particular, the approach taken by Veenboer and Bloem (2023) is of pivotal interest here since it draws on a method similar to ours and tries to make predictions for collexemes of the *X waiting-to-happen* construction using the masked language model functionality of BERT. Collected example sentences, in which the original fillers of the slots are masked are used as a basis. The results are then compared with a CA conducted on the same data. While they find significant overlap between both methods, they also state that "only items that exist in the model vocabulary as a single token can be predicted" (Veenboer and Bloem 2023: 12941).

With the method we have developed, this problem can be solved by learning so-called decontextualized embeddings (Wada et al. 2022) from a large corpus beforehand. We also implement a mechanism to filter predicted collexeme candidates. To this end, we take into account POS restrictions of the construction slot and do not count substitutes that include the source word, since these are preferentially predicted. For example, among the top 10 predictions for the

word *Jahr* ‘year’ in the N1 PREP N1 construction we find the following words: *Kalenderjahr* ‘calendar year’, *Jahrs* ‘year (inflected)’, *Jahrzehnt* ‘century’, *Jahres* ‘year (inflected)’ and *Jahrgänge* ‘years / generations’ (cf. Feldmüller et al. in prep. for details). Overall, the operation of our method can be summarized as follows:

- In a first step, we collect as many example sentences for a construction as possible and annotate instances of construction elements.
- Using decontextualized embeddings and a slightly modified implementation of unsupervised lexical substitution (Wada et al., 2022), we let BERT predict possible substitutes for the fillers of the slots, which we filter with respect to their POS tag as well as overlapping word segments.
- Finally, for every sentence, the scores of the top five predictions are summed up to generate a collo-profile.

This procedure allows us to supplement our rather small sets of annotated example sentences with vast distributional information provided by BERT and our trained embeddings. By doing so, we aim at acquiring accurate predictions which go beyond the annotated slot-fillers (i.e., instances of construction elements).

3.3 Putting neural embeddings into practice

To test our procedure, we selected two constructions: the X macht Y ADJP construction as exemplified in (4) and the N1 PREP N1 construction in the continuity meaning exemplified in (5).

- (4) Dafür kann man schlecht die EU verantwortlich machen. (TIGER) X: man ‘you’, macht: machen ‘make’, Y: dafür ‘for this’, ADJP: verantwortlich ‘responsible’
 ‘You can hardly hold the EU responsible for this / You can hardly blame the EU for this’
- (5) Diese teilte ihre Bestellungen Jahr für Jahr so auf, dass alle Anbieter genügend ausgelastet waren. (Wikipedia)
 N1: Jahr ‘year’, PREP: für ‘by’
 ‘The latter divided its orders year by year so that all suppliers were sufficiently busy.’

We opted for the former construction because Fehrmann (2018) conducted comprehensive research on it using CA. As the TIGER corpus (Brants et al. 2004), which is one of the corpora employed in Fehrmann’s analysis, is publicly accessible, and, thankfully, we were able to obtain the complete analysis results and annotated example sentences from the author, the construction was an ideal case for our comparative study. In addition, we selected the second construction because it is well searchable in a corpus and in this respect, it gave us the possibility to investigate on the basis of a large corpus⁴ and with sampling experiments to what extent the quality of the predictions depends on the number of example sentences.

While we were able to take the example sentences and results for the *X macht Y ADJP* construction from Fehrmann, we used a Python script for the *N1 PREP N1* construction to extract all sentences from the German Wikipedia matching the construction’s formal pattern.⁵ For the CA, we used Flach’s R package (2021) and approximated the number of comparable constructions contained in the corpus via the frequency of phrases with a similar syntactic function, i.e. adverbials.⁶ Following Fehrmann (2018), we used the negative decadic log transformed p-value of Fisher-Yates test as a statistical measure. The results for both constructions and each of the two methods used are shown in Table 2. Note that the scores of our procedure – as in CA – are comparable only within one slot. For different slots or constructions, only the ranking is comparable.

⁴ Both for the training of the decontextualized embeddings and for finding occurrences of the *N1 PREP N1* construction, we used the 2020 Wikipedia corpus (available under <https://sfb833-a3.github.io/tueba-ddp/>). As a BERT model we used the pretrained BERT-base-german-dbmdz-cased model, available under <https://huggingface.co/dbmdz/bert-base-german-cased> and trained by MDZ Digital Library at Bavarian State Library.

⁵ We collected all sentences in which one of the prepositions relevant to the construction (*auf*, *für* or *um*) is preceded and followed by the same type. After a review of the sentences, we further filtered out cases in which the surrounding type is tagged as pronoun, determiner or punctuation or where the pattern is preceded by the word *von*, indicating a different construction.

⁶ We counted all occurrences of the Universal Dependency Relations tags *obl*, *advcl* and *advmod* in the corpus.

Table 2: Top 20 results with scores of both CA and our method, for both constructions examined here

X macht Y ADJP		N1 PREP N1 (continuity)	
Top 20 CA	Top 20 our method	Top 20 CA	Top 20 our method
deutlich 'clear' (77.48)	deutlich 'clear' (46.62)	Schritt 'step' (inf)	Schritt 'step' (3015.66)
verantwortlich 'responsible' (71.22)	klar 'clear' (35.53) 'responsible' (28.25)	Stück 'piece' (inf)	Hals 'throat' (1239.90)
aufmerksam 'alert' (58.46)	sichtbar 'visible' (30.18)	Jahr 'year' (inf)	Tag 'day' (1210)
geltend 'valid' (34.56)	verantwortlich 'responsible' (28.25)	Zug 'move' (inf)	Monat 'month' (1059.34)
abhängig 'dependend' (31.63)	spürbar 'perceptible' (25.87)	Tag 'day' (inf)	Woche 'week' (1055.18)
rückgängig 'reversed' (23.63)	bemerkbar 'noticeable' (21.49)	Maß 'measure' (inf)	Jahr 'year' (1001.59)
zunichte 'nullified' (23.63)	mitverantwortlich 'co-responsible' (20.94)	Schicht 'layer' (282.15)	Stück 'piece' (886.73)
bemerkbar 'noticeable' (20.25)	mitschuldig 'complicit' (20.90)	Auge (277.87) 'eye'	stückweise 'piecewise' (766.17)
mobil (20.10) 'mobile'	aufmerksam 'alert' (19.52)	Zahn 'tooth' (270.71)	Zug 'move' (490.75)
überflüssig 'redundant' (18.24)	zuständig 'responsible' (18.02)	Wort 'word' (225.97)	Exemplar 'exemplar' (447.17)
fit 'fit' (17.59)	geltend 'valid' (17.65)	Stein 'stone' (216.23)	abschnittsweise 'sectionwise' (397.44)
stark 'strong' (16.14)	schuld 'guilty' (17.09)	Zeile 'line' (162.63)	Abend 'evening' (354.22)
sichtbar 'visible' (16.13)	verständlich 'understandable' (14.97)	Schlag 'strike' (160.10)	zeilenweise 'linewise' (303.28)

Table 2: (continued)

X macht Y ADJP		N1 PREP N1 (continuity)	
klar 'clear' (16.09)	vertraut 'familiar' (14.36)	Woche 'week' (139.50)	Maß 'measure' (154.04)
schuldig 'guilty' (13.74)	neugierig 'curious' (14.07)	Nacht 'night' (136.41)	Wort 'word' (151.43)
streitig 'disputed' (13.50)	unmöglich 'impossible' (13.75)	Monat 'month' (112.91)	mühevoll 'laborious' (145.27)
locker 'loose' (10.64)	bekräftigt 'affirmed' (12.44)	Steinchen 'small stone' (112.56)	Stein 'stone' (143.84)
möglich 'possible' (10.41)	lustig 'fun' (12.20)	Buchstabe 'letter' (97.94)	Auge 'eye' (143.12)
ausfindig 'located' (10.12)	überflüssig 'redundant' (12.11)	Punkt 'point' (97.41)	Schicht 'layer' (128.26)
gefügig 'compliant' (10.12)	abhängig 'dependent' (10.75)	Abend 'evening' (96.38)	stufenlos 'continuously variable' (125.98)

As shown in the table, the results exhibit noticeable overlap. For the X macht Y ADJP construction, there is a 40% agreement among the top 10 results, and a 50% agreement for the N1 PREP N1 construction. Furthermore, when considering the top 20 results, the agreement rate rises to 45% and 65%, respectively. If we analyze which of the top 10 results, although not ranked equally high, were still identified as collexemes through CA, we find 70% agreement for both constructions.

It is important to note that, unlike Veenboer and Bloem (2023), we do not consider CA a gold standard in our analysis, but rather a point of reference and comparison. Consequently, the remaining 30% of the results are not necessarily false positives, as one of the strengths of our method lies in identifying suitable fillers that go beyond the lexical content of the examples used in prompting.

In addition to more detailed quantitative results, in Feldmüller et al. (in prep.), we also elaborate on linguistic regularities that we observe in the output of the LLM, which is to some extent a black box. In contrast to CA, we cannot assume that all the collexemes proposed by our procedure are actually observable fillers of the construction. This is particularly evident from the results for the N1 PREP N1 construction, which also contain noun-to-adjective derivatives (*stückweise* ‘piecewise’, *abschnittsweise* ‘sectionwise’, *zeilenweise* ‘line-wise’, *mühevoll* ‘laborious’, *stufenlos* ‘continuously variable’) incorrectly tagged as nouns in the corpus. We therefore check for all top 20 results of our procedure whether they occur in the construction in DeReKo (Kupietz et al. 2018), i.e. the largest reference corpus available for German (55 billion tokens). The result is that there is no support for the predictions *schuld* ‘guilty’ and *bekräftigt* ‘affirmed’ in the X macht Y ADJP construction. For the N1 PREP N1 construction, in addition to the incorrectly tagged adjectives mentioned above, we do not find corpus evidence for *Hals* ‘throat’ as a filler.⁷

Yet to some extent it remains unclear why our model predicts these collexemes (for a discussion cf. Feldmüller et al. in prep.). Regardless of this shortcoming, we implemented a prediction-based approach in the German construction. In the final section, we elaborate on the reasons for doing so.

3.4 Implementation in the German Constructicon

Count-based methods such as CA and prediction-based approaches such as our proposed method follow different logics and produce different results, as e.g. Fankhauser and Kupietz (2022) have shown. An essential quality of the first-mentioned approaches is, first, that one can assume that all results actually occur in the underlying corpus and respectively in language use, and second, that their collexeme status can be precisely understood through the transparent statistical methods. These are very important qualities, so that it is not our intention to present a better method *per se*. However, the advantage of our approach is, in particular, that it does not require annotated constructions for a whole corpus and yet achieves a fairly high level of agreement with the established method. It is thus much more economical with respect to the limited resources available in constructicon projects. Moreover, by suggesting fillers that do not occur in the corpus, it can act as a complement to CA.

⁷ We find one occurrence of *Hals* in the formal pattern that does, however, not realise the continuity meaning of the construction.

Considering the favorable evaluation results and the impracticality of conducting CA on a large scale, we are currently implementing the system into the German Constructicon as outlined in Table 3. Consequently, for every identified collexeme, it will be transparent which score was assigned to it and for how many sentences it was predicted as a suitable filler. By clicking on the latter value, the corresponding sentences are accessible.

Of course, our method comes with its own set of limitations. As has already been pointed out, not every predicted collexeme proves to be a good filler for the target construction. Additionally, the ranking of the results may not match the one provided by CA; it could be influenced by factors that are not desirable from a construction grammar perspective, potentially favoring hyperonyms, for example (cf. Feldmüller et al. in prep.). Furthermore, only individual words in the slots can ever be considered for predictions and output by our method. Hence, for phrasal fillers, we select the head of the phrase determined by a parser. We also anticipate problems of representativeness for constructions with a limited number of example sentences. For these reasons, we will label the collo-profiles as automatically generated and point to potential shortcomings, particularly the limited data set. However, the model and its parameters as well as the procedural steps for generating collo-profiles and the corpora they operate on are made transparent so that users of our constructicon can decide to what extent prediction-based collo-profiles are suitable and relevant for their research questions.

Table 3: Information on collo-profiles included in construction entries

N1 PREP N1 (continuity): N1	
Collexeme	Score and number of sentences
Schritt	3015.66 (4191)
Hals	1239.90 (2159)
Tag	1210 (1821)
Monat	1059.34 (1630)
Woche	1055.18 (1669)
Jahr	1001.59 (1329)
Stück	886.73 (1178)
stückweise	766.17 (1332)
Zug	490.75 (618)
Exemplar	447.17 (806)

4 Why collo-profiles matter for constructicography: conclusions and outlook

Documenting the meaning of a construction is one of the key tasks of constructicography – and it is a specifically challenging task since it is not sufficient to solely rely on lexicographic procedures. Lexicographic analyses fall short not least because fully and partially schematic constructions are semantically much more complex than lexical constructions. Therefore, the schematicity of constructions has to be taken into account at the outset. What is the function of the construction elements (slots) of a target construction? What do they contribute to its meaning? What is the relationship between the construction elements? Which instances typically occur in the slots, and to what extent are they informative for the construction's meaning? How is the target construction related to its neighbors, and how do their construction elements relate to one another? It goes without saying that the dimensions of constructional meanings addressed are manifold – nonetheless, construction entries should attempt to cover them all.

With such a goal in mind, the present paper has introduced various dimensions of constructional meanings with a specific focus on collo-profiles. Collo-profiles are considered a semantically prominent dimension of constructional meaning in that they not only provide information about usage preferences of constructions but also serve as a helpful indicator for semantic peculiarities of constructions. However, given that a construction does not consist of dozens but of many hundreds and even thousands of constructions, collostructional analyses as proposed by Stefanowitsch and Gries are simply not feasible on such a large scale. Drawing on the neural-network-based LLM BERT, the present article has introduced a way to cope with this challenge.

Specifically, we have developed an automated system that generates collo-profiles based on a limited number of annotated instances. To this end, we use an adapted version of Wada et al.'s approach and predict lexical substitutes for the original fillers of the example sentences, which we then filter and count. For illustration, we used the N_1 PREP N_1 (' N_1 PREP N_1 ') and the x macht y ADJ (' x makes y ADJ') construction. The results achieved show considerable agreement with a CA on the same data. However, as expected, there are also shortcomings; one of them concerns the handling of predicted inappropriate collexemes.

So far, our method only provides equivalent functionality to one variant of CA, namely the simple collexeme analysis. It would be very promising to extend the framework presented here and also incorporate procedures mimicking distinctive collexeme analysis and covarying collexeme analysis.

While the procedure for predicting collo-profiles has already been implemented in the German Constructicon, there are still some issues we need to address in the near future. One of them relates to developing new methods for identifying as many new constructions as possible and for subsequently creating collo-profiles for them following the procedure explained above. The aim is that each construction entry should come with an illustration of its collo-profile. For the identification of new constructions, we currently use distributional semantics and neural embeddings that measure the degree of attraction between the slots of a construction (construction elements) and those linguistic expressions that are licensed to fill them (Barteld and Ziem 2020).

Another future task concerns applications of collo-profiles in different domains. For example, we intend to use collo-profiles for detecting semantic similarities and differences between constructions in different languages. For such cross-linguistic alignments of constructions, we are developing an interface that takes into account language-specific as well as cross-linguistic requirements (Czulo et al. in press). The long-term goal here is to link constructions across constructicons in German, Brazilian Portuguese and Swedish and Japanese (for an overview see Lyngfelt et al. 2018), yielding a multilingual resource (Czulo et al. in press). By doing so, frames evoked by constructions may be included as well. This is possible since numerous projects use similar annotation procedures for building their repositories (Lyngfelt et al. 2018; Boas, Torrent and Lyngfelt 2019); specifically, the Brazilian, Japanese, and German constructicon closely follow the agenda formulated in Fillmore et al. (2012) and thus share a battery of basic features.

Yet another domain of application relates to language pedagogy. Including collo-profiles in construction entries invites to specifically address special users, most prominently language learners. Even though the German Constructicon currently focuses on documenting (relations between) constructions in German, it goes far beyond a mere reference work: As the coverage of constructions progresses, the emerging repository becomes more useful as a reference tool for various practical purposes; it may serve, for example, as a resource for language learning and teaching by helping to create learner-specific gap-filling tasks based on formal and semantic constraints as evidenced in the construction entries.

Needless to say that collo-profiles as introduced and elaborated prominently by Thomas Herbst are invaluable for any kind of constructicon.

Literature

- Bäckström, Linnéa, Lars Borin, Markus Forsberg, Benjamin Lyngfelt, Julia Prentice & Emma Sköldbberg. 2013. Automatic Identification of Construction Candidates for a Swedish Constructicon. *Proceedings of the Workshop on Lexical Semantic Resources for NLP at NODALIDA 2013* (NEALT Proceedings Series 19 / Linköping Electronic Conference Proceedings 88), 2–11.
- Barteld, Fabian & Alexander Ziem. 2020. Construction mining: Identifying construction candidates for the German constructicon. *Belgian Journal of Linguistics* 34. 5–16. <https://doi.org/10.1075/bjl.00030.bar>
- Boas, Hans & Alexander Ziem. 2018. Constructing a constructicon for German: Empirical, theoretical, and methodological issues. In Ben Lyngfelt, Lars Borin, Kyoko Ohara & Tiago Torrent (eds.), *Constructicography: Constructicon development across languages*, 183–228. Amsterdam: Benjamins.
- Boas, Hans C. 2010. The syntax-lexicon continuum in construction grammar: A case study of English communication verbs. *Belgian Journal of Linguistics* 24. 57–86.
- Boas, Hans C., Benjamin Lyngfelt & Tiago Torrent. 2019. Framing constructicography. In *Lexicographica* 35. 41–95.
- Brants, Sabine, Stephanie Dipper, Peter Eisenberg, Silvia Hansen-Schirra, Esther König, Wolfgang Lezius, Christian Rohrer, George Smith & Hans Uszkoreit. 2004. TIGER: Linguistic interpretation of a German corpus. *Research on Language and Computation* 2(4). 597–620. <https://doi.org/10.1007/s11168-004-7431-3>
- Broccias, Cristiano. 2012. The syntax-lexicon continuum. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 735–747. Oxford: Oxford University Press.
- Czulo, Oliver, Tiago Torrent, Alexander Willich & Alexander Ziem. 2023. A multilingual approach to the interaction between frames and constructions: towards a joint framework and methodology. *Frames and Constructions* 15(1). 59–90.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee & Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Jill Burstein, Christy Doran, Tamar Solorio (eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol. 1 (Long and Short Papers)*, 4171–4186. Minneapolis, Minnesota: Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Digitales Wörterbuch der deutschen Sprache [Digital Dictionary of the German Language]*. <https://www.dwds.de/>
- Dobrovolskij, Dmitrij. 2011. Phraseologie und Konstruktionsgrammatik. In Alexander Lasch & Alexander Ziem (eds.), *Konstruktionsgrammatik III: Aktuelle Fragen und Lösungsansätze*, 111–130. Tübingen: Narr.
- Fankhauser, Peter, & Kupietz, Marc. 2022. Count-based and predictive language models for exploring DeReKo. In Piotr Bański, Adrien Barbaresi, Simon Clematide, Marc Kupietz, & Harald Lungen (eds.), *Proceedings of the LREC 2022 Workshop on Challenges in the Management of Large Corpora*, 27–31. Marseilles: European Language Resources Association (ELRA).

- Fehrmann, Ingo. 2018. *Kausative Konstruktionen mit dem Verb 'machen' im Deutschen*. Berlin: Humboldt-Universität zu Berlin Doctoral thesis. <https://edoc.hu-berlin.de/handle/18452/20169>
- Feldmüller, Tim, Fabian Barteld & Alexander Ziem. in prep. Can BERT predict construction fillers? A new approach to collo-profiles for the German Constructicon.
- Fillmore, Charles J. 1988. The mechanisms of “construction grammar”. In Shelley Axmaker (ed.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 35–55. Berkeley: Berkeley Linguistics Society.
- Fillmore, Charles J., Russell Lee-Goldman & Russell Rhodes. 2012. The FrameNet constructicon. Hans C. Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 283–299. Stanford: CSLI Publications.
- Fillmore, Charles J. & Paul Kay & Mary Catherine O’Connor. 1988. Regularity and idiomatity in grammatical constructions. The case of let alone. *Language* 64(3), 501–538.
- Finkbeiner, Rita. 2018. Wie soll die Grammatikschreibung mit Konstruktionen umgehen? In Eric Fuß & Angelika Wöllstein (eds.), *Grammatikographie und Grammatiktheorie*, 139–173. Tübingen: Narr.
- Firth, John R. 1951/1957. Modes of meaning. In Firth, John R. (ed.), *Papers in linguistics, 1934-1951*, 190–215. Oxford: Oxford University Press.
- Flach, Susanne. 2021. Collostructions: An R implementation for the family of collostructional methods (v.0.2.0) [Computer software]. <https://sfla.ch/collostructions/>
- Fuß, Eric & Angelika Wöllstein. 2018. Einleitung: Grammatiktheorie und Grammatikographie. In Eric Fuß & Angelika Wöllstein (eds.), *Grammatiktheorie und Grammatikographie*, 7–30. Tübingen: Narr.
- Goldberg, Adele. 2003. Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics* 9(1). 97–129.
- Herbst, Thomas. 2018. Is language a collostruction? – a proposal for looking at collocations, valency, argument structure and other constructions. In Pascual Cantos-Gómez & Moisés Almela-Sánchez (eds.), *Lexical Collocation Analysis: Advances and Applications*, 1–22. Cham: Springer.
- Herbst, Thomas. 2020. Constructions, generalizations, and the unpredictability of language: Moving towards collostruction grammar. *Constructions and Frames* 12(1). 56–96.
- Herbst, Thomas & Thomas Hoffmann. 2018. Construction grammar for students: A constructionist approach to syntactic analysis (CASA). *Yearbook of the German Cognitive Linguistics Association* 6. 197–218.
- Janda, Laura A., Anna Endresen, Valentina Zhukova, Daria Mordashova & Ekaterina Rakhilina. 2020. How to build a constructicon in five years: The Russian example. *Belgian Journal of Linguistics* 34(1), 162–175.
- Jurafsky, Dan. 1992. An on-line computational model of human sentence interpretation. *AAAI’92: Proceedings of the tenth national conference on Artificial intelligence*, 302–308
- Kupietz, Marc, Harald Lungen, Paweł Kamocki & Andreas Witt. 2018. The German reference corpus DeReKo: New developments – new opportunities. *Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018)*, 4353–4360.

- Lyngfelt, Benjamin. 2018. Constructicon and constructicography. In Benjamin Lyngfelt, Lars Borin, Kyoko Ohara & Tiago Torrent (eds.), *Constructicography: Constructicon development across languages* (Constructional Approaches to Language 22), 1–13. Amsterdam, Philadelphia: John Benjamins.
- Lyngfelt, Benjamin, Lars Borin, Kyoko Ohara, Tiago Torrent (eds.). 2018. *Constructicography: Constructicon development across languages* (Constructional Approaches to Language 22). Amsterdam, Philadelphia: John Benjamins.
- Mikolov, Tomas, Kai Chen, Greg Corrado & Jeffrey Dean. 2013. Efficient estimation of word representations in vector space (Version 3). *arXiv*. <https://doi.org/10.48550/ARXIV.1301.3781>
- Mosel, Ulrike. 2006. Grammaticography. The art and craft of writing grammars. In Felix Ameka, Alan Charles & Nicholas Evans (eds.), *Catching language. The standing challenge of grammar writing*, 41–68. Berlin: de Gruyter.
- Osswald, Rainer. 2015. Syntax and lexicography. In Aretmis Alexiadou & Tibor Kiss (eds.), *Syntax - theory and analysis* (Handbuch Sprache und Kommunikation 3), 1963–2000. Berlin & New York: de Gruyter.
- Oxford Advanced Learner's Dictionary*. 2020. 10th edn. Oxford: Oxford University Press.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243. <https://doi.org/10.1075/ijcl.8.2.03ste>
- Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Co-varying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1). 1–43.
- Tayyar Madabushi, Harish, Laurence Romain, Dagmar Divjak & Petar Milin. 2020. CxGBERT: BERT meets construction grammar. In Donia Scott, Nuria Bel & Chengqing Zong (eds.), *Proceedings of the 28th International Conference on Computational Linguistics*, 4020–4032. <https://doi.org/10.18653/v1/2020.coling-main.355>
- Veenboer, Tim & Jelke Bloem. 2023. Using collostructional analysis to evaluate BERT's representation of linguistic constructions. *Findings of the Association for Computational Linguistics: ACL 2023*, 12937–12951. <https://aclanthology.org/2023.findings-acl.819.pdf>
- Wada, Takashi, Timothy Baldwin, Yuji Matsumoto & Jey Han Lau. 2022. Unsupervised lexical substitution with decontextualised embeddings. *Proceedings of the 29th International Conference on Computational Linguistics*, 4172–4185. <https://aclanthology.org/2022.coling-1.366>
- Weissweiler, Leonie, Taiqi He, Naoki Otani, David R. Mortensen, Lori Levin & Hinrich Schütze. 2023. Construction grammar provides unique insight into neural language models. *Proceedings of the First International Workshop on Construction Grammars and NLP (CxGs+NLP, GURT/SyntaxFest 2023)*, 85–95. <https://aclanthology.org/2023.cxgsnlp-1.10>
- Ziem, Alexander & Johanna Flick. 2019. Constructicography at work: implementation and application of the German constructicon. *Yearbook of the German Cognitive Linguistics Association* 7. 201–214.
- Ziem, Alexander. 2018. Tag für Tag Arbeit über Arbeit: konstruktionsgrammatische Zugänge zu Reduplikationsstrukturen im Deutschen. In Kathrin Steyer (ed.), *Sprachliche Verfestigung. Wortverbindungen, Muster, Phrasem-Konstruktionen* (Studien zur Deutschen Sprache 79), 25–48. Tübingen: Narr.
- Ziem, Alexander. 2020. Wenn sich FrameNet und Konstruktikon begegnen: erste Annäherungsversuche zwischen zwei neuen Repositorien zum Deutschen. In Michel Lefèvre & Katharina Mucha (eds.), *Konstruktionen, Kollokationen, Muster*, 13–38. Tübingen: Stauffenburg.

- Ziem, Alexander. 2022. Konstruktionelle Arbeitsteilung im Lexikon-Grammatik-Kontinuum: das Beispiel sprachlicher Kodierungen von Quantität. In Carmen Mellado Blanco, Fabio Mollica & Elmar Schafroth (eds.), *Konstruktionen zwischen Lexikon und Grammatik. Phrasem-Konstruktionen monolingual, bilingual und multilingual*, 19–54. Berlin: de Gruyter.
- Ziem, Alexander & Alexander Willich. 2023. Familienähnlichkeiten im Konstruktikon: von Frame-Familien zu Konstruktionsfamilien. In Fabio Mollica & Sören Stumpf (eds.), *Konstruktionsgrammatik IX. Konstruktionsfamilien im Deutschen*, 55–102. Tübingen: Stauffenburg.