

Towards a Constructicon for German

Alexander Ziem, Hans C. Boas

University of Düsseldorf, Germany, and The University of Texas at Austin, USA
ziem@phil.hhu.de
hcb@mail.utexas.edu

Abstract

This paper presents the ongoing activities of a pilot project aimed at building a prototype constructicon of German. To limit the scope and domain, our pilot project focuses on full text annotation of a first year textbook of German developed at the University of Texas at Austin. Based on an existing lexical database of words evoking semantic frames in the German on-line textbook (Boas/Dux 2013; Boas/Dux/Ziem 2016), we show how these lexical entries can be systematically complemented by construction entries.

Introduction

Our paper presents first results achieved in a pilot project on the development of a constructicon for German. It builds on previous efforts to create an on-line learner's dictionary of German for first year students at the University of Texas at Austin (Boas/Dux 2013, Boas/Dux/Ziem 2016). Designed as a collaboration between UT Austin and HHU Düsseldorf, the German constructicon project uses the first-year online German textbook "Deutsch im Blick" (<http://coerll.utexas.edu/dib/>) for full-text annotation of both lexical items (frame-based) and grammatical structures (construction-based). By linking resources for both the manual annotation work and the web-based storing of constructions and frames in a database it is in line with the FrameNet constructicon project (Fillmore et al. 2012). The constructicon we have in mind primarily builds on thoroughly annotated corpus examples illustrating (a) the construction evoking elements (CEEs), (b) the range of construction elements (CEs) specifying the construction, and (c) the syntactic variation of these CEs, together with information about any semantic frames that are evoked by the construction.

The aims of our paper are threefold: First, we discuss empirical, theoretical, and methodological issues that arise in the course of creating a constructicon for German. By discussing a set of grammatical constructions in contemporary German and comparing them with their equivalents in English, we aim to single out to what extent constructions

in German exhibit commonalities but also idiosyncrasies that need to be taken into account when constructing a constructicon for German. Second, based on our results, we show that the benefit of mapping English constructions, as, for example, documented in the prototype of the Berkeley FrameNet constructicon (Fillmore et al. 2012), to their counterparts in German is limited to a relatively small number of constructions. Other constructions require additional treatments both in terms of their syntactic behavior as well as their grammatical realization patterns and their semantic properties including pragmatic constraints. Third, we discuss ways of further extending our database by way of integrating annotated and analyzed constructions at different levels of abstraction and complexity (following the continuum from lexical constructions to complex sentence-level constructions, such as sentence types).

Syntactic peculiarities of German and the limits of re-using existing constructicon resources

German syntax is characterized by some typologically interesting peculiarities that should be taken into account when building a constructicon. For example, important differences from English syntax include (a) word order (Webelhuth 1992, Kathol 2000), (b) topological fields (Wöllstein 2010), (c) the case system (Zifonun et al. 1997), (d) the passive (Ackerman & Webelhuth 1998, Lasch 2016), and (e) (semi-)idiomatic constructions (Oya 1999, Boas 2003, Engelberg et al. 2010, Ziem/Staffeldt 2011). Such syntactic particularities of German belong to the core of German grammar. They have important consequences for the architecture of a German constructicon, particularly if (parts of) the English prototype constructicon is reused (for a description a English constructicon cf. Fillmore 2008; Fillmore et al. 2012; for an overview: Ziem 2014a).

In contrast to the language-internal strategy pursued by the constructicon projects for Swedish (Lyngfelt et al. 2012), Japanese (Ohara 2014), and Brazilian Portuguese (Torrent et al. 2014), we are also interested in exploring

what types of information from English constructions can be re-used for developing comparable construction entries for German. The results of our discussion form the basis for outlining a contrastive methodology that relies on both (a) a contrastive extension of English construction entries to German, and (b) language-internal analysis and writing of construction entries in cases in which the contrastive approach is not fruitful (see Boas 2014).

How could (and should) the empirical and theoretical insights about the syntax-lexicon continuum drive the design of a constructicon for German? Building on prior research such as Boas (2014), Ziem (2014a), Ziem/Boas/Ruppenhofer (2013), Ziem/Ellsworth (2016), and Boas/Dux/Ziem (2016), our project starts off with investigating what types of construction entries from the English constructicon (Fillmore et al. 2012) can be re-used for creating parallel construction entries for a German constructicon (similar to proposals in Boas (2002) for re-using English semantic frames for other languages). More specifically, we discuss and compare three constructions in German and English, ranging from quasi-synonymous and structurally homologous ones, such as the *just because ... doesn't mean* construction, to constructions with significant language-specific characteristics, such as the *way*-construction (Goldberg 1995: 199-218, Oya 1999) and the family of exclamative constructions (d'Avis 2013, Michaelis 2001, Ziem/Ellsworth 2016). The empirical evidence leads us to propose a “continuum of constructional correspondence” to argue that re-using English construction entries has only limited benefits. We therefore propose a language-specific corpus-based methodology that focuses on the creation of German-specific construction entries by primarily relying on syntactic and semantic categories of German.

What's in a German Constructicon?

Following Fillmore (2008; Fillmore et al. 2012) our pilot project seeks to integrate constructions into a lexical frame-type database. Since we use the formalisms employed in the FrameNet constructicon in a slightly simplified way, we briefly introduce the most important annotation categories before turning to three types of constructions illustrating the continuum of (non-)correspondences between English constructions and their German counterparts.

We begin with the linguistic unit evoking a construction, which is called a ‘Construction Evoking Element’ (CEE; see also Fillmore et al. 2012: Section 2.2). To illustrate, consider (1), an instantiation of an exclamative construction (Ziem/Ellsworth 2016).

- (1) Was für ein spektakulärer Blick!
 What for a spectacular view
 ‘What a spectacular view!’

In (1), the lexical items *was für* (‘what’) serve as the CEE, that is, these elements make up the lexical ‘anchors’ of this (subtype of the) exclamative construction. The complete expression, the so-called construct licensed by the exclamative construction, comprises the scope of the surprise conveyed by the exclamative construction. Since the meaning of the construction is determined by the *Experience_obj* frame (see <http://framenet.icsi.berkeley.edu>), its construction elements (CEs) can also be annotated with recourse to the FEs constituting the *Experience_obj* frame. Specifically, the scope of the surprise equates with the frame element *STIMULUS*. Hence, CEs can be defined as those constituents, or slots, of a syntagmatically complex construction that are instantiated by the respective parts of constructs.

Constructional annotations help describe and define a construction appropriately. In a first step, we identify the CEE. Note that, in contrast to frame-semantic annotations, a target LU providing an explicit link to the construction is often missing. This is because not all constructions entail (more or less) fixed lexical constituents. The more schematic a construction gets, the more likely it is that it does not include one or more fixed lexical items. Hence, in the case of fully schematic constructions (like, for example, the subject-predicate construction), structural properties rather than lexical items evoke a construction. Thus, in these cases, a configuration of structural properties serves as CEE for the construction.

In a second step, we name those parts of sample sentences that form the constituents of the constructs licensed by the construction. These components are labeled as elements of the construction.

Following this procedure, (2) exemplifies the annotation regarding (a) the CEE, (b) the CEs and their functions within the construction, and (c) the construct that is licensed by the construction. Following FrameNet annotation conventions, we tag CEs with square brackets and constructs with curly brackets, while labeling the meanings or functions of these elements with the help of subscripts.

- (2) {[_{CEE}<What>] a [_{DEGREE} spectacular] [_{STIMULUS}view of the city]!}.

Note that (2) does not yet include annotations of the grammatical functions and phrase types of each of the CEs (if applicable). In line with the descriptions of the respective FEs in the *Experience_obj* frame in FrameNet (Fillmore/Baker 2010; <http://framenet.icsi.berkeley.edu>), the CEs realized in (2) can be defined as follows:

- DEGREE is the degree to which the STIMULUS brings about an emotion – that is, surprise in the case of an exclamative construction – in the EXPERIENCER.
- STIMULUS is the event or entity, which brings about the emotional or psychological state – that is, surprise in the case of an exclamative construction – of the EXPERIENCER.

Overall, there is a plethora of information that goes into a constructional entry in a German constructicon. Full descriptions of grammatical constructions should include, but are not limited to the following:

- lists of the construction-evoking elements (CEEs)
- descriptions of the construction's lexical head, if applicable
- descriptions of construction elements (CEs), including the function of each CE within a construction as well as the phrase types in which each CE may be realized
- illustrations and descriptions of the realization patterns of a construction
- reports on pragmatic, semantic, and syntactic constraints (preemption)
- explanations of collostructional preferences for each CE, if applicable
- explanations of covariational preferences of CEs, if applicable
- annotated sample sentences illustrating the range of realization patterns
- definitions of both form- and meaning-related relations connecting a construction to other constructions in the constructicon.

Clearly, providing all information for each grammatical construction in German is a very ambitious endeavor. Even in the case of well-documented constructions, not all information required for a full construction entry is available.

Towards a German Constructicon

It is a hard and winding road from an English to a German constructicon. Even though there are one-to-one constructional correspondences between English and German constructions, such as the *just because doesn't mean*-construction and its German counterpart, many English constructions do not have clear-cut German equivalents. The *way*-construction and the German reflexive motion construction fall into this category. The fact that numerous basic German constructions do not have a straightforward English counterpart at all makes the situation even more complicated.

We take these findings as empirical support for doubting the usefulness of the Berkeley FrameNet constructicon, or any other constructicon, for creating parallel construction entries without questioning the annotation schema developed there. To be as comprehensive and precise as possi-

ble, a constructicon's architecture is required that meets the most fundamental grammatical requirements peculiar to German. In this view, the empirical evidence discussed so far suggests that re-using English construction entries is not always helpful. We therefore propose to start with parallel construction entries, focusing solely on language-internal evidence from German as the basis for construction entries. This will ensure that the German constructicon will evolve in the style of the FrameNet constructicon while remaining at the same time conceptually independent of it. The corpus-based methodology we have in mind first focuses on the creation of German-specific construction entries by primarily relying on syntactic and semantic categories of German. This approach has the advantage of first providing detailed lexico-syntactic construction entries for German, linking these in larger networks of (families of) constructions.

References

- Ackerman, F. and Webelhuth, G. eds. 1998. *A Theory of Predicates*. Stanford: CSLI Publications.
- Boas, H. C. 2002. Bilingual FrameNet dictionaries for machine translation. In *Proceedings of the Third International Conference on Language Resources and Evaluation*, Vol. IV: 1364–1371. Las Palmas, Spain.
- Boas, H. C. 2003. *A constructional approach to resultatives*. Stanford: CSLI Publications.
- Boas, H. C. 2014. Zur Architektur einer konstruktionsbasierten Grammatik des Deutschen. In *Grammatik als Inventar von Konstruktionen? Sprachliches Wissen im Fokus der Konstruktionsgrammatik*, 37–63. Berlin, New York: de Gruyter.
- Boas, H.C., and Dux, R. 2013. Semantic frames for foreign-language education: Towards a German frame-based dictionary. *Veridas On-Line* 1/2013, 81–100.
- Boas, H. C., Dux, R. and Ziem, A. 2016. Frames and constructions in an online learner's dictionary of German. In *Applied Construction Grammar*. Berlin: de Gruyter.
- D'Avis, F. 2013. Exklamativsatz. In *Satztypen des Deutschen*, 171–201. Berlin, Boston: de Gruyter.
- Engelberg, S.; König, S.; Proost, K. and Winkler, E. 2011. Argumentstrukturmuster als Konstruktionen? In *Sprachliches Wissen zwischen Lexikon und Grammatik*, 71–112. Berlin, New York: de Gruyter.
- Fillmore, C. J. 2008. Border Conflicts: FrameNet meets Construction Grammar. *Proceedings of the XIII EURALEX International Congress* 15–19 July 2008: 49–68.
- Fillmore, C. J. and Baker C. 2010. A frame approach to semantic analysis. In *The Oxford Handbook of Linguistic Analysis*, 313–340. Oxford: Oxford University Press.
- Fillmore, C. J.; Lee-Goldman, R. R. and Rhomieux, R. 2012. The FrameNet Constructicon. In *Sign-based Construction Grammar*, 309–372. Stanford: CSLI Publications.
- Goldberg, A. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

- Kathol, A. 2000. *Linear Syntax*. Oxford: Oxford University Press.
- Lasch, A. 2016. *Nonagentive Konstruktionen des Deutschen*. Berlin, Boston: de Gruyter.
- Lyngfelt, B.; Borin L.; Forsberg M.; Prentice J.; Rydstedt R.; Sköldberg E. and Tingsell, S. 2012. Adding a construction to the Swedish resource network of Språkbanken. In *Proceedings of KONVENS 2012*, 452–461. Vienna.
- Michaelis, L. A. 2001. Exclamative Constructions. In *Language Universals and Language Typology: An International Handbook*, 1038-1050. Berlin: de Gruyter.
- Ohara, K. H. 2014. Relating Frames and Constructions in Japanese FrameNet. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC '14), 2474–2477.
- Oya, T. 1999. Er bettelt sich durchs Land: Eine one's way Konstruktion im Deutschen? *Deutsche Sprache* 27, 356–369.
- Torrent, T.; Lange, L. M.; Sampaio, T. F.; da Silva Tavares, T. and da Silva Matos, E. E. 2014. Revisiting border conflicts between FrameNet and Construction Grammar: Annotation policies for the Brazilian Portuguese Constructicon. *Constructions and Frames* 6.1, 33–50.
- Webelhuth, G. 1992. *Principles and Parameters of Syntactic Saturation*. Oxford: Oxford University Press.
- Wöllstein, A. 2010. *Topologisches Satzmodell*. Heidelberg: Winter.
- Ziem, A. and Staffeldt, S. 2011. Compositionality and embodied meanings of phraseologisms: the case of somatisms. In *Converging evidence: Methodological and Theoretical Issues for Linguistic Research*, 195–219. Amsterdam, Philadelphia: Benjamins.
- Ziem, A.; Boas, H. C. and Ruppenhofer J. 2014. “Grammatische Konstruktionen und semantische Frames fuer die Textanalyse,” *Syntaxtheorien. Analysen im Vergleich*, 297–333.
- Ziem, A. 2014a. Von der Kasusgrammatik zum FrameNet: Frames, Konstruktionen und die Idee eines Konstruktikons. In *Grammatik als Inventar von Konstruktionen? Sprachwissen im Fokus in der Konstruktionsgrammatik*, 263–290. Berlin, New York: de Gruyter.
- Ziem, A. and Ellsworth, M. 2016. Exklamativsätze im FrameNet-Konstruktikon. In: *Satztypen und Konstruktionen im Deutschen*, 146–191. Berlin, Boston: de Gruyter.
- Zifonun, G.; Hoffmann, L. and Strecker, B. 1997. *Grammatik der deutschen Sprache*. 3 volumes. Berlin, New York: Mouton de Gruyter.