Chapter 32

HPSG and Construction Grammar

Stefan Müller
Humboldt-Universität zu Berlin

This chapter discusses the main tenets of Construction Grammar (CxG) and shows that HPSG adheres to them. The discussion includes surface orientation, language acquisition without UG, and inheritance networks and shows how HPSG (and other frameworks) are positioned along these dimensions. Formal variants of CxG will be briefly discussed and their relation to HPSG will be pointed out. It is argued that lexical representations of valence are more appropriate than phrasal approaches, which are assumed in most variants of CxG. Other areas of grammar seem to require headless phrasal constructions (e.g., the NPN construction and certain extraction constructions) and it is shown how HPSG handles these. Derivational morphology is discussed as a further example of an early constructionist analysis in HPSG.

This chapter deals with Construction Grammar (CxG) and its relation to Head-Driven Phrase Structure Grammar (HPSG). The short version of the message is: HPSG is a Construction Grammar. It had constructional properties right from the beginning and over the years – due to influence by Construction Grammarians like Fillmore and Kay – certain aspects were adapted, making it possible to better capture generalizations over phrasal patterns. In what follows I will first say what Construction Grammars are (Section 1), and I will explain why HPSG as developed in Pollard & Sag (1987; 1994) was a Construction Grammar and how it was changed to become even more Constructive (Section 1.2.3). Section 2 deals

1This does not mean that HPSG is not a lot of other things at the same time. For instance, it is also a Generative Grammar in the sense of Chomsky (1965: 4), that is, it is explicit and formalized. HPSG is also very similar to Categorial Grammar (Müller 2013; Kubota 2021, Chapter 29 of this volume). Somewhat ironically, Head-Driven Phrase Structure Grammar is not entirely head-driven anymore (see Section 4.1), nor is it a phrase structure grammar (Richter 2021, Chapter 3 of this volume).
with so-called argument structure constructions, which are usually dealt with by assuming phrasal constructions in CxG, and explains why this is problematic and why lexical approaches are more appropriate. Section 3 explains Construction Morphology, Section 4 shows how cases that should be treated phrasally can be handled in HPSG, and Section 5 sums up the chapter.

1 What is Construction Grammar?

Construction Grammar was developed as a theory that can account for non-regular phenomena as observed in many idioms (Fillmore, Kay & O’Connor 1988). It clearly set itself apart from theories like Government & Binding (Chomsky 1981), which assumes very abstract schemata for the combination of lexical items (\(\overline{X}\) rules). The argument was that grammatical constructions are needed to capture irregular phenomena and their interaction with more regular ones. In contrast, Chomsky (1981: 7) considered rules for passive or relative clauses as epiphenomenal; everything was supposed to follow from general principles.\(^2\) According to Chomsky, grammars consisted of a set of general combinatorial rules and some principles. The Minimalist Program (Chomsky 1995) is even more radical, since only two combinatorial rules are left (External and Internal Merge). Various forms of CxG object to this view and state that several very specific phrasal constructions are needed in order to account for language in its entirety and full complexity. Phenomena for which this is true will be discussed in Section 4. However, the case is not as clear in general, since one of the authors of Fillmore, Kay & O’Connor (1988) codeveloped a head-driven, lexical theory of idioms that is entirely compatible with the abstract rules of Minimalism (Sag 2007; Kay, Sag & Flickinger 2015; Kay & Michaelis 2017). This theory will be discussed in Section 1.3.2.1. Of course, the more recent lexical theory of idioms is a constructional theory as well. So the first question to answer in a chapter like this is: what is a construction in the sense of Construction Grammar? What is Construction Grammar? While it is relatively clear what a construction is, the answer to the question regarding Construction Grammar is less straight-forward (see also Fillmore 1988: 35 on this). Section 1.1 provides the definition for the term construction.

\(^2\)The passive in GB is assumed to follow from suppression of case assignment and the Case Filter, which triggers movement of the object to SpecIP. The important part of the analysis is the combination of the verb stem with the passive morphology. This is where suppression of case assignment takes place. This morphological part of the analysis corresponds to the Passive Construction in theories like HPSG and SBCG: a lexical rule (Pollard & Sag 1987: 215; Müller 2003a; Müller & Ørsnes 2013; Wechsler, Koenig & Davis 2021: Section 5.3, Chapter 9 of this volume). So in a sense there is a Passive Construction in GB as well.
and Section 1.2 states the tenets of CxG and discusses to what extent the main frameworks currently on the market adhere to them.

1.1 Defining “construction”

Fillmore, Kay & O’Connor (1988) discuss sentences like (1) and notice that they pose puzzles for standard accounts of syntax and the syntax/semantics interface.

(1) a. The more carefully you do your work, the easier it will get.
    b. I wouldn’t pay five dollars for it, let alone ten dollars.

The the -er the -er Construction is remarkable, since it combines aspects of normal syntax (clause structure and extraction) with idiosyncratic aspects like the special use of the. In (1a) the adverb phrase more carefully does not appear to the left of work but is fronted and the appears without a noun. The second clause in (1a) is structured in a parallel way. There have to be two of these the clauses to form the respective construction. Fillmore, Kay & O’Connor (1988) extensively discuss the properties of let alone, which are interesting for syntactic reasons (the fragments following let alone) and for semantic and information structural reasons. I will not repeat the discussion here but refer the reader to the paper.³

In later papers, examples like (2) were discussed:

(2) a. What is this scratch doing on the table? (Kay & Fillmore 1999: 3)
    b. Frank dug his way out of prison. (Goldberg 1995: 199)

Again, the semantics of the complete sentences is not in an obvious relation to the material involved. The question in (2a) is not about a scratch’s actions, but rather the question is why there is a scratch. Similarly, (2b) is special in that there is a directional PP that does not normally go together with verbs like dug. It is licensed by way in combination with a possessive pronoun.

Fillmore et al. (1988), Goldberg (1995), Kay & Fillmore (1999) and Construction Grammarians in general argue that the notion of “construction” is needed for adequate models of grammar, that is, for models of grammar that are capable of analyzing the examples above. Fillmore et al. (1988: 501) define construction as follows:

Constructions on our view are much like the nuclear family (mother plus daughters) subtrees admitted by phrase structure rules, EXCEPT that (1) constructions need not be limited to a mother and her daughters, but may span

³For an analysis of comparative correlative constructions as in (1a) in HPSG, see Abeillé & Chaves (2021: Section 3.3), Chapter 16 of this volume and the papers cited there.
wider ranges of the sentential tree; (2) constructions may specify, not only syntactic, but also lexical, semantic, and pragmatic information; (3) lexical items, being mentionable in syntactic constructions, may be viewed, in many cases at least, as constructions themselves; and (4) constructions may be idiomatic in the sense that a large construction may specify a semantics (and/or pragmatics) that is distinct from what might be calculated from the associated semantics of the set of smaller constructions that could be used to build the same morphosyntactic object. (Fillmore et al. 1988: 501)

A similar definition can be found in Goldberg’s work. Goldberg (2006: 5) defines construction as follows:

Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency. (Goldberg 2006: 5)

The difference between this definition and earlier definitions by her and others is that patterns that are stored because of their frequencies are included. This addition is motivated by psycholinguistic findings that show that forms may be stored even though they are fully regular and predictable (Bybee 1995; Pinker & Jackendoff 2005: 228).

Goldberg provides Table 32.1 as examples of constructions. In addition to such

<table>
<thead>
<tr>
<th>Table 32.1: Examples of constructions, varying in size and complexity according to Goldberg (2009: 94)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word</strong></td>
</tr>
<tr>
<td><strong>Word (partially filled)</strong></td>
</tr>
<tr>
<td><strong>Complex word</strong></td>
</tr>
<tr>
<td><strong>Idiom (filled)</strong></td>
</tr>
<tr>
<td><strong>Idiom (partially filled)</strong></td>
</tr>
<tr>
<td><strong>Covariational Conditional</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Ditransitive</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Passive</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
constructions with a clear syntax-semantics or syntax-function relation, Goldberg (2013: 453) assumes a rather abstract VP construction specifying “statistical constraints on the ordering of postverbal complements, dependent on weight and information structure”.

If one just looks at Goldberg’s definition of construction, all theories currently on the market could be regarded as Construction Grammars. As Peter Staudacher pointed out in the discussion after a talk by Knud Lambrecht in May 2006 in Potsdam, lexical items are form-meaning pairs and the rules of phrase structure grammars come with specific semantic components as well, even if it is just functional application. So, Categorial Grammar, GB-style theories paired with semantics (Heim & Kratzer 1998), GPSG, TAG, LFG, HPSG, and even Minimalism would be Construction Grammars. If one looks at the examples of constructions in Table 32.1, things change a bit. Idioms are generally not the focus of work in Mainstream Generative Grammar (MGG). MGG is usually concerned with explorations of the so-called Core Grammar as opposed to the Periphery, to which the idioms are assigned. The Core Grammar is the part of the grammar that is supposed to be acquired with help of innate domain specific knowledge, something whose existence Construction Grammar denies. But if one takes Hauser, Chomsky & Fitch (2002) seriously and assumes that only the ability to form complex linguistic objects out of less complex linguistic objects (Merge) is part of this innate knowledge, then the core/periphery distinction does not have much content and after all, Minimalists could adopt a version of Sag’s local, selection-based analysis of idioms (Sag 2007; Kay, Sag & Flickinger 2015; Kay & Michaelis 2017). However, as is discussed in the next subsection, there are other aspects that really set Construction Grammar apart from MGG.

1.2 Basic tenets of Construction Grammar

Goldberg (2003: 219) names the following tenets as core assumptions standardly made in CxG:

**Tenet 1** All levels of description are understood to involve pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully abstract phrasal patterns. (See Table 32.1.)
Tenet 2 An emphasis is placed on subtle aspects of the way we conceive of events and states of affairs.

Tenet 3 A “what you see is what you get” approach to syntactic form is adopted: no underlying levels of syntax or any phonologically empty elements are posited.

Tenet 4 Constructions are understood to be learned on the basis of the input and general cognitive mechanisms (they are constructed), and are expected to vary cross-linguistically.

Tenet 5 Cross-linguistic generalizations are explained by appeal to general cognitive constraints together with the functions of the constructions involved.

Tenet 6 Language-specific generalizations across constructions are captured via inheritance networks much like those that have long been posited to capture our non-linguistic knowledge.

Tenet 7 The totality of our knowledge of language is captured by a network of constructions: a “constructicon”.

I already commented on Tenet 1 above. Tenet 2 concerns semantics and the syntax-semantics interface, which are part of most HPSG analyses. In what follows I want to look in more detail at the other tenets. Something that is not mentioned in Goldberg’s tenets but is part of the definition of construction by Fillmore et al. (1988: 501) is the non-locality of constructions. I will comment on this in a separate subsection.

1.2.1 Surface orientation and empty elements

Tenet 3 requires a surface-oriented approach. Underlying levels and phonologically empty elements are ruled out. This excludes derivational models of transformational syntax assuming an underlying structure (the so-called D-structure) and some derived structure or more recent derivational variants of Minimalism. There was a time where representational models of Government & Binding (GB, Chomsky 1981) did not assume a D-structure but just one structure with traces (Koster 1978: 1987: 235; Kolb & Thiersch 1991; Haider 1993: Section 1.4; Frey 1993: 14; Lohnstein 1993: 87–88, 177–178; Fordham & Crocker 1994: 38; Veenstra 1998: 58). Some of these analyses are rather similar to HPSG analyses as they are assumed today (Kiss 1995; Bouma & van Noord 1998; Meurers 2000; Müller 2005;
Chomsky’s Minimalist work (Chomsky 1995) assumes a derivational model and comes with a rhetoric of building structure in a bottom-up way and sending complete phases to the interfaces for pronunciation and interpretation. This is incompatible with Tenet 3, but in principle, Minimalist approaches are very similar to Categorial Grammar, so there could be representational approaches adhering to Tenet 3.\footnote{There is a variant of Minimalist Grammars (Stabler 2011), namely Top-down Phase-based Minimalist Grammar (TPMG) as developed by Chesi (2012; 2007) and Bianchi & Chesi (2006; 2012). There is no movement in TPMG. Rather, wh-phrases are linked to their “in situ” positions with the aid of a short-term memory buffer that functions like a stack. See also Hunter (2010; 2019) for a related account where the information about the presence of a wh-phrase is percolated in the syntax tree, like in GPSG/HPSG. For a general comparison of Minimalist grammars and HPSG, see Müller (2013: Section 2.3) and Müller (2020b: 177–180), which includes the discussion of a more recent variant suggested by Torr (2109).}

A comment on empty elements is in order: all articles introducing Construction Grammar state that CxG does not assume empty elements. Most of the alternative theories do use empty elements: see König (1999) on Categorial Grammar, Gazdar, Klein, Pullum & Sag (1985: 143) on GPSG, Bresnan (2001: 67) on LFG, Bender (2001) and Sag, Wasow & Bender (2003: 464) on HPSG/Sign-Based Construction Grammar. There are results from the 60s that show that phrase structure grammars containing empty elements can be translated into grammars that do not contain empty elements (Bar-Hillel, Perles & Shamir 1961: 153, Lemma 4.1) and sure enough there are versions of GPSG (Uszkoreit 1987: 76–77), LFG (Kaplan & Zaenen 1989; Dalrymple et al. 2001), and HPSG (Bouma et al. 2001; Sag 2010: 508) that do not use empty elements. Grammars with empty elements often are more compact than those without empty elements and express generalizations more directly. See for example Bender (2001) for copulaless sentences in African American Vernacular English and Müller (2014) on nounless NPs in German. The argument against empty elements usually refers to language acquisition: it is argued that empty elements cannot be learned since they are not detectable in the input. However, if the empty elements alternate with visible material, it can be argued that what is learned is the fact that a certain element can be left out. What is true, though, is that things like empty expletives cannot be learned since these empty elements are neither visible nor do they contribute to meaning. Their only purpose in grammars is to keep uniformity. For example, Grewendorf (1995) working in GB suggests an analysis of the passive in German that is parallel to the movement-based analysis of English passives (Chomsky 1981: 124). In order to account for the fact that the subject does not move to initial position in German, he suggests an empty expletive pronoun that takes the
subject position and that is connected to the original nonMoved subject. Such elements cannot be acquired without innate knowledge about the IP/VP system and constraints about the obligatory presence of subjects. The CxG criticism is justified here.

A frequent argumentation for empty elements in MGG is based on the fact that there are overt realizations of an element in other languages (e.g., object agreement in Basque and focus markers in Gungbe). But since there is no language-internal evidence for these empty elements, they cannot be learned and one would have to assume that they are innate. This kind of empty element is rightly rejected (by proponents of CxG and others).

Summing up, it can be said that all grammars can be turned into grammars without empty elements and hence fulfill Tenet 3. It was argued that the reason for assuming Tenet 3 (problems in language acquisition) should be reconsidered and that a weaker form of Tenet 3 should be assumed: empty elements are forbidden unless there is language-internal evidence for them. This revised version of Tenet 3 would allow one to count empty element versions of CG, GPSG, LFG, and HPSG among Construction Grammars.

1.2.2 Language acquisition without the assumption of UG

Tenets 4 and 5 are basically what everybody should assume in MGG if Hauser, Chomsky & Fitch (2002) are taken seriously. Of course, this is not what is done in large parts of the field. The most extreme variant is Cinque & Rizzi (2010), who assume at least 400 functional heads (p. 57) being part of Universal Grammar (UG) and being present in the grammars of all languages, although sometimes invisibly (p. 55). Such assumptions beg the question why the genera of Bantu languages should be part of our genome and how they got there. Researchers working on language acquisition realized that the Principles & Parameters approach (Meisel 1995) makes wrong predictions. They now talk about Micro-Cues instead of parameters (Westergaard 2014) and these Micro-Cues are just features that can be learned. However, Westergaard still assumes that the features are determined by UG, a dubious assumption seen from a CxG perspective (and from the perspective of Hauser, Chomsky & Fitch and genetics in general; Bishop 2002).

Note that even those versions of Minimalism that do not follow the Rizzi-style Cartographic approaches are far from being minimalist in their assumptions. Some distinguish between strong and weak features, some assume enumerations of lexical items from which a particular derivation draws its input, and some assume that all movement has to be feature-driven. Still others assume that deriva-
tions work in so-called phases and that a phase, once completed, is “shipped to the interfaces”. Construction of phases is bottom-up, which is incompatible with psycholinguistic results (see also Borsley & Müller 2021: Section 5.1, Chapter 28 in this volume). None of these assumptions is a natural one to make from a language acquisition point of view. Most of these assumptions do not have any empirical motivation; the only motivation usually given is that they result in “restrictive theories”. But if there is no motivation for them, this means that the respective architectural assumptions have to be part of our innate domain-specific knowledge, which is implausible according to Hauser, Chomsky & Fitch (2002).

As research in computational linguistics shows, our input is rich enough to form classes, to determine the part of speech of lexical items, and even to infer syntactic structure thought to be underdetermined by the input. For instance, Bod (2009) shows that the classical auxiliary inversion examples that Chomsky still uses in his Poverty of the Stimulus arguments (Chomsky 1971: 29–33; Berwick, Pietroski, Yankama & Chomsky 2011) can also be learned from language input available to children. See also Freudenthal et al. (2006; 2007) on input-based language acquisition.

HPSG does not make any assumptions about complicated mechanisms like feature-driven movement and so on. HPSG states properties of linguistic objects like part of speech, case, gender, etc., and states relations between features like agreement and government. In this respect it is like other Construction Grammars and hence experimental results regarding theories of language acquisition can be carried over to HPSG. See also Borsley & Müller (2021: Section 5.2), Chapter 28 of this volume on language acquisition.

1.2.3 Inheritance networks

This leaves us with Tenets 6 and 7, that is, inheritance networks and the constructicon. Inheritance is something that is used in the classification of knowledge. For example, the word animal is very general and refers to entities with certain properties. There are subtypes of this kind of entity: mammal and further subtypes like mouse. In inheritance hierarchies, the knowledge of superconcepts is not restated at the subconcepts but instead, the superconcept is referred to. This is like Wikipedia: the Wikipedia entry of mouse states that mice are mammals without listing all the information that comes with the concept of mammal. Such inheritance hierarchies can be used in linguistics as well. They can be used to classify roots, words, and phrases. An example of such a hierarchy used for the classification of adjectives and adjectival derivation is discussed in Section 3. See
also Davis & Koenig (2021: Section 4), Chapter 4 of this volume on inheritance in the lexicon.

MGG does not make reference to inheritance hierarchies. HPSG did this right from the beginning in 1985 (Flickinger, Pollard & Wasow 1985) for lexical items and since 1995 also for phrasal constructions (Sag 1997). LFG rejected the use of types, but used macros in computer implementations. The macros were abbreviatory devices specific to the implementation and did not have any theoretical importance. This changed in 2004, when macros were suggested in theoretical work (Dalrymple, Kaplan & King 2004). And although any connection to constructionist work is vehemently denied by some of the authors, recent work in LFG has a decidedly constructional flavor (Asudeh, Dalrymple & Toivonen 2008; 2014).\(^6\) LFG differs from frameworks like HPSG, though, in assuming a separate level of c-structure. c-structure rules are basically context-free phrase structure rules, and they are not modeled by feature value pairs (but there is a model-theoretic formalization; Kaplan 1995: 12). This means that it is not possible to capture a generalization regarding lexical items, lexical rules, and phrasal schemata, or any two-element subset of these three kinds of objects. While HPSG describes all of these elements with the same inventory and hence can use common supertypes in the description of all three, this is not possible in LFG (Müller 2018b: Section 23.1).\(^7\) For example, Höhle (1997) argued that complementizers and finite verbs in initial position in German form a natural class. HPSG can capture this since complementizers (lexical elements) and finite verbs in initial position (results of lexical rule applications or a phrasal schema, see Müller 2021: Section 5.1, Chapter 10 of this volume) can have a common supertype. TAG is also using inheritance in the Meta Grammar (Lichte & Kallmeyer 2017).

Since HPSG’s lexical entries, lexical rules, and phrasal schemata are all described by typed feature descriptions, one could call the set of these descriptions the constructicon. Therefore, Tenet 7 is also adhered to.

1.2.4 Non-locality

Fillmore, Kay & O’Connor (1988: 501) stated in their definition of constructions that constructions may involve more than mothers and immediate daughters (see

\[^6\]See Toivonen (2013: 516) for an explicit reference to construction-specific phrase structure rules in the sense of Construction Grammar. See Müller (2018a) for a discussion of phrasal LFG approaches.

\[^7\]One could use templates (Dalrymple et al. 2004; Asudeh et al. 2013) to specify properties of lexical items and of mother nodes in c-structure rules, but usually c-structure rules specify the syntactic categories of mothers and daughters, so this information has a special status within the c-structure rules.
That is, daughters of daughters can be specified as well. A straightforward example of such a specification is given in Figure 32.1, which shows the TAG analysis of the idiom *take into account* following Abeillé & Schabes (1989: 7). The fixed parts of the idiom are just stated in the tree. NP↓ stands for an open slot into which an NP has to be inserted. The subscript NA says that adjunction to the respectively marked nodes is forbidden. Theories like Constructional HPSG can state such complex tree structures like TAG can. Dominance relationships are modeled by feature structures in HPSG and it is possible to have a description that corresponds to Figure 32.1. The NP slots would just be left underspecified and can be filled in models that are total (see Richter 2007 and Richter 2021, Chapter 3 of this volume for formal foundations of HPSG).

It does not come without some irony that the theoretical approach that was developed out of Berkeley Construction Grammar and Constructional HPSG, namely Sign-Based Construction Grammar (Sag, Boas & Kay 2012; Sag 2012), is strongly local: it is made rather difficult to access daughters of daughters (Sag 2007). So, if one would stick to the early definition, this would rule out SBCG as a Construction Grammar. Fortunately, this is not justified. First, there are ways to establish nonlocal selection (see Section 1.3.2.1) and second, there are

---

8This subsection is based on a much more thorough discussion of locality and SBCG in Müller (2016: Section 10.6.2.1.1 and Section 18.2).
ways to analyze idioms locally. Sag (2007), Kay, Sag & Flickinger (2015), and Kay & Michaelis (2017) develop a theory of idioms that is entirely based on local selection. For example, for *take into account*, one can state that *take* selects two NPs and a PP with the fixed lexical material *into* and *account*. The right form of the PP is enforced by means of the feature LEXICAL IDENTIFIER (LID). A special word *into* with the LID value *into* is specified as selecting a special word *account*. What is done in TAG via direct specification is done in SBCG via a series of local selections of specialized lexical items. The interesting (intermediate) conclusion is: if SBCG can account for idioms via local selection, then theories like Categorial Grammar and Minimalism can do so as well. So, they cannot be excluded from Construction Grammars on the basis of arguments concerning idioms and non-locality of selection.

However, there may be cases of idioms that cannot be handled via local selection. For example, Richter & Sailer (2009) discuss the following idiom:

(3) glauben, X_Acc tritt ein Pferd
    believe X kicks a horse
    'be utterly surprised'

The X-constituent has to be a pronoun that refers to the subject of the matrix clause. If this is not the case, the sentence becomes ungrammatical or loses its idiomatic meaning.

(4) a. Ich glaube, mich / # dich tritt ein Pferd.
    I believe me.ACC you.ACC kicks a horse
b. Jonas glaubt, ihn tritt ein Pferd.10
    Jonas believes him kicks a horse
    'Jonas is utterly surprised.'
c. # Jonas glaubt, dich tritt ein Pferd.
    Jonas believes you kicks a horse
    'Jonas believes that a horse kicks you.'

Richter & Sailer (2009: 313) argue that the idiomatic reading is only available if the accusative pronoun is fronted and the embedded clause is V2. The examples in (5) do not have the idiomatic reading:

---

9Of course this theory is also compatible with any other variant of HPSG. As Flickinger, Pollard & Wasow (2021: xxiii), Chapter 2 of this volume point out, it was part of the grammar fragment that has been developed at the CSLI by Dan Flickinger (Flickinger, Copestake & Sag 2000; Flickinger 2000; 2011) years before the SBCG manifesto was published.

They develop an analysis with a partly fixed configuration and some open slots, similar in spirit to the TAG analysis in Figure 32.1. However, their restrictions on *Pferd* clauses are too strict since there are variants of the idiom that do not have the accusative pronoun in the *Vorfeld*:

> ich glaube, dass mich ein Pferd tritt.
> I believe that a horse kicks me.

(6)

> ich glaube, ein Pferd tritt mich.
> I believe a horse kicks me.

(5)  

a. Ich glaube, dass mich ein Pferd tritt.
> 'I believe that a horse kicks me.'

b. Ich glaube, ein Pferd tritt mich.
> 'I believe that a horse kicks me.'

So it might be the case that the organization of the embedded clause can be stated clause-internally, and hence it is an open question whether there are idioms that make nonlocal Constructions necessary.

What is not an open empirical question, though, is whether humans store chunks with complex internal structure or not. It is clear that we do, and much Construction Grammar literature emphasizes this. Constructional HPSG can represent such chunks directly in the theory, but SBCG cannot, since linguistic signs do not have daughters. So here, Constructional HPSG and TAG are the theories that can represent complex chunks of linguistic material with its internal structure, while other theories like GB, Minimalism, CG, LFG, SBCG, and DG cannot.

### 1.2.5 Summary

If all these points are taken together, it is clear that most variants of MGG are not Construction Grammars. However, CxG had considerable influence on other frameworks so that there are constructionist variants of LFG, HPSG, and TAG. HPSG in the version of Sag (1997) (also called Constructional HPSG) and the HPSG dialect Sign-Based Construction Grammar are Construction Grammars that follow all the tenets mentioned above.

---

1.3 Variants of Construction Grammar

The previous section discussed the tenets of CxG and to what degree other frameworks adhere to them. This section deals with frameworks that have Construction Grammar explicitly in their name or are usually grouped among Construction Grammars:

- Berkeley Construction Grammar (Fillmore 1985b; 1988; Kay & Fillmore 1999; Fried 2015)
- Cognitive Construction Grammar (Lakoff 1987; Goldberg 1995; 2006)
- Radical Construction Grammar (Croft 2001)
- Embodied Construction Grammar (Bergen & Chang 2005)
- Fluid Construction Grammar (Steels & De Beule 2006; Steels 2011)
- Sign-Based Construction Grammar (Sag 2010; 2012)

Berkeley Construction Grammar, Embodied Construction Grammar, Fluid Construction Grammar, and Sign-Based Construction Grammar are the ones that are more formal. All of these variants use feature value pairs and are constraint-based. They are sometimes also referred to as unification-based approaches. Berkeley Construction Grammar never had a consistent formalization. The variant of unification assumed by Kay & Fillmore (1999) was formally inconsistent (Müller 2006a: Section 2.4) and the computation of construction-like objects (CLOs) suggested by Kay (2002) did not work either (Müller 2006a: Section 3). Berkeley Construction Grammar was dropped by the authors, who joined forces with Ivan Sag and Laura Michaelis and eventually came up with an HPSG variant named Sign-Based Construction Grammar (Sag 2012). The differences between Constructional HPSG (Sag 1997) and SBCG are to some extent cosmetic: semantic relations got the suffix -fr for frame (like-rel became like-fr), phrases were called constructions (hd-subj-ph became subj-head-cxt), and lexical rules were called derivational constructions.\(^{12}\) While this renaming would not have changed anything in terms of expressiveness of theories, there was another change that

\(^{12}\) This renaming trick was so successful that it even confused some of the co-editors of the volume about SBCG (Boas & Sag 2012). See for example Boas (2014) and the reply in Müller & Wechsler (2014b).
was not motivated by any of the tenets of Construction Grammar but rather by the wish to get a more restrictive theory: Sag, Wasow & Bender (2003) and Sag (2007) changed the feature geometry of phrasal signs in such a way that signs do not contain daughters. The information about mother-daughter relations is contained in lexical rules and phrasal schemata (constructions) only. The phrasal schemata are more like GPSG immediate dominance schemata (phrase structure rules without constraints on the order of the daughters) in licensing a mother node when certain daughters are present, but without the daughters being represented as part of the AVM that stands for the mother node, as was common in HPSG from 1985 till Sag, Wasow & Bender (2003). This differs quite dramatically from what was done in Berkeley Construction Grammar, since BCxG explicitly favored a non-local approach (Fillmore 1988: 37; Fillmore et al. 1988: 501). Arguments were not canceled but passed up to the mother node. Adjuncts were passed up as well, so that the complete internal structure of an expression is available at the top-most node (Kay & Fillmore 1999: 9). The advantage of BCxG and Constructional HPSG (Sag 1997) is that complex expressions (e.g., idioms and other more transparent expressions with high frequency) can be stored as chunks containing the internal structure. This is not possible with SBCG, since phrasal signs never contain internal structures. For a detailed discussion of Sign-Based Construction Grammar see Section 1.3.2 and Müller (2016: Section 10.6.2).

Embodied Construction Grammar (Bergen & Chang 2005) uses typed feature descriptions for the description of linguistic objects and allows for discontinuous constituents. As argued by Müller (2016: Section 10.6.3), it is a notational variant of Reape-style HPSG (Reape 1994; see also Müller 2021: Section 6, Chapter 10 of this volume for discontinuous constituents in HPSG).

Fluid Construction Grammar is also rather similar to HPSG. An important difference is that FCG attaches weights to constraints, something that is usually not done in HPSG. But in principle, there is nothing that forbids adding weights to HPSG as well, and in fact it has been done (Brew 1995; Briscoe & Copestake 1999; Miyao & Tsujii 2008), and it should be done to a larger extent (Miller 2013). Van Trijp (2013) tried to show that Fluid Construction Grammar is fundamentally different from SBCG, but I think he failed in every single respect. See Müller (2017b) for a detailed discussion, which cannot be repeated here for space reasons.

In what follows I will compare Constructional HPSG (as assumed in this volume) with SBCG.

---

13The two approaches will be discussed in more detail in Section 1.3.1 and Section 1.3.2.
1.3.1 Constructional HPSG

As is discussed in other chapters in more detail (Richter 2021: Section 2; Abeillé & Borsley 2021: Section 3), HPSG uses feature value pairs to model linguistic objects. One important tool is structure sharing. For example, determiner, adjective, and noun agree with respect to certain features in languages like German. The identity of properties is modeled by identity of feature values and this identity is established by identifying the values in descriptions. Now, it is obvious that certain features are always shared simultaneously. In order to facilitate the statement of respective constraints, feature value pairs are put into groups. This is why HPSG feature descriptions are very complex. Information about syntax and semantics is represented under SYNTAX-SEMANTICS (SYNSEM), information about syntax under CATEGORY (CAT), and information that is projected along the head path of a projection is represented under HEAD. All feature structures have to have a type. The type may be omitted in the description, but there has to be one in the model. Types are organized in hierarchies. (7) shows an example lexical item for the word *ate*:

(7) Lexical item for the word *ate*:

\[
\begin{aligned}
\text{word} & \quad \text{PHONOLOGY (ate)} \\
\text{SYNTAX-SEMANTICS} & \quad \text{local} \\
\text{CATEGORY} & \quad \text{HEAD} \\
\text{SPR} & \quad \text{NP[nom]} \\
\text{COMPS} & \quad \text{NP[acc]} \\
\text{CONTENT} & \quad \text{ACTOR} \\
\end{aligned}
\]

14The first ‘…’ stands for the feature LOCAL, which is irrelevant in the present discussion. It plays a role in the treatment of nonlocal dependencies (Borsley & Crysmann 2021, Chapter 13 of this volume).

15To keep things simple, I omitted the feature ARG-ST here. ARG-ST stands for argument structure. The value of ARG-ST is a list containing all arguments, that is, the elements of SPR and COMPS are also contained in the ARG-ST. Linking constraints are formulated with respect to the argument structure list. See Wechsler, Koenig & Davis (2021), Chapter 9 of this volume for a discussion of linking. The way arguments are linked to the valence features SPR and COMPS is language- or language-class-specific. See Chapter 9 and also Müller (2021: Section 4), Chapter 10 of this volume.
The information about part of speech and finiteness is bundled under head. The selection of a subject is represented under spr (sometimes the feature subj is used for subjects) and the non-subject arguments are represented as part of a list under comps. The semantic indices 1 and 2 are linked to thematic roles in the semantic representation (for more on linking, see Wechsler, Koenig & Davis 2021, Chapter 9 of this volume).

Dominance structures can also be represented with feature value pairs. While Pollard & Sag (1987) and Pollard & Sag (1994) had a daughters feature and then certain phrasal types constraining the daughters within the daughters feature, Sag (1997) represented the daughters and constraints upon them at the top level of the sign.16 This move made it possible to have subtypes of the type phrase, e.g., filler-head-phrase, specifier-head-phrase, and head-complement-phrase. Generalizations over these types can now be captured within the type hierarchy together with other types for linguistic objects like lexical items and lexical rules (see Section 1.2.3). (8) shows an implicational constraint on the type head-complement-phrase.17

(8) Head-Complement Schema adapted from Sag (1997: 479):

$$\text{head-complement-phrase} \Rightarrow [\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{COMPS} \langle \rangle, \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{COMPS} \langle 1, \ldots, 2 \rangle, \text{NON-HEAD-DTRS} \langle [\text{SYNSEM} 1], \ldots, [\text{SYNSEM} 2] \rangle]$$

The constraint says that feature structures of type head-complement-phrase have to have a SYNSEM value with an empty COMPS list, a HEAD-DTR feature, and a list-valued NON-HEAD-DTRS feature. The list has to contain elements whose SYNSEM values are identical to respective elements of the COMPS list of the head daughter (1, ..., 2).

Dominance schemata (corresponding to grammar rules in phrase structure grammars) refer to such phrasal types. (9) shows how the lexical item in (7) can be used in a head-complement configuration:

16 The top level is the outermost level. So in (7), phonology and syntax-semantics are at the top level.
17 The schema in (8) licenses flat structures. See Müller (2021: xi), Chapter 10 of this volume for binary branching structures.
Analysis of *ate a pizza* in Constructional HPSG:

The description in the `comps` list of the head is identified with the `synsem` value of the non-head daughter (3). The information about the missing specifier is represented at the mother node (2). Head information is also shared between head daughter and mother node. The respective structure sharings are enforced by principles: the Subcategorization Principle or, in more recent versions of HPSG, the Valence Principle makes sure that all valents of the head daughter that are not realized in a certain configuration are still present at the mother node. The Head Feature Principle ensures that the head information of a head daughter in headed structures is identical to the head information on the mother node, that is, head features are shared.

This is a very brief sketch of Constructional HPSG and is by no means intended to be a full-blown introduction to HPSG, but it provides a description of properties that can be used to compare Constructional HPSG to Sign-Based Construction Grammar in the next subsection.

### 1.3.2 Sign-Based Construction Grammar

Having discussed some aspects of Constructional HPSG, I now turn to SBCG. SBCG is an HPSG variant, so it shares most properties of HPSG but there are
some interesting properties that are discussed in this section. Locality constraints are discussed in the next subsection, and changes in feature geometry in the subsections to follow. Subsection 1.3.2.7 discusses Frame Semantics.

### 1.3.2.1 Locality constraints

As mentioned in Section 1.2.4, SBCG assumes a strong version of locality: phrasal signs do not have daughters. This is due to the fact that phrasal schemata (= phrasal constructions) are defined as in (10):

\[
\text{head-comp-cx} \rightarrow \\
\begin{cases}
\text{MOTHER|SYN|VAL|COMPS} \langle \rangle \\
\text{HEAD-DTR} [\text{word} \text{SYN|VAL|COMPS} A] \\
\text{DTRS} \langle \emptyset \rangle \oplus A \text{nelist}
\end{cases}
\]

Rather than specifying syntactic and semantic properties of the complete linguistic object at the top level (as earlier versions of HPSG did), these properties are specified as properties under mother. Hence a construction licenses a sign (a phrase or a complex word), but the sign does not include daughters. The daughters live at the level of the construction only. While earlier versions of HPSG licensed signs directly, SBCG needs a statement saying that all objects under mother are objects licensed by the grammar (Sag, Wasow & Bender 2003: 478):

\[
\Phi \text{ is a Well-Formed Structure according to a grammar } G \text{ if and only if:}
\]

1. there is a construction \(C\) in \(G\), and
2. there is a feature structure \(I\) that is an instantiation of \(C\), such that \(\Phi\) is the value of the mother feature of \(I\).

The idea behind this change in feature geometry is that heads cannot select for daughters of their valents and hence the formal setting is more restrictive and hence reducing computational complexity of the formalism (Ivan Sag, p.c. 2011). However, this restriction can be circumvented by just structure sharing an element of the daughters list with some value within mother. The xarg feature making one argument available at the top level of a projection (Bender & Flickinger 1999) is such a feature. So, at the formal level, the mother feature

---

18A less formal version of this constraint is given as the Sign Principle by Sag (2012: 105): "Every sign must be listemically or constructionally licensed, where: a sign is listemically licensed only if it satisfies some listeme, and a sign is constructionally licensed if it is the mother of some well-formed construct."
alone does not result in restrictions on complexity. One would have to forbid such structure sharings in addition, but then one could keep MOTHER out of the business and state the restriction for earlier variants of HPSG (Müller 2018b: Section 10.6.2.1.3).

Note that analyses like the one of the Big Mess Construction by Van Eynde (2018: 841), also discussed in Van Eynde (2021: xxx), Chapter 8 of this volume, cannot be directly transferred to SBCG since in the analysis of Van Eynde, this construction specifies the phrasal type of its daughters, something that is excluded by design in SBCG: all MOTHER values of phrasal constructions are of type phrase and this type does not have any subtypes (Sag 2012: 98). Daughters in syntactic constructions are of type word or phrase. So, it is impossible to require a daughter to be of type regular-nominal-phrase as in the analysis of Van Eynde. In order to capture the Big Mess Construction in SBCG, one would have to specify the properties of the daughters with respect to their features rather than specifying the types of the daughters, that is, one has to explicitly provide the features that are characteristic for feature structures of type regular-nominal-phrase in Van Eynde’s analysis rather than just naming the type. See Kay & Sag (2012) and Kim & Sells (2011) for analyses of the Big Mess Construction in SBCG.

1.3.2.2 SPR and COMPS vs. VALENCE

Sag, Wasow & Bender (2003) differentiated between specifiers and complements, but this distinction was given up in later work on SBCG. Sag (2012) has just one valence list that includes both subjects and non-subjects. This is a return to the valence representations of Pollard & Sag (1987). An argument for this was never given, despite arguments for a separation of valence information by Borsley (1987). With one single valence feature, a VP would be an unsaturated projection and generalizations concerning phrases cannot be captured. For example, a generalization concerning extraposition (in German) is that maximal projections (that is projections with an empty COMPS list) can be extraposed (Müller 1999: Section 13.1.2). It is impossible to state this generalization in SBCG in a straightforward way (Müller 2018b: Section 10.6.2.3).

1.3.2.3 The Head Feature Principle

There have been some other developments as well. Sag (2012) got rid of the Head Feature Principle and stated identity of information explicitly within constructions. Structure sharing is not stated with boxed numbers but with capital letters instead. An exclamation mark can be used to specify information that is
not shared (Sag 2012: 125). While the use of letters instead of numbers is just a presentational variant, the exclamation mark is a non-trivial extension. (12) provides an example: the constraints on the type \textit{pred-hd-comp-cxt}:

\begin{equation}
\text{pred-hd-comp-cxt} \Rightarrow \\
\begin{array}{l}
\text{MOTHER|SYN X ! [VAL } \langle Y \rangle ] \\
\text{HEAD-DTR Z: } \begin{array}{l}
\text{word} \\
\text{SYN X: [VAL } \langle Y \rangle \oplus L]\end{array} \\
\text{DTRS} \end{array} \oplus L: \text{nelist}
\end{equation}

The X stands for all syntactic properties of the head daughter. These are identified with the value of SYM of the mother with the exception of the VAL value, which is specified to be a list with the element Y. It is interesting to note that the !-notation is not without problems: Sag (2012: 145) states that the version of SBCG that he presents is “purely monotonic (non-default)”, but if the SYM value of the mother is not identical due to overwriting of VAL, it is unclear how the type of SYM can be constrained. ! can be understood as explicitly sharing all features that are not mentioned after the !. Note, though, that the type has to be shared as well. This is not trivial, since structure sharing cannot be applied here, since structure sharing the type would also identify all features belonging to the respective value. So one would need a relation that singles out a type of a structure and identifies this type with the value of another structure. Note also that information from features behind the ! can make the type of the complete structure more specific. Does this affect the shared structure (e.g., HEAD-DTR|SYM in (12))? What if the type of the complete structure is incompatible with the features in this structure? What seems to be a harmless notational device in fact involves some non-trivial machinery in the background. Keeping the Head Feature Principle makes this additional machinery unnecessary.

1.3.2.4 Feature geometry and the FORM feature

The phrasal sign for \textit{ate a pizza} in Constructional HPSG was given in (9). (13) is the Predicational Head Complement Construction with daughters and mother filled in.
As was explained in the previous subsection, Constructional HPSG groups all selectable information under SYNSEM and then differentiates into CAT and CONT. SBCG goes back to Pollard & Sag (1987) and uses SYN and SEM. The idea behind SYNSEM was to exclude the selection of phonological information and daughters (Pollard & Sag 1994: 23). Since daughters are outside of the definition of synsem, they cannot be accessed from within valence lists. Now, SBCG pushes this idea one step further and also restricts the access to daughters in phrasal schemata (constructions in SBCG terminology): since signs do not have daughters, constructions may not refer to the daughters of their parts. But obviously signs need to have a form part, since signs are per definition form-meaning pairs. It follows that the form part of signs is selectable in SBCG. This will be discussed in more detail in the following subsection. Subsection 1.3.2.6 discusses the omission of the LOCAL feature.

1.3.2.5 Selection of PHON and FORM values

The feature geometry of Constructional HPSG has the PHON value outside of SYNSEM. Therefore verbs can select for syntactic and semantic properties of their arguments but not for their phonology. For example, they can require that an object has accusative case but not that it starts with a vowel. SBCG allows for the selection of phonological information (the feature is called FORM here) and one example of such a selection is the indefinite article in English, which has to be either a or an depending on whether the noun or nominal projection it is combined with starts with a vowel or not (Flickinger, Mail to the HPSG mailing list, 01.03.2016):
(14)  a. an institute  
    b. a house

The distinction can be modeled by assuming a selection feature for determiners.\footnote{In the 1994 version of HPSG there is mutual selection between the determiner and the noun. The noun selects the determiner via \textit{spr} and the determiner selects the noun via a feature called \textit{specified} (Pollard & Sag 1994: 45–54).} An alternative would be, of course, to capture all phonological phenomena by formulating constraints on phonology at the phrasal level (see Bird & Klein 1994, citealpHoehle99a-u, and Walther 1999 for phonology in HPSG).

Note also that the treatment of raising in SBCG admits nonlocal selection of phonology values, since the analysis of raising in SBCG assumes that the element on the valence list of the embedded verb is identical to an element in the \textit{arg-st} list of the matrix verb (Sag 2012: 159). Hence, both verbs in (15) can see the phonology of the subject:

(15) Kim can eat apples.

In principle, there could be languages in which the form of the downstairs verb depends on the presence of an initial consonant in the phonology of the subject. English allows for long chains of raising verbs and one could imagine languages in which all the verbs on the way are sensitive to the phonology of the subject. Such languages probably do not exist.

Now, is this a problem? Not really, but if one develops a general setup in a way to exclude everything that is not attested in the languages of the world (as for instance the selection of arguments of arguments of arguments), then it is a problem that heads can see the phonology of elements that are far away.

There are two possible conclusions for practitioners of Sign-Based Construction Grammar: either the \textit{mother} feature could be given up, since one agrees that theories that do not make wrong predictions are sufficiently constrained and one does not have to explicitly state what cannot occur in languages, or one would have to address the problem with nonlocally selected phonology values and therefore assume a \textit{synsem} or \textit{local} feature that bundles information that is relevant in raising and does not include the phonology. In the latter case, the feature geometry of SBCG would get more complicated. This additional complication is further evidence against \textit{mother}, adding to the argument I made about \textit{mother} in Subsection 1.3.2.1.
1.3.2.6 The local feature and information shared in nonlocal dependencies

Similarly, elements of the ARG-ST list contain information about FORM. In nonlocal dependencies, this information is shared in the GAP list (SLASH set or list in other versions of HPSG) and is available all the way to the filler (Sag 2012: Section 10). In other versions of HPSG, only local information is shared and elements in valence lists do not have a PHON feature. If the sign that is contained in the GAP list were identified with the filler, the information about phonological properties of the filler would be available at the extraction site and SBCG could be used to model languages in which the phonology of a filler is relevant for a head from which it is extracted. So for instance, *likes* could see the phonology of *bagels* in (16):

(16) Bagels, I think that Peter likes.

It would be possible to state constraints saying that the filler has to contain a vowel or two vowels or that it ends with a consonant. In addition, all elements on the extraction path (*that* and *think*) can see the phonology of the filler as well. While there are languages that mark the extraction path (Bouma et al. 2001: 4–5; Borsley & Crysmand 2021: xiv–xv, Chapter 13 of this volume), I doubt that there are languages that have phonological effects over unbounded dependencies. This problem can be and has been solved by assuming that the filler is not shared with the information in the GAP list, but parts of the filler are shared with parts in the GAP list: Sag (2012: 166) assumes that SYN, SEM, and STORE information is identified individually. Originally, the feature geometry of HPSG was motivated by the wish to structure share information. Everything within local was shared between filler and extraction site. This kind of motivation is given up in SBCG.

Note, also, that not sharing the complete filler with the gap means that the FORM value of the element in the ARG-ST list at the extraction site is not constrained. Without any constraints, the theory would be compatible with infinitely many models, since the FORM value could be anything. For example, the FORM value of an extracted adjective could be ⟨Donald Duck⟩ or ⟨Dunald Dock⟩ or any arbitrary chaotic sequence of letters/phonemes. To exclude this, one can stipulate the FORM values of extracted elements to be the empty list, but this leaves one with the unintuitive situation that the element in GAP has an empty FORM list while the corresponding filler has a different, filled one.

See also Borsley & Crysmand (2021: Section 10), Chapter 13 of this volume for a comparison of the treatment of unbounded dependencies in Constructional HPSG and SBCG.
1.3.2.7 Frame Semantics

Another difference between SBCG and other variants of HPSG is the use of Frame Semantics (Fillmore 1982; Fillmore 1985a). The actual representations in SBCG are based on MRS (Minimal Recursion Semantics, Copestake et al. 2005, see also Koenig & Richter 2021, Chapter 22 of this volume) and the change seems rather cosmetic (relations have the suffix -fr for frame rather than -rel for relation and the feature is called FRAMES rather than RELATIONS), but there is one crucial difference: the labels of semantic roles are more specific than what is usually used in other variants of HPSG.\(^{20}\) Sag (2012: 89) provides the following representation for the meaning contribution of the verb *eat*:

\[
\begin{array}{c}
\text{sem-obj} \\
\text{INDEX s} \\
\text{FRAMES} \\
\end{array}
\begin{array}{c}
\text{eating-fr} \\
\text{LABEL l} \\
\text{SIT s} \\
\text{INGESTOR i} \\
\text{INGESTIBLE j} \\
\end{array}
\]

While some generalizations over verbs of a certain type can be captured with role labels like INGESTOR and INGESTIBLE, this is limited to verbs of ingestion. More general role labels like AGENT and PATIENT (or PROTO-AGENT and PROTO-PATIENT, Dowty 1991, or ACTOR and UNDERGOER Van Valin 1999) allow for more generalizations of broader classes of verbs (see Davis & Koenig 2000, Davis 2001: Section 4.2.1, and Wechsler, Koenig & Davis 2021, Chapter 9 of this volume).

1.3.3 Summary

This section enumerated various flavors of Construction Grammars and briefly discussed the more formal variants. It was noted that the formal underpinnings are rather similar in many cases. What is different, though, is the kind of approach taken towards the representation of valence and argument structure constructions. Constructional HPSG and SBCG differ from other Construction Grammars in taking a strongly lexicalist stance (Sag & Wasow 2011: Section 10.4.3; Wasow 2021: Section 3.4, Chapter 24 of this volume): argument structure is encoded lexically. A ditransitive verb is a ditransitive verb since it selects for three NP

\(^{20}\)Pollard & Sag (1987: 95) and Pollard & Sag (1994) use role labels like KISSEr and KISSEE that are predicate-specific. Generalizations over these feature names are impossible within the standard formal setting of HPSG (Pollard & Sag 1994: Section 8.5.3; Müller 1999: 24, Fn. 1; Davis 2001: Section 4.2.1).
arguments. This selection is encoded in valence features of lexical items. It is not assumed that phrasal configurations can license additional arguments as it is in basically all other variants of Construction Grammar. The next section discusses phrasal CxG approaches in more detail. Section 4 then discusses patterns that should be analyzed phrasally and which are problematic for entirely head-driven (or rather functor-driven) theories like Categorial Grammar, Dependency Grammar, and Minimalism.

2 Valence vs. phrasal patterns

Much work in Construction Grammar starts from the observation that children acquire patterns and, in later acquisition stages, abstract from these patterns to schemata containing open slots to be filled by variable material, for example subjects and objects (Tomasello 2003). The conclusion that is drawn from this is that language should be described with reference to phrasal patterns. Most Construction Grammar variants assume a phrasal approach to argument structure constructions (Goldberg 1995; 2006; Goldberg & Jackendoff 2004), with Constructional HPSG (Sag 1997), Boas’s (2003) work, and SBCG (Sag, Boas & Kay 2012; Sag 2012) being the three exceptions. So, for examples like the resultative construction in (18), Goldberg (1995: Chapter 8) assumes that there is a phrasal construction [Subj [V Obj Obl]] into which material is inserted and which contributes the resultative semantics as a whole.

(18) She fished the pond empty.

HPSG follows the lexical approach and assumes that fish- is inserted into a lexical construction (lexical rule), which licenses the combination with other parts of the resultative construction (Müller 2002: Section 5.2).

I argued in several publications that the language acquisition facts can be explained in lexical models as well (Müller 2010: Section 6.3; Müller & Wechsler 2014a: Section 9). While a pattern-based approach claims that (19) is analyzed by inserting Kim, loves, and Sandy into a phrasal schema stating that NP[nom] verb NP[acc] or subject verb object are possible sequences in English, a lexical approach would state that there is a verb loves selecting for an NP[nom] and an NP[acc] (or for a subject and an object).

(19) Kim loves Sandy.

21Goldberg does not state the resultative construction, but the Caused-Motion Construction, which is syntactically parallel to the Resultative Construction, is specified this way on p. 152. She describes the syntax of resultative constructions on p. 192.
Since objects follow the verb in English (modulo extraction) and subjects precede the verb, the same sequence is licensed in the lexical approach. The lexical approach does not have any problems accounting for patterns in which the sequence of subject, verb, and object is discontinuous. For example, an adverb may intervene between subject and verb:

(20) Kim really loves Sandy.

In a lexical approach it is assumed that verb and object may form a unit (a VP). The adverb attaches to this VP and the resulting VP is combined with the subject. The phrasal approach has to assume either that adverbs are part of phrasal schemata licensing cases like (20) (see Uszkoreit 1987: Section 6.3.2 for such a proposal in a GPSG approach to German) or that the phrasal construction may license discontinuous patterns. Bergen & Chang (2005: 170) follow the latter approach and assume that subject and verb may be discontinuous but verb and object(s) have to be adjacent. While this accounts for adverbs like the one in (20), it does not solve the general problem, since there are other examples showing that verb and object(s) may appear discontinuously as well:

(21) Mary tossed me a juice and Peter a water.

Even though *tossed* and *Peter a water* are discontinuous in (21), they are an instance of the ditransitive construction. The conclusion is that what has to be acquired is not a phrasal pattern but rather the fact that there are dependencies between certain elements in phrases (see also Behrens 2009 for a similar view from a language acquisition perspective). I return to ditransitive constructions in Section 2.3.

I discussed several phrasal approaches to argument structure and showed where they fail (Müller 2006a,b; 2007; 2010; Müller & Wechsler 2014a,b; Müller 2018a). Of course, the discussion cannot be reproduced here, but I want to repeat four points showing that lexical valence representation is necessary and that effects that are the highlight of phrasal approaches can be achieved in lexical proposals as well. The first two are problems that were around in GPSG times and basically were solved by abandoning the framework and adopting a new framework which was a fusion of GPSG and Categorial Grammar: HPSG.\footnote{For further criticism of GPSG see Jacobson (1987). A detailed discussion of reasons for abandoning GPSG can be found in Müller (2016: Section 5.5).}

### 2.1 Derivational morphology and valence

The first argument (Müller 2016: Section 5.5.1) is that certain patterns in derivational morphology refer to valence. For example, the *-bar ‘able’ derivation pro-
ductively applies to transitive verbs only, that is, to verbs that govern an accusative.

(22)  

a. unterstützbar  
    supportable  

b. * helfbar  
    helpable  

c. * schlafbar  
    sleepable  

Note that the -bar ‘able’ derivation is like the passive in that it suppresses the subject and promotes the accusative object: the accusative object is the element adjectives derived with the -bar ‘able’ derivation predicate over. There is no argument realized with the adjective unterstützbaren ‘supportable’ attaching to Arbeitsprozessen ‘work.processes’ in unterstützbaren Arbeitsprozessen. Hence one could not claim that the stem enters a phrasal construction with arguments and -bar attaches to this phrase. It follows that information about valence has to be present in stems.

Note also that the resultative construction interacts with the -bar ‘able’ derivation. (23) shows an example of the resultative construction in German in which the accusative object is introduced by the construction: it is the subject of leer ‘empty’ but not a semantic argument of the verb fischt ‘fishes’.

(23)   Sie fischt den Teich leer.  
            she fishes the  pond empty  

So even though the accusative object is not a semantic argument of the verb, the -bar ‘able’ derivation is possible and an adjective like leerfischbar ‘empty.fishable’ meaning ‘can be fished empty’ can be derived. This is explained by lexical analyses of the -bar ‘able’ derivation and the resultative construction, since if one assumes that there is a lexical item for the verb fisch- selecting an accusative object and a result predicate, then this item may function as the input for the -bar ‘able’ derivation. See Section 3 for further discussion of -bar ‘able’ derivation.

\[23\] Adjectives realize their arguments preverbally in German:

(i)   der [seiner Frau treue]  
      the  [his  wife faithful  
      ‘the man who is faithful to his wife’  

unterstützbaren ‘supportable’ does not take an argument; it is a complete adjectival projection like seiner Frau treue.
and Verspoor (1997), Wechsler (1997), Wechsler & Noh (2001), and Müller (2002: Chapter 5) for lexical analyses of the resultative construction in the framework of HPSG.

2.2 Partial verb phrase fronting

The second argument concerns partial verb phrase fronting (Müller 2016: Section 5.5.2). (24) gives some examples: in (24a) the bare verb is fronted and its arguments are realized to the right of the finite verb in the so-called middle field, in (24b) one of the objects is fronted together with the verb, and in (24c) both objects are fronted with the verb.

(24)  
   a. Erzählen wird er seiner Tochter ein Märchen können.  
       tell will he his daughter a fairy.tale can  
   b. Ein Märchen erzählen wird er seiner Tochter können.  
       a fairy.tale tell will he his daughter can  
   c. Seiner Tochter ein Märchen erzählen wird er können.  
       his daughter a fairy.tale tell will he can  
       ‘He will be able to tell his daughter a fairy tale.’

The problem with sentences such as those in (24) is that the valence requirements of the verb *erzählen* ‘to tell’ are realized in various positions in the sentence. For fronted constituents, one requires a rule which allows a ditransitive to be realized without its arguments or with one or two objects. This basically destroys the idea of a fixed phrasal configuration for the ditransitive construction and points again in the direction of dependencies.

Furthermore, it has to be ensured that the arguments that are missing in the prefield are realized in the remainder of the clause. It is not legitimate to omit obligatory arguments or realize arguments with other properties like a different case, as the examples in (25) show:

(25)  
   a. Verschlungen hat er ihn nicht.  
       devoured has he.NOM him.ACC not  
       ‘He did not devour it/him.’  
   b. * Verschlungen hat er nicht.  
       devoured has he.NOM not  
   c. * Verschlungen hat er ihm nicht.  
       devoured has he.NOM him.DAT not
Stefan Müller

The obvious generalization is that the fronted and unfronted arguments must add up to the total set of arguments selected by the verb. This is scarcely possible with the rule-based representation of valence in GPSG (Nerbonne 1986; Johnson 1986). In theories such as Categorial Grammar, it is possible to formulate elegant analyses of (25) (Geach 1970). Nerbonne (1986) and Johnson (1986) both suggest analyses for sentences such as (25) in the framework of GPSG which ultimately amount to changing the representation of valence information in the direction of Categorial Grammar. With a switch to CG-like valence representations in HPSG, the phenomenon of partial verb phrase fronting found elegant solutions (Höhle 2018a: Section 4; Müller 1996; Meurers 1999).

2.3 Coercion

An important observation in constructionist work is that, in certain cases, verbs can be used in constructions that differ from the constructions they are normally used in. For example, verbs that are usually used with one or two arguments may be used in the ditransitive construction:

(26)  a. She smiled.
     b. She smiled herself an upgrade.  
     c. He baked a cake.
     d. He baked her a cake.

The usual explanation for sentences like (26b) and (26d) is that there is a phrasal pattern with three arguments into which intransitive and strictly transitive verbs may enter. It is assumed that the phrasal patterns are associated with a certain meaning (Goldberg 1996; Goldberg & Jackendoff 2004). For example, the benefactive meaning of (26d) is contributed by the phrasal pattern (Goldberg 1996: Section 6; Asudeh, Giorgolo & Toivonen 2014: 81).

The insight that a verb is used in the ditransitive pattern and thereby contributes a certain meaning is of course also captured in lexical approaches. Briscoe & Copestake (1999: Section 5) suggested a lexical rule-based analysis mapping a transitive version of verbs like bake onto a ditransitive one and adding the benefactive semantics. This is parallel to the phrasal approach in that it says: three-place bake behaves like other three-place verbs (e.g., give) in taking three arguments and by doing so, it comes with a certain meaning (see Müller 2018a for a lexical rule-based analysis of the benefactive constructions that works for both

English and German, despite the surface differences of the respective languages. The lexical rule is a form-meaning pair and hence a construction. As Croft put it 18 years ago: lexical rule vs. phrasal schema is a false dichotomy (Croft 2003). But see Müller (2018a; 2006a; 2013) and Müller & Wechsler (2014a) for differences between the approaches.

Briscoe & Copestake (1999) paired their lexical rules with probabilities to be able to explain differences in productivity. This corresponds to the association strength that van Trijp (2011: 141) used in Fluid Construction Grammar to relate lexical items to phrasal constructions of various kinds.

2.4 Non-predictability of valence

The last subsection discussed phrasal models of coercion that assume that verbs can be inserted into constructions that are compatible with the semantic contribution of the verb. Müller & Wechsler (2014a: Section 7.4) pointed out that this is not sufficiently constrained. Müller & Wechsler discussed the examples in (27), among others:

(27) a. John depends on Mary. (counts, relies, etc.)
    b. John trusts (*on) Mary.

While depends can be combined with a on-PP, this is impossible for trusts. Also the form of the preposition of prepositional objects is not always predictable from semantic properties of the verb. So there has to be a way to state that certain verbs go together with certain kinds of arguments and others do not. A lexical specification of valence information is the most direct way to do this. Phrasal approaches sometimes assume other means to establish connections between lexical items and phrasal constructions. For instance, Goldberg (1995: 50) assumes that verbs are “conventionally associated with constructions”. The more technical work in Fluid CxG assumes that every lexical item is connected to various phrasal constructions via coapplication links (van Trijp 2011: 141). This is very similar to Lexicalized Tree Adjoining Grammar (LTAG; Schabes, Abeillé & Joshi 1988), where a rich syntactic structure is associated to a lexical anchor. So, phrasal approaches that link syntactic structure to lexical items are actually lexical approaches as well. As in GPSG, they include means to ensure that lexical items enter into correct constructions. In GPSG, this was taken care of by a number. I already discussed the GPSG shortcomings in previous subsections.

Concluding this section, it can be said that there has to be a connection between lexical items and their arguments and that a lexical representation of argument structure is the best way to establish such a relation.
3 Construction Morphology

The first publications in Construction Morphology were the master’s thesis of Riehemann (1993), which later appeared as Riehemann (1998), and Koenig’s 1994 WCCFL paper and thesis (Koenig & Jurafsky 1995; Koenig 1994; 1999). Riehemann called her framework Type-Based Derivational Morphology, since it was written before influential work like Goldberg (1995) appeared and before the term Construction Morphology (Booij 2005) was used. Riehemann did a careful corpus study on adjective derivations with the suffix `-bar` `-able`. She noticed that there is a productive pattern that can be analyzed by a lexical rule relating a verbal stem to the adjective suffixed with `-bar`. The productive pattern applies to verbs governing an accusative as in (28a) but is incompatible with verbs taking a dative as in (28b):

(28)  

a. unterstüzbar  
    supportable
b. * helfbar  
    helpable
c. * schlafbar  
    sleepable

Intransitive verbs are also excluded, as (28c) shows. Riehemann suggests a schema like the one in (29):

---

She did not call her rule a lexical rule, but the difference between her template and the formalization of lexical rules by Müller (2002: 26) is the naming of the feature MORPH-B vs. LEX-DTR. Copestake & Briscoe (1992: Section 8.2.3), Briscoe & Copestake (1999: Section 2), and Meurers (2001: 176) use a representation with IN and OUT features that actually corresponds to the MOTHER/DTRS format of SBCG. See Section 1.3.2.1.
Schema for productive adjective derivations with the suffix -\textit{bar} in German adapted from Riehemann (1998: 68):

\[
\begin{array}{c}
\text{reg-bar-adj} \\
\text{PHON} \mathbb{1} \oplus \langle \text{bar} \rangle \\
\text{trans-verb} \\
\text{PHON} \mathbb{1} \\
\text{MORPH-B} \\
\text{SYNSEM|LOC} \\
\text{CAT|COMPS} \langle \text{NP} \text{[acc]}^{[2]} \oplus \mathbb{5} \rangle \\
\text{CONT|NUC} \mathbb{4} \\
\text{RELN} \circ \\
\text{UND} \mathbb{2} \\
\text{SOA-ARG} \mathbb{1} \\
\text{SYNSEM|LOC} \\
\text{CAT} \\
\text{adj} \\
\text{SUBJ} \langle \text{NP}^{[2]} \rangle \\
\text{COMPS} \mathbb{3} \\
\text{CONT|NUC} \\
\end{array}
\]

\text{MORPH-B} is a list that contains a description of a transitive verb (something that governs an accusative object which is linked to the undergoer role (\textbf{2}) and has an actor). The phonology of this element (\textbf{1}) is combined with the suffix -\textit{bar} and forms the phonology of the complete lexical item. The resulting object is of category \textit{adj} and the index of the accusative object of the input verb (\textbf{2}) is identified with the one of the subject of the resulting adjective and with the

Note that the specification of the type \textit{trans-verb} in the list under \text{MORPH-B} is redundant, since it is stated that there has to be an accusative object and that there is an actor and an undergoer in the semantics. Depending on further properties of the grammar, the specification of the type is actually wrong: productively derived particle verbs may be input to the -\textit{bar} `able' derivation, and these are not a subtype of \textit{trans-verb}, since the respective particle verb rule derives both transitive (\textit{anlachen} `laugh at somebody') and intransitive verbs (\textit{loslachen} `start to laugh') (Müller 2003b: 296). \textit{Anlachen} does not have an undergoer in the semantic representation suggested by Stiebels (1996). See Müller (2003b: 308) for a version of the -\textit{bar} `able' derivation schema that is compatible with particle verb formations as input.

The original formulation of Riehemann shares the \textit{CONT} value of the semantics of the accusative NP with the subject of the adjective and the value of the \textit{UNDERGOER} feature. I adapted the rule here to just share the index, since values of \textit{ACTOR} and \textit{UNDERGOER} features are of type \textit{index}. Jean-Pierre Koenig pointed out to me that sharing of the whole content of the accusative object and the subject of the adjective is necessary, since otherwise the \textit{CONT} value of the accusative object would be unrestricted and -- according to the formal basics of HPSG -- could vary in infinitely many ways. Such an explicit sharing of semantics is not necessary in Müller’s approach, since he distinguishes between structural and lexical case (Przepiórkowski 2021: Section 2, Chapter 7 of this volume) and this makes it possible to structure share the complete description of the accusative object with the subject of the adjective.
value of the UNDERGOER feature in the semantic representations of the adjective. The semantics of the input verb \( \square \) is embedded under a modal operator in the semantics of the adjective.

While the description of the -bar ‘able’ derivation given so far captures the situation quite well, there are niches and isolated items that are exceptions. According to Riehemann (1998: 5), this was the case for 7% of the adjectives she looked at in her corpus study. Examples are verbs ending in -ig like entschuldigen ‘to excuse’. The -ig is dropped in the derivation:

(30) entschuldbar
     excuseable

Other cases are lexicalized forms like essbar ‘safely edible’, which have a special lexicalized meaning. Exceptions of the accusative requirement are verbs selecting a dative (31a), a prepositional object (31b), reflexive verbs (31c), and even intransitive, mono-valet verbs (31d):

(31) a. unentrinnbar
     inescapable
b. verfügbar
     available
c. regenerierbar
     regenerable
d. brennbar
     inflammable

To capture generalizations about productive, semi-productive and fixed patterns/items, Riehemann suggests a type hierarchy, parts of which are provided in Figure 32.2. The type bar-adj stands for all -bar adjectives and comes with the constraints that apply to all of them. One subtype of this general type is trans-bar-adj, which subsumes all adjectives that are derived from transitive verbs. This

![Figure 32.2: Parts of the type hierarchy for -bar ‘able’ derivation adapted from Riehemann (1998: 15)](image-url)
includes all regularly derived -bar-adjectives, which are of the type reg-bar-adj but also essbar ‘edible’ and sichtbar ‘visible’.

As this recapitulation of Riehemann’s proposal shows, the analysis is a typical CxG analysis: V-bar is a partially filled word (see Goldberg’s examples in Table 32.1). The schema in (29) is a form-meaning pair. Exceptions and subregularities are represented in an inheritance network.

4 Phrasal patterns

Section 2 discussed the claim that constructions in the sense of CxG have to be phrasal. I showed that this is not true and that in fact lexical approaches to valence have to be preferred under the assumptions usually made in non-transformational theories. However, there are other areas of grammar that give exclusively head-driven approaches like Categorial Grammar, Minimalism, and Dependency Grammar a hard time. In what follows I discuss the NPN construction and various forms of filler gap constructions.

4.1 The NPN Construction

Matsuyama (2004) and Jackendoff (2008) discuss the NPN Construction, examples of which are provided in (32):

(32) a. Student after student left the room.
    b. Day after day after day went by, but I never found the courage to talk to her. (Bargmann 2015)

The properties of the NPN construction (with after) are summarized by Bargmann (2015) in a concise way and I will repeat his examples and summarization below to motivate his analysis in (40).

The examples in (32) show that the N-after-N Construction has the distribution of NPs. As (33) shows, the construction is partially lexically fixed: after cannot be replaced by any other word (Matsuyama 2004: 73).

(33) Alex asked me question { after / * following / * succeeding } question.

The construction is partially lexically flexible: the choice of Ns is free, except for the fact that the Ns must be identical (34a), the Ns must be count nouns (34b), Ns must be in the singular (34c), and the Ns must be bare (34d).

(34) a. * bus after car  \hspace{1cm} \text{ (N1 \neq N2) } \\
    b. * water after water  \hspace{1cm} \text{ (Ns = mass nouns) }
The construction is *syntactically fixed*: N-after-N cannot be split by syntactic operations as the contrast in (35) shows (Matsuyama 2004):

(35)  
   a. Man after man passed by.  
   b. *Man passed by after man.

If extraposition of the after-N constituent were possible, (35b) with an extraposed after man should be fine but it is not, so NPN seems to be a fixed configuration. There is a syntax-semantics mismatch: while N-after-N is syntactically singular, as (36) shows, it is plural semantically, as (37) shows:

(36) Study after study { reveals / *reveal } the dangers of lightly trafficked streets.

(37)  
   a. John ate { apple after apple / apples / *an apple } for an hour.  
   b. John ate { *apple after apple / *apples / an apple } in an hour.

Furthermore there is an aspect of semantic sequentiality: N-after-N conveys a temporal or spatial sequence: as Bargmann (2015) states, the meaning of (38a) is something like (38b).

(38)  
   a. Man after man passed by.  
   b. First one man passed by, then another(, then another(, then another(, then ... ))).

The Ns in the construction do not refer to one individual each; rather, they contribute to a holistic meaning.

The NPN construction allows adjectives to be combined with the nouns, but this is restricted. N1 can only be preceded by an adjective if N2 is preceded by the same adjective:

(39)  
   a. bad day after bad day (N1 and N2 are preceded by the same adjective.)  
   b. *bad day after awful day (N1 and N2 are preceded by different adjectives.)  
   c. *bad day after day (Only N1 is preceded by an adjective.)  
   d. day after bad day (Only N2 is preceded by an adjective.)
Finally, after N may be iterated to emphasize the fact that there are several referents of N, as the example in (32b) shows.

This empirical description is covered by the following phrasal construction, which is adapted from Bargmann (2015).

(40) NPN Construction as formalized by Bargmann (2015):

There is a list of daughters consisting of a first daughter and an arbitrarily long list of after N pairs. The ‘+’ means that there has to be at least one after N pair. The nominal daughters select for a determiner via SPR, so they can be either bare nouns or nouns modified by adjectives. The semantic representation, non-standardly represented as the value of SR, says that there have to be several objects in a set X (\(\exists X. |X| > 1\)) and for all of them, the meaning of the \(\overline{N}\) has to hold (\(\forall x \in X; N'(x)\)). Furthermore there is an order between the elements of X as stated by \(\exists Rorder \subseteq X^2\).

From looking at this construction, it is clear that it cannot be accounted for by standard X rules. Even without requiring X syntactic rules, there seems to be no

---

27Jackendoff and Bargmann assume that the result of combining N, P, and N is an NP. However this is potentially problematic, as Matsuyama’s example in (i) shows (Matsuyama 2004: 71):

(i) All ranks joined in hearty cheer after cheer for every member of the royal family ...

As Matsuyama points out, the reading of such examples is like the reading of old men and women in which old scopes over both men and women. This is accounted for in structures like the one indicated in (ii):

(ii) hearty [cheer after cheer]

Since adjectives attach to \(\overline{N}\)s and not to NPs, this means that NPN constructions should be \(\overline{N}\)s. Of course (ii) cannot be combined with a determiner, so one would have to assume that NPN constructions select for a determiner that has to be dropped obligatorily. Determiners are also dropped in noun phrases with mass nouns with a certain reading.

Prepublished draft of 5th February 2021, 13:48
Stefan Müller

way to capture these constructions in head-based approaches like Minimalism, Categorial Grammar, or Dependency Grammar. For simple NPN constructions, one could claim that after is the head. After would be categorized as a third-person singular mass noun and select for two Ns. It would (non-compositionally) contribute the semantics stated above. But it is unclear how the general schema with arbitrarily many repetitions of after N could be accounted for. If one assumes that day after day forms a constituent, then the first after in (41) would have to combine an N with an NPN sequence.

(41) day after [day [after day]]

This means that we would have to assume two different items for after: one for the combination of Ns and another one for the combination of N with NPN combinations. Note that an analysis of the type in (41) would have to project information about the Ns contained in the NPN construction, since this information has to be matched with the single N at the beginning. In any case, a lexical analysis would require several highly idiosyncratic lexical items (prepositions projecting nominal information and selecting items they usually do not select). It is clear that a reduplication account of the NPN construction as suggested by G. Müller (2011) does not work, since patterns with several repetitions of PN as in (41) cannot be accounted for as reduplication. G. Müller (p. 241) stated that reduplication works for word-size elements only (in German) and hence his account does not extend to the English examples given above. (42) shows an attested German example containing adjectives, which means that G. Müller’s approach is not appropriate for German either.

(42) Die beiden tauchten nämlich geradewegs wieder aus dem heimischen Legoland auf, wo sie im Wohnzimmer, schwarzen Stein um schwarzen Stein, vermeintliche Schusswaffen nachgebaut hatten.29

‘The two surfaced straightaway from their home Legoland where they had recreated alleged firearms black brick after black brick.’

28 Hudson (2021: xxx), Chapter 31 of this volume provides an analysis of the NPN Construction in the framework of Word Grammar. Since Word Grammar is a Dependency Grammar, this seems to falsify my claim, but it does not since Word Grammar is more powerful than usual Dependency Grammars. Hudson uses a network with some extra syntactic primitives that allow him to account for loops.

29 Attested example from the newspaper taz, 05.09.2018, p. 20
Travis (2003: 240) suggested a syntactic approach to the NPN construction. The trees she provides are broken and contain symbols like Spec, so the details of the analysis are unclear, but she assumes that the preposition is of category Q and Q heads are special reduplication heads. An element from inside of the complement of Q is moved to SpecQP. The analysis begs several questions: why can incomplete constituents move to SpecQP? How is the external distribution of NPN constructions accounted for? Are they QPs? Where can QPs appear? Why do some NPN constructions behave like NPs? How is the meaning of this construction accounted for? If it is assigned to a special Q, the question is: how are examples like (32b) accounted for? Are two Q heads assumed? And if so, what is their semantic contribution?

This subsection showed how a special phrasal pattern can be analyzed within HPSG. The next section will discuss filler-gap constructions, which were analyzed as instances of a single schema by Pollard & Sag (1994: 164) but which were later reconsidered and analyzed as a family of subconstructions by Sag (1997; 2010).

### 4.2 Specialized sub-constructions

HPSG took over the treatment of nonlocal dependencies from GPSG (Gazdar 1981; see also Flickinger, Pollard & Wasow 2021, Chapter 2 of this volume on the history of HPSG and Borsley & Crysmann 2021, Chapter 13 of this volume on unbounded dependencies). Pollard & Sag (1994: Chapters 4 and 5) had an analysis of topicalization constructions like (43) and an analysis of relative clauses. However, more careful examination revealed that more fine-grained distinctions have to be made. Sag (2010: 491) looked at the following examples:

(43) a. [My bagels,] she likes. (topicalized clause)
b. [What books] do they like? (wh-interrogative)
c. (the person) [who (se book)] they like (wh-relative)
d. [What a play] he wrote! (wh-exclamative)
e. [the more books] they read ... (the-clause)

As Sag shows, the fronted element is specific to the construction at hand:

(44) a. * [Which bagels] / [Who], she likes. (topicalized clause)
c. % the thing [[what] they like] (wh-relative)
A topicalized clause should not contain a *wh*-item (44a), a *wh*-interrogative should not contain a *what a* sequence appropriate for *wh*-exclamatives (44b), and so on. Furthermore, some of these constructions allow non-finite clauses and others do not:

(45)  
a. *Bagels, (for us) to like. (topicalized clause)  
b. *It’s amazing [what a dunce (for them) to talk to]. (wh-exclamative)  
c. *The harder (for them) to come, the harder (for them) to fall. (the-clause)  
d. I know how much time (*for them) to take. (wh-interrogative)  
e. The time in which (*for them) to finish. (wh-relative)

So there are differences as far as fillers and sentences from which something is extracted are concerned. Sag discussed further differences like inversion/non-inversion in the clauses out of which something is extracted. I do not repeat the full discussion here but refer the reader to the original paper.

In principle, there are several ways to model the phenomena. One could assume empty heads as Pollard & Sag (1994: Chapter 5) suggested for the treatment of relative clauses. Or one could assume empty heads as they are assumed in Minimalism: certain so-called operators have features that have to be checked and cause items with the respective properties to move (Adger 2003: 330–331). Borsley (2006) discussed potential analyses of relative clauses involving empty heads and showed that one would need a large number of such empty heads, and since there is no theory of the lexicon in Minimalism, generalizations are missed (see also Borsley & Müller 2021: Section 4.1.5, Chapter 28 of this volume). The alternative suggested by Sag (2010) is to assume a general Filler-Head Schema of the kind assumed in Pollard & Sag (1994: 164) and then define more specific sub-constructions. To take an example, the *wh*-exclamative is a filler-head structure, so it inherits everything from the more general construction, but in addition, it specifies that the filler daughter must contain a *what a* part and states the semantics that is contributed by the exclamative construction.

5 Summary

This paper summarized the properties of Construction Grammar, or rather Construction Grammars, and showed that HPSG can be seen as a Construction Grammar, since it fulfills all the tenets assumed in CxG: it is surface-based, grammatical constraints pair form and function/meaning, the grammars do not rely...
on innate domain-specific knowledge, and the grammatical knowledge is represented in inheritance hierarchies. This sets HPSG and CxG apart from other generative theories that either assume innate language-specific knowledge (Minimalism, e.g., Chomsky 2013; Kayne 1994; Cinque & Rizzi 2010) or do not assume inheritance hierarchies for all linguistic levels (e.g., LFG).

I showed why lexical analyses of argument structure should be preferred over phrasal ones and that there are other areas in grammar where phrasal analyses are superior to lexical ones. I showed that they can be covered in HPSG, while they are problematic for proposals assuming that all structures have to have a head.

Acknowledgements

I thank Anne Abeillé, Bob Borsley, Rui Chaves, and Jean-Pierre Koenig for comments on earlier versions of this chapter and for discussion in general. I thank Frank Richter for discussion of formal properties of SBCG. Thanks go to Frank Van Eynde for discussion of the Big Mess Construction in relation to SBCG.

References


Stefan Müller


xlvi Prepublished draft of 5th February 2021, 13:48


Stefan Müller


