Head-Driven Phrase Structure Grammar

to appear in András Kertész, Edith Moravcsik, and Csilla Rákosi (Eds): Current approaches to syntax – a comparative handbook. Berlin: de Gruyter Mouton

Stefan Müller & Antonio Machicao y Priemer
Humboldt Universität Berlin
Institut für deutsche Sprache und Linguistik
Dorotheenstraße 24
10117 Berlin
St.Mueller@hu-berlin.de

January 30, 2019

Abstract:

Head-Driven Phrase Structure Grammar is a constraint-based theory. It uses features and values to model linguistic objects. Values may be complex, e.g. consist of feature values pairs themselves. The paper shows that such feature value pairs together with identity of values and relations between feature values are sufficient to develop a complete linguistic theory including all linguisitc levels of discription. The paper explains the goals of researchers working in the framework and the way they deal with data and motivate their analyses.

The framework is explained with respect to an example sentence that involves the following phenomena: valence, constituent structure, adjunction/modification, raising, case assignment, nonlocal dependencies, relative clauses.

1 General remarks

Head-Driven Phrase Structure Grammar (HPSG) was developed in the 1980s by Carl Pollard and Ivan Sag. Ivan Sag was one of the developers of Generalized Phrase Structure Grammar (Gazdar, Klein, Pullum & Sag 1985), which strongly influenced HPSG. Carl Pollard worked in Categorial Grammar (Pollard 1984), from where some ideas were brought into the new framework. Categorial Grammar is an approach that puts a lot of emphasis on lexical information and by adopting a lexical view a lot of the shortcomings of GPSG (Jacobson 1987; Müller 2016: Section 5.5) were avoided. With the advent of HPSG, research on GPSG came to an end almost entirely and most researchers switched to HPSG. Since 1993 annual conferences have been held rotating between the US, Europe, and Asia. The HPSG online bibliography at `https://hpsg.hu-berlin.de/HPSG-Bib/` lists over 1500 papers published in journals, books, or conference proceedings.

The following sections deal with data, goals, tools and evaluation as they are used in the HPSG research community. These sections have a rather general character with the exception of Section 4, which deals with some of the formal foundations of HPSG. The formal foundations will be the basis of the analysis of an example sentence in Section 6.

2 Data

There is no dogma concerning the kind of evidence that should be used in work within HPSG. Most of the early practitioners of HPSG were coming from a Chomskyan research tradition and hence, worked mainly based on introspection. Since HPSG was well-formalized from the beginning it was used in grammar implementations and these implementations were used in research prototypes or applications. These were built with respect to naturally occurring data like spoken dialogues for appointment scheduling in the Verb*mobil* project, which was running from 1992–2000 (Wahlster 2000). Languages that are covered in Verb*mobil* are German (Müller & Kasper 2000), English (Flickinger et al. 2000), and Japanese (Siegel 2000). Other projects were set up with the goal of parsing parts of Wikipedia (Flickinger et al. 2010). So, computationally oriented researchers were working with corpus data right from the beginning. While more theoretically oriented papers started out with introspectively produced data, many researchers started using data from the web and those who were reflecting deeper about the possibility of reproducing the results of search queries and about accessibility of their data sources used corpora like the British National Corpus, Tiger (Brants et al. 2004) or COW (Schäfer & Bildhauer 2012). Nevertheless, introspection might be a useful guide but it alone is not sufficient since certain phenomena are just not accessible through introspection. See Müller (2007) and Meurers & Müller (2009) on wrong claims about particle verbs, extraposition and subjacency, and apparent multiple frontings in German that were based on introspective data.

On the other hand, attested examples alone are not sufficient either. For instance the following example from a newspaper is ungrammatical since it contains both *das* and *dem*, where only *das* would be grammatical:

(1) * *Dagegen hatte nach dem Bekanntwerden des ersten Berichtsentwurfs*
there.against had after the release of.the first report.draft
nicht nur das dem Umweltbundesamt protestiert.[1]
not only the the Umweltbundesamt protested
Intended: 'Not just the Umweltbundesamt protested against this after the release of the first draft of the report.'

[1] taz, 22.06.2017, p. 9 (The taz is a Berlin-based nation-wide newspaper.)

One can speculate how this error came about. Probably, it is due to some reformulation of the text. Unless such questions are the focus of inquiry, examples like (1) are not of interest and have to be filtered out. So a mix of methods is required: corpus data and introspection or corpus data and experiments. Since (1) is a clear case, introspection is sufficient here.

Apart from corpus data as used for example by Bildhauer (2011), all available evidence from experiments (e.g., speaker judgments, eye tracking data, and results from speeded reading tasks) are seen as relevant data. Since HPSG is a performance-compatible competence theory (Sag & Wasow 2011), it can be paired with a performance model and hence experimental data can feed back into theory development. The performance model can be seen as an additional layer of constraints referring to the linguistic knowledge and hence certain structures may be ruled out for performance reasons rather than being ruled out by the theory itself. An example of the latter is the question of why there are no languages that form questions by reversing the order of words in a string, a question often asked in Mainstream Generative Grammar. The answer of those working in HPSG is that our short-term memory is just not large enough to do such complex computations, a fact that is independent of our linguistic knowledge. A linguistic competence theory does not have to explain the non-existence of such languages. To take another, more relevant example, consider Sag et al. (2007: 228), who argue that Subjacency and the Complex NP Constraint should not be part of a competence grammar.

While some HPSG analyses are developed without testing their psycholinguistic plausibility, the analyses do make certain predictions that are open for testing. For example, Wittenberg & Piñango (2011) and Wittenberg et al. (2014) examined various analyses of complex predicates and found that Goldberg's analysis (2003) makes wrong predictions while the one in Müller (2010) is compatible with the psycholinguistic findings.

3 Goals

The goal of research is similar to the goals of Construction Grammar and hence also includes many of the goals of Mainstream Generative Grammar (GB, Minimalism and variants thereof): We want to understand language as a cognitive system, we want to understand which properties are common to all languages and how languages may vary, we want to understand how natural language can be acquired and processed. The explorations are not limited to a core grammar in the Chomskyan sense since it is believed that the so-called periphery interacts in interesting ways with what is thought of as the core. There is interesting research on idioms (Soehn & Sailer 2008, Sag 2007, Richter & Sailer 2009, Kay et al. 2015) and in fact the distinction between core and periphery does not play an important role in theorizing (Müller 2014b).

Research in HPSG is not limited to syntax. Many papers address semantic phenomena and make the syntax-semantics interface explicit. HPSG theories are declarative statements about language and this linguistic knowledge can be used in various ways. One way is to find the meaning of a given utterance (parsing) and the other way is to find the phonological or orthographic representation for a given meaning (production or generation). Apart from work on syntax and semantics, there is work on phonology (Bird & Klein 1994, Bird 1995, Orgun 1996, Höhle 1999, Klein 2000, Alexopoulou & Kolliakou 2002), morphology (Riehemann 1997, Crysmann & Bonami 2016), information structure (Engdahl & Vallduví 1996, Kuhn 1996, Wilcock 2005, De Kuthy 2002, Paggio 2005, Bildhauer 2008, Bildhauer & Cook 2010) and dialogue (Schlangen et al. 2003, Ginzburg & Cooper 2004).

Since the work is formalized it can be implemented and used in computer applications. One field for applications is machine translation (Oepen et al. 2007), another one information extraction. For further details see the webpage of the DELPH-IN consortium.[2]

4 Tools

HPSG is a model-theoretic approach (Pullum & Scholz 2001, Richter 2007) and hence belongs to the family of constraint-based theories.[3] Linguistic theories are sets of constraints (mostly feature-value pairs) that constrain the number of linguistic objects licensed by the theory. Theories like HPSG are surface oriented, that is, there is no underlying structure from which another representation is derived. A sentence like (2a) is analyzed directly involving the words that are visible. That is (2a) is not derived from (2b). Neither is (2b) derived from (2c):

(2) a. *This book, Kim was given as a present.*
  b. *Kim was given this book as a present.*
  c. *Somebody gave Kim this book as a present.*

Of course, Chomsky (1957) was right in pointing out that simple phrase structure grammars are inappropriate for modeling linguistic phenomena since they cannot account for the fact that these sentences are related. HPSG has means of capturing these relations, but they do not involve different levels like Deep and Surface Structure, and they do not involve transformations of complete sentential structures (or their equivalent in Minimalist theories). Rather than employing a passive transformation or more general transformations that account for passive, lexical rules are used deriving passive participles (like *given* in (2a)) from word stems (see Section 4.10). Similarly, extraction phenomena like the fronting of *this book* in (2a) are modeled by establishing a relation between the fronted element and the head of which the fronted element depends, but this does not involve movement in the literal sense (see Section 4.9). In what follows, we nevertheless use terms like *extraction* and *scrambling* since these terms are established in the literature. But as will become clear these phenomena are not described with respect to several trees but it is always just one linguistic structure that is assumed for one utterance.

The fact that HPSG is surface-oriented and transformationless is a huge advantage when it comes to psycholinguistical plausibility (Fodor, Bever & Garrett 1974: 320–

[2] `http://www.delph-in.net/wiki/index.php/Background`, 2018-06-23.

[3] These theories are sometimes also called unification-based theories but constraint-based is the more general term.

328). HPSG just assigns complex categories to linguistic objects (for instance 'complete nominal projection in the accusative') and states the dependency between elements (for instance head and its argument). These entities and the relations are directly observable. No statements regarding the processing order of constraints are made in HPSG theories. HPSG theories are just declarative statements about linguistic objects. This has the advantage that the linguistic knowledge can be paired with various processing models and processing regimes. The linguistic knowledge can be used for parsing utterances and for generation/production as well. HPSG grammars can also account for fragmentary input (Pullum & Scholz 2001: Section 3.2). This is not the case for some of its competitors, e.g., all those models that see languages as (infinite) sets that are enumerated by a grammar (Chomsky & Miller 1963: 283). See Sag & Wasow (2011), Müller (2018c) and Wasow (2019) for further discussion.

4.1 Features, values, structure sharing and relational constraints

HPSG makes use of a very small number of descriptive tools: features that have values of a certain type. Values of features may be complex. For instance, an AGREEMENT feature may have a complex value providing information about case, gender and number. Types are organized in hierarchies, which makes it possible to capture generalizations by positing abstract types and more specific subtypes (see also Section 4.6). Values of features can be identified with values of other features (structure sharing, explained below in more detail) or they can be related to other features by relational constraints (explained below). As will be shown in the remainder of the paper, this is sufficient to express everything one has to say about language: roots, stems, words, lexical rules, phrases can all be described using feature-value pairs.

HPSG assumes feature structures as models of linguistic objects.[4] These feature structures are described by feature descriptions, which are also called *attribute value matrix* (AVM). Such AVMs consist of feature value pairs. The values can be atomic (e.g., *sg* and *pl*) or feature descriptions. Every feature structure is of a certain type. Types are written in italics. They are ordered in hierarchies with the most general type at the top of the hierarchy and the most specific types at the bottom. Figure 1 shows an example hierarchy for the type *number* and its subtypes.

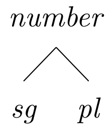

Figure 1: Subtypes of *number* in a grammar of English

Types in a model of a linguistic object are maximally specific, that is, a noun in a

[4]Feature structures are usually depicted as graphs (Pollard & Sag 1994: 16–17; Richter 2007). Due to space limitations we do not give an example here but provide feature descriptions only, which are used to formulate theories about possible feature structures.

model of an actual utterance has a NUMBER value that is *sg* or *pl*. The linguist develops theories that describe possible feature structures. In contrast to feature structures, feature descriptions can be partial. For instance it is not necessary to specify a NUMBER value for the word *sheep* since *sheep* can be used both in singular and in plural NPs.

(3) a. *one sheep*
b. *two sheep*

(4) is an example of a complex AVM. It shows a description of a referential index as it is used in the semantic representation of *man*:

(4) $$\begin{bmatrix} \textit{ref} & \\ \text{PER} & \textit{3} \\ \text{NUM} & \textit{sg} \\ \text{GEN} & \textit{mas} \end{bmatrix}$$

There are nouns like *cousin* that are underspecified with respect to their gender. They could either have male or female gender as is shown by examples with coreferring pronouns:

(5) *I met Peter's cousin. She/he is very tall.*

There are two ways to specify this. The first is to use a disjunction (*fem* ∨ *mas*) and the second is to use a common supertype for *fem* and *mas*. While both solutions are equivalent when it comes to models, descriptions without disjunctions are often more compact, which is why representation in the latter is preferred over the former. Figure 2 shows the type hierarchy for gender with a special type for objects that can be either *fem* or *mas*. The value of GEN in the description of the referential index of *cousin* is

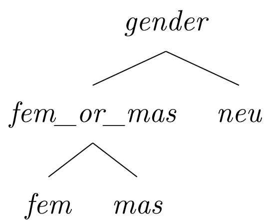

Figure 2: Subtypes of *gender* in a grammar of English

fem_or_mas. (See Müller (2016: Section 14.3) for alleged problems for model-theoretic syntax with cases like this)

One very important part of the formalism is *structure sharing*. It is used to express that information in feature structures is identical, that is, token-identical rather than just type identical. Structure sharing is indicated by boxed numbers in feature descriptions. An identical number at several places in an AVM expresses the fact that the respective values are identical.

To give an example of structure sharing, let us consider case agreement in German noun phrases:

(6) a. *der* *Mann*
the.NOM man.NOM ∨ DAT ∨ ACC

b. *des* *Mannes*
the.GEN man.GEN

The determiner has to agree with the noun in case. *Mann* can be nominative, dative or accusative. *Mannes* is in the genitive. The form of the determiner has to be *des* if the noun is genitive. So the case specification for *Mann* would be *nom* ∨ *dat* ∨ *acc* or an equivalent type. One could be tempted to suggest that the noun *Mann* has a case value that is *nom* ∨ *dat* ∨ *acc* and that it has to be combined with a determiner that has the CASE value *nom* ∨ *dat* ∨ *acc*. The following AVM depicts this without assuming anything about the theory that will be developed later:

(7)
$$\begin{bmatrix} \text{DETERMINER} & \begin{bmatrix} \text{CASE} & \textit{nom} \vee \textit{dat} \vee \textit{acc} \end{bmatrix} \\ \text{NOUN} & \begin{bmatrix} \text{CASE} & \textit{nom} \vee \textit{dat} \vee \textit{acc} \end{bmatrix} \end{bmatrix}$$

But a specification of the values as in (7) is not sufficient since when an actual determiner is chosen (*der* or *des*), the case of the complete NP is unambiguously determined, but this is not reflected in (7). With the setting above, we would get the following for inserting the determiner *der*:

(8)
$$\begin{bmatrix} \text{COMBINATION} & \begin{bmatrix} \text{CASE} & \textit{??} \end{bmatrix} \\ \text{DETERMINER} & \begin{bmatrix} \text{CASE} & \textit{nom} \end{bmatrix} \\ \text{NOUN} & \begin{bmatrix} \text{CASE} & \textit{nom} \vee \textit{dat} \vee \textit{acc} \end{bmatrix} \end{bmatrix}$$

The disjunctive specification of the determiner is resolved to nominative, but the other disjunction is not affected. Furthermore it is unclear what the case of the whole combination would be. If the case value of the whole NP is determined by the head (the noun alone), it would be *nom* ∨ *dat* ∨ *acc* and this would mean that the whole phrase *der Mann* could be used as a dative or an accusative object. Obviously, this is not what is wanted. What is needed instead is that the case of the determiner is token-identical to the case of the noun (agreement) and to the case of the complete phrase (projection of feature values). This is ensured by structure sharing:

(9)
$$\begin{bmatrix} \text{COMBINATION} & \begin{bmatrix} \text{CASE} & \boxed{1} \end{bmatrix} \\ \text{DETERMINER} & \begin{bmatrix} \text{CASE} & \boxed{1} \end{bmatrix} \\ \text{NOUN} & \begin{bmatrix} \text{CASE} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

With such a setting the case of the NP *der Mann* is nominative and the one of *des Mannes* is genitive, as expected. Note also that the case of *die Frau* is *nom* ∨ *acc* since

Frau is compatible with all four cases and *die* is *nom* ∨ *acc*. Depending on the governing verb this disjunction can be resolved to either *nom* or *acc*.

While structure sharing is the most important expressive means in HPSG there is one extension of the basic formalism that plays a crucial role in most HPSG analyses: *relational constraints*. Relational constraints are used to relate several values in a feature structure to each other. The relational constraint that is used most often in HPSG is *append* (⊕), which is used to concatenate two lists. The Schema 1, which will be discussed in Section 4.7, is an example for an application of such a constraint.

4.2 Descriptive levels, feature geometry and modularization

The following AVM is the description of the word *man*. The grouping of features – the so-called feature geometry – is the one of Pollard & Sag (1994) and Sag (1997).[5] (10) shows parts of the lexical item for *man*:

(10) lexical item for *man*:

$$\begin{bmatrix} \textit{word} \\ \text{PHONOLOGY} \ \langle \text{m,æ,n} \rangle \\ \text{SYNTAX-SEMANTICS} \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \textit{local} \\ \text{CATEGORY} \begin{bmatrix} \textit{category} \\ \text{HEAD} \ \textit{noun} \\ \text{SPR} \ \langle \text{DET} \rangle \end{bmatrix} \\ \text{CONTENT} \ \ldots \begin{bmatrix} \textit{man} \\ \text{INST} \ x \end{bmatrix} \end{bmatrix} \\ \text{NONLOCAL} \ \ldots \end{bmatrix} \end{bmatrix}$$

The first feature value pair describes the phonological form of the word. The value of PHONOLOGY is a list of phonemes. For reasons of readability usually the orthographic form is given in HPSG papers and phonological structure is omitted, but see Bird & Klein (1994), Bird (1995), Orgun (1996), Höhle (1999), Klein (2000), and Alexopoulou & Kolliakou (2002) for phonological analyses. The second feature is SYNTAX-SEMANTICS (SYNSEM) and its value is a description of all properties of a linguistic object that are syntactically and semantically relevant and can be selected by other heads. Information that is locally relevant (LOCAL) is distinguished from information that plays a role in non-local dependencies (NONLOCAL, see Section 4.9). Syntactic information is represented under CATEGORY (CAT) and semantic information under CONTENT (CONT). The example shows the HEAD value, which provides information about all syntactic aspects that are relevant for the external distribution of a maximal projection of a lexical head. In

[5] There are various HPSG variants around that differ mainly in the way features are grouped. One example is Sign-Based Construction Grammar (Sag 2012). For a discussion of SBCG's feature geometry see Müller (2018c: Section 10.6.2).

particular the part of speech information (*noun*) is represented under HEAD. As well as information regarding the head features, valence information also belongs under CAT. The example shows the SPR feature, which is used for the selection of a specifier (see Section 4.3 for details on valence).

The AVM in (10) shows a description of phonological, morpho-syntactic, and semantic aspects of a word. But of course other aspects can be and have been described by feature value pairs as well. For instance Engdahl & Vallduví (1996), Kuhn (1996), Wilcock (2005), De Kuthy (2002), Paggio (2005), Bildhauer (2008), and Bildhauer & Cook (2010) show how information structure can be modeled in HPSG in general and how the interaction between phonology, syntax, semantics, and information structure can be captured in a constraint-based setting. For general discussion of interfaces between the linguistic levels of description see Kuhn (2007).

As is clear from looking at (10), information about the descriptive linguistic levels is represented in one structure. This makes it possible to connect syntax and semantics (see Section 4.3), phonology and information structure (Bildhauer 2008), and syntax and information structure (Bildhauer & Cook 2010), phonology and semantics (Halliday 1970) and whatever other descriptive levels have to be connected. Since all information is in the same structure, HPSG is compatible with psycholinguistic findings that tell us that all available information is processed in parallel (Tanenhaus et al. 1995). This sets HPSG apart from other models like GB and Minimalism that assume that there are post-syntactic modules like Phonological Form and Logical Form. Minimalism has a conception that differs from the GB architecture (Richards 2015: 812, 830) but it is psycholinguistically as implausible as the GB model for the same reason. Language processing is not a bottom up combination with shipping complete phrases or phases to the interfaces. For a discussion of psycholinguistics from an HPSG perspective see Wasow (2019).

In principle, all the information in (10) could be provided in a simple, unstructured list of feature value pairs (as is done in Minimalism, for instance). However, having phonology, syntax, semantics and information structure in different parts of the structure provides a cleaner representation. Furthermore, the specific groupings of information are motivated by the need to share this information (see Section 4.1 on structure sharing). As will be shown in Section 4.7, the SYNSEM value of an argument will be structure shared with the respective representation in the valence list of its head. Similarly the information under LOCAL is shared between an extracted element and the place where it is missing (see Section 4.9). Everything under CAT (valence, part of speech, case, …) is identified in symmetric coordinations (Pollard & Sag 1994: 202) and finally, all the head features under HEAD are projected from head daughters to their mothers. In principle, it would be possible to share single values of an unstructured list of features but having these groupings allows for a more general treatment: nouns share other properties with their projections than verbs. A noun has a certain case and the respective value is relevant for the whole nominal projection. Similarly the verb form (*fin* = *finite*, *bse* = infinitive without *to*, *inf* = infinitive with *to*, *pas* = passive) is relevant for the maximal verbal projection. (11) shows two example HEAD values for nouns and verbs:

(11) a. $\begin{bmatrix} \textit{noun} \\ \text{CASE}\ \ \textit{gen} \end{bmatrix}$ b. $\begin{bmatrix} \textit{verb} \\ \text{VFORM}\ \ \textit{pas} \end{bmatrix}$

With such an encoding of information it is possible to provide general constraints for head information without making recourse to specific features and their values. See Section 4.5 on the Head Feature Principle.

4.3 Valence and linking

The previous subsections dealt with the basic formal apparatus that is used in HPSG and made some general remarks about foundational assumptions. In what follows, we will look at the analysis of the example sentence in (12) to explain further basic assumptions.[6]

(12) *After Mary introduced herself to the audience, she turned to a man that she had met before.*

As was said in the introduction, HPSG is a lexicalist theory. So, much information about the combinatorial potential of a head is represented in its lexical item. For instance the verbs in (12) can take two NPs, an NP and a PP, or two NPs and a PP. The required form of the arguments is described in a list, the so-called argument structure list (ARG-ST list). (13) provides some prototypical examples:

(13)

	ARG-ST
a. meet	⟨ NP, NP ⟩
b. turn	⟨ NP, PP[*to*] ⟩
c. introduce	⟨ NP, NP, PP[*to*] ⟩

NP and PP[*to*] are abbreviations. They stand for AVMs describing fully saturated nominal or prepositional objects. Square brackets are used in such abbreviations for specification of some values like case of NPs or the form of the preposition in PPs.

The elements on the argument structure list are ordered according to their obliqueness (Keenan & Comrie 1977). That is, subjects are ordered before primary objects before secondary objects before obliques. Since the order is fixed one can use this list to establish the particular linking patterns between arguments and the semantic roles they fill. For example, the first NP in the ARG-ST list of *turn* is linked to the agent (ARG1 in (14)) and the PP is linked to the person or object that is turned to (ARG2 in (14)).

(14) Linking for the verb *turn*:

$$\begin{bmatrix} \text{CAT}|\text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}},\ \text{PP}[\textit{to}]_{\boxed{2}} \right\rangle \\ \text{CONT}|\text{RELS} & \left\langle \begin{bmatrix} \textit{turn} \\ \text{ARG1}\ \ \boxed{1} \\ \text{ARG2}\ \ \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix}$$

[6] This sentence was provided as a shared task by the editors of this volume.

We do not use feature names like AGENT and THEME in the AVMs but rather more abstract features like ARG1. This is compatible with Dowty's approach (1991), which uses proto-roles.

For SVO languages it is useful to distinguish the arguments that are realized before the verb from those that follow it. The respective elements are represented in different valence lists. We assume that the subject (the first element of the ARG-ST list) is represented in a list called SPECIFIER list and that all other arguments are represented as elements of the COMPLEMENTS list.[7] These mappings are language-dependent. For example, for German, all arguments of finite verbs are mapped to COMPS (Pollard 1996: 295–296, Kiss 1995: 80).[8] (15) shows the respective mapping for *turn*:

(15) Mapping from ARG-ST to SPR and COMPS for verbs like *turn*:

$$\begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \\ \text{ARG-ST} & \langle \boxed{1}\ \text{NP},\ \boxed{2}\ \text{PP}[\textit{to}] \rangle \end{bmatrix} \end{bmatrix}$$

The structure sharing $\boxed{1}$ indicates that the first element of the ARG-ST list is identical to the element in SPR and the $\boxed{2}$ indicates that the second element in ARG-ST is identical to the element in the COMPS list. SPR is used as well for nouns selecting a determiner (see (10)). Figure 3 shows how *she turned to a man* is analyzed. The left figure shows the analysis with traditional abbreviations and the right one shows the individual features for valence and part of speech. An N′ is a linguistic object of category *noun* that does not select any complements but needs a specifier. A preposition is a linguistic object that selects an NP, that is, a linguistic object of category *noun* that is complete as far as valence is concerned. The abbreviation VP corresponds to a verbal projection of a verb with an empty COMPS list and one element in the SPR list. If the element in SPR is saturated as well, we get an S. Looking at the analysis of *a man* we see that the description in the SPR list of the noun is identified with the determiner that is combined with the noun ($\boxed{4}$). Elements combined with a lexical head selecting them are not represented in the valence list of the resulting mother node. All other elements in valence lists are passed up. For instance *turned* selects both for a PP and an NP. The PP is combined with the verb first so that the description of the PP is not contained in the mother node (but see Section 6.3 for a modification). But the NP requirement is passed up: *turned to a man* selects an NP via its SPR list. After combination with *she* we get a fully saturated phrase, a maximal projection, that is, something with empty

[7] Borsley (1987) argues for a SUBJ, a SPR, and a COMPS feature. We follow Borsley (1989) and the German tradition and assume a head feature SUBJ that is used for control and raising constructions (Pollard 1996: 295–296, Kiss 1995: Section 3.1.1, Müller 1999b: Section 1.7). Under this assumption, SUBJ is not a valence feature, a head cannot be combined with anything in SUBJ. We assume that those subjects that can be combined with their head are either in SPR or in COMPS, depending on the language (Müller 2018b).

[8] German is a V2+SOV language. Hence, orders in which the subject appears in front of the verb are analyzed as extraction structures in which the subject is fronted. (Note: we are using the term *fronted* here since this is a handy way to describe German. In the analysis there is no underlying sentence from which something is fronted. See Section 4.9 on nonlocal dependencies.)

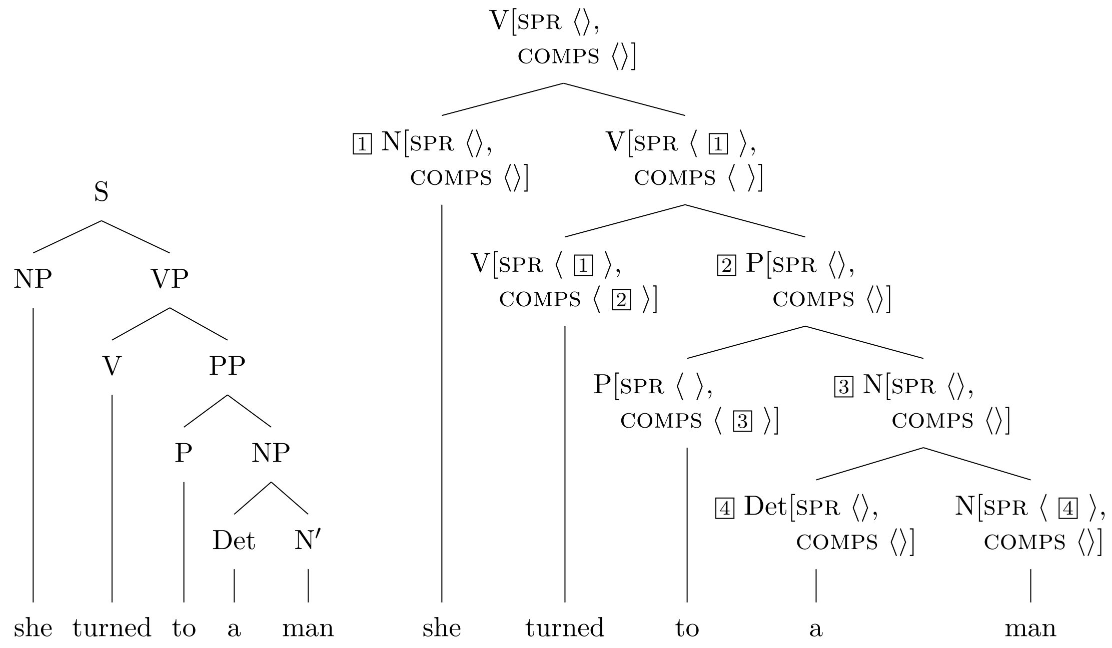

Figure 3: Analysis of *she turned to a man*

SPR and COMPS list.

Note that this representation of valence avoids unary projections as they are common in $\overline{\text{X}}$ theory: the pronoun *she* is just an NP without the intermediate projections from N^0 to $\overline{\text{N}}$ to NP. Similarly, an intransitive verb in English can be treated as a VP.[9]

In this paragraph and throughout the paper, the combination of items is described in a bottom-up way, but it is important to note that all statements are purely declarative, which means that there is no order in which constraints have to apply.

In the introduction, it was mentioned that in HPSG everything is done with feature value pairs, that is, without trees and phrase structure rules. How this is achieved will be explained in the next subsection.

4.4 Constituent Structure

While other theories that are similar to HPSG in using feature value pairs for describing complex categories use special phrase structure rules to model constituent structure (Bresnan & Kaplan 1982), HPSG uses feature descriptions also for constraints on constituent structure. For example, the structure for *a man* can be represented by using features whose values correspond to the daughters in the tree. (16) shows parts of the structure for *a man*:

[9] Pollard & Sag (1994) and Ginzburg & Sag (2000: 34, 364) assume that a lexical verb is projected to the VP level in any case. See Müller (2013b: 935) for some discussion.

(16) $\begin{bmatrix} \text{PHON} & \langle\ a,\ man\ \rangle \\ \text{DTRS} & \left\langle \begin{bmatrix}\text{PHON}\ \langle\ a\ \rangle\end{bmatrix}, \begin{bmatrix}\text{PHON}\ \langle\ man\ \rangle\end{bmatrix} \right\rangle \end{bmatrix}$

Note that HPSG differs from many theories in that the phonology is represented at the mother nodes. So HPSG is not like other theories where just the leaves in a tree are concatenated. Rather every linguistic object has its own PHON value. This makes it possible to specify constraints on phonology that are dependent on structure without assuming that these phonological constraints are somehow external or post-syntactic.

In addition to what is given in (16), structure sharing is used to point to the daughter that contains the head:

(17) $\begin{bmatrix} \text{PHON} & \langle\ a,\ man\ \rangle \\ \text{HEAD-DTR} & \boxed{1} \\ \text{DTRS} & \left\langle \begin{bmatrix}\text{PHON}\ \langle\ a\ \rangle\end{bmatrix}, \boxed{1} \begin{bmatrix}\text{PHON}\ \langle\ man\ \rangle\end{bmatrix} \right\rangle \end{bmatrix}$

The so-called head daughter is the daughter that contains the head. In the case of *a man* this is simply the noun *man* but for [[*introduced herself*] *to the audience*] the head daughter would be *introduced herself* since this phrase contains the head *introduced*.

Both flat (Pollard & Sag 1987, 1994, Sag 1997, Ginzburg & Sag 2000) and binary branching (Kiss 1995, Meurers 1999a, Kathol 2001, Müller 2002, 2018b) structures have been suggested in the literature. We will assume binary branching structures in this paper. The basic combinatorial schemata, which are introduced in Section 4.7, are similar to forward and backward application in Categorial Grammar (Ajdukiewicz 1935, Steedman & Baldridge 2006) and to Merge in Minimalism (Chomsky 1995). See Müller (2013b) for a detailed comparison and some remarks concerning the history of ideas. The reason for assuming binary branching structures is that this makes it possible to assume the same set of schemata for head-argument combinations for many languages (if not for all) and hence allowing to capture crosslinguistic generalizations. Note though that the representation of daughters is sufficiently general to allow for flat structures. The N-P-N construction, suggested by Jackendoff (2008) to analyze phrases like *student after student*, is an example where we would assume flat structures with more than two daughters (Bargmann 2015, Müller 2019).

4.5 Principles: Implicational constraints

HPSG publications often contain prose statements stating principles. One such principle is the Head Feature Principle which says that in headed structures the head features of the head daughter are identical to the head features of the mother. This principle is formalized by an implicational constraint:

(18) *headed-phrase* ⇒

$\begin{bmatrix} \text{SYNSEM|LOC|CAT|HEAD}\ \boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|HEAD}\ \boxed{1} \end{bmatrix}$

(18) will be explained in more detail below, what is important here is the formal aspect: HPSG can formulate implicational constraints that are to be interpreted as logical, that is, if the left-hand side of the implication holds, the right-hand side must hold as well. In (18) the left-hand side is a single type (*headed-phrase*) but in principle a complex description could be used at the left-hand side as well.

As mentioned above, types are organized in hierarchies (see also Section 4.6). An implication such as (18) holds for all structures of type *headed-phrase* and this includes of course all subtypes of *headed-phrase*.

Note that the constraint above does not entail that all structures are headed, it only states that certain constraints must hold for structures of type *headed-phrase*. The implicational constraint does not say anything about structures of another type.

4.6 Inheritance hierarchies

In Section 4.1, we explained type hierarchies and showed how they may be useful for specifying features and leaving them underspecified. But this is not the only advantage of using types in a theory. Types play an important role for capturing generalizations. Figure 4 shows a type hierarchy for the subtypes of *sign*. *sign* is the most general type for feature structures of linguistic objects. Signs can be of type *stem*, *word* or *phrase*.

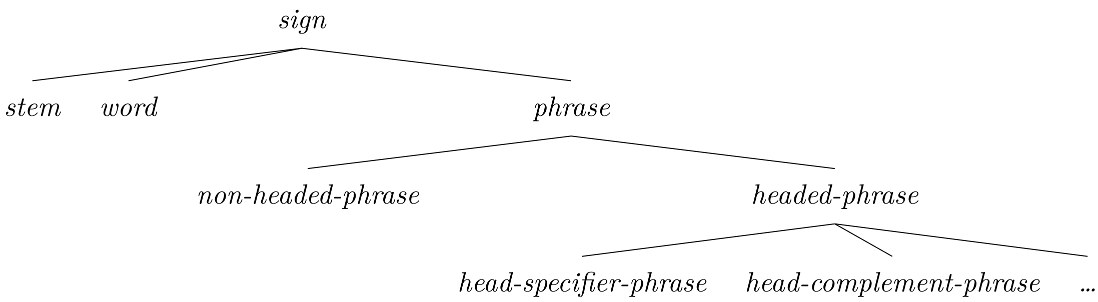

Figure 4: Type hierarchy for *sign*: all subtypes of *headed-phrase* inherit constraints

Types are associated with features. For instance, feature structures of type *sign* always have a PHON value and a SYNSEM value. Signs of type *phrase* have a DTRS value in addition to the features introduced by *sign* and phrases that are of type *headed-phrase* have a HEAD-DTR in addition to everything that *phrase* has. Besides the introduction of features at certain types, values of features can be specified. So, type hierarchies can be used to capture generalizations over certain linguistic objects. They can be used to capture lexical generalizations as well as to classify phrases. Construction Grammar is well-known for using inheritance hierarchies for capturing generalizations (Goldberg 1996, Croft 2001) but HPSG uses inheritance since the very first HPSG paper (Flickinger, Pollard & Wasow 1985). The early work on inheritance was work about the lexicon since HPSG is a lexicalist framework but later this was extended to phrasal types (Sag 1997, 2010).

4.7 Head-Argument Schemata

With the kind of representation of constituent structure introduced in Section 4.4, we can now formalize the treatment of valence depicted in Figure 3. The following schema licenses Head-Complement structures:

Schema 1 (Head-Complement Schema [preliminary])
head-complement-phrase ⇒

$$\begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{comps}\ \boxed{1} \\ \textsc{head-dtr}|\ \boxed{2}\ \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{comps}\ \langle\ \boxed{3}\ \rangle \oplus \boxed{1} \\ \textsc{dtrs}\ \left\langle \boxed{2},\ \begin{bmatrix}\textsc{synsem}\ \boxed{3}\end{bmatrix} \right\rangle \end{bmatrix}$$

The schema expresses constraints on structures of type *head-complement-phrase*.[10] The schema states that the COMPS list of the head daughter is split in two parts: a list with one element (⟨ $\boxed{3}$ ⟩) and the remainder of the list ($\boxed{1}$).[11] The COMPS list contains descriptions of the syntactic and semantic properties of arguments. One such description is identified with the SYNSEM value of the second element in the DTRS list, which is the non-head daughter. The remainder of the list ($\boxed{1}$) is identified with the COMPS list of the whole phrase. The SPR list of the head daughter is not affected in head-complement phrases and hence the SPR value of the mother node is identical to the SPR value of the head daughter. This is not shown in the schema. Since this passing on of the SPR value holds for further phrasal types of phrases (e.g., head-adjunct phrases), the constraints are formulated as constraints on a supertype from which *head-complement-phrase* inherits.

The head daughter ($\boxed{2}$) is identified with the first element of the DTRS list. The second element of the DTRS list corresponds to one element of the COMPS list of the head daughter ($\boxed{3}$). The DTRS list is assumed to be ordered in the way the elements are serialized. Therefore, the PHON value of the mother node is the concatenation of the PHON values of the daughters. In English, the complements of verbs, nouns and adjectives follow the head, but in languages like German and Dutch the complements follow nouns and prepositions but they precede adjectives and verbs. One of the cases where one could claim a head-final order in English is the postposition *ago* (cf. the preposition *in*):

(19) a. *one year ago*
 b. *in one year*

In order to account for the two possibilities (19a) vs. (19b), one could simply state another version of the Head-Complement Schema or one could assume a more abstract

[10] The same kind of implicational constraint is used for stating principles, but although principles and schemata look similar, principles usually are defined as constraints on more general types. For example, the type *headed-phrase* is a supertype of *specifier-head-phrase*, *head-complement-phrase*, *filler-head-phrase* and so on. See Sag (2010: 533) for an elaborate type hierarchy of English clause types.

[11] The alternative is to combine a head with all of its complements in one go (Ginzburg & Sag 2000: 33–34). The result of this alternative is a flat structure (see Section 4.4).

representation of the schema, one that is neutral with respect to serialization of the daughters. The more abstract version is provided as Schema 2.

Schema 2 (Head-Complement Schema)
head-complement-phrase ⇒

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{COMPS}\ \boxed{1} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{COMPS}\ \langle\ \boxed{2}\ \rangle \oplus \boxed{1} \\ \text{NON-HEAD-DTRS}\ \left\langle \begin{bmatrix} \text{SYNSEM}\ \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix}$$

This schema does not constrain the order of the daughters. Since head daughter and non-head daughter are represented as values of two different features nothing is said about the order. One can then assume two different subtypes: one in which the head daughter is the first element in the DTRS list and the non-head daughter the second one and another subtype in which the head daughter is the second element in the DTRS list and the non-head daughter is the first one. The latter version is used to analyze orders like (19a).[12]

In Section 4.3, two valence features SPR and COMPS were introduced. We provided the schema for head-complement phrases above and a parallel schema for specifier-head combinations is given as Schema 3:

Schema 3 (Head-Specifier Schema)

head-specifier-phrase ⇒

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{SPR}\ \boxed{1} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}\ \begin{bmatrix} \text{SPR} & \boxed{1} \oplus \langle\ \boxed{2}\ \rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS}\ \left\langle \begin{bmatrix} \text{SYNSEM}\ \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix}$$

The last element of the SPR list is realized as the non-head daughter. The remaining list is passed up to the mother node. Note that the non-head daughter is taken from the end of the SPR list, while the non-head daughter in head-complement phrases is taken from the beginning. For heads that have exactly one specifier this difference is irrelevant, but in the analysis of object shift in Danish suggested by Müller & Ørsnes (2013b), the authors assume multiple specifiers and hence the difference in order of combination is relevant.

The COMPS values of mother and head daughter are identified in specifier-head phrases, as it is the case for the SPR value in head-complement phrases. The respective constraints are inherited from a supertype and are not given here. The head daughter's COMPS value

[12] Due to space limitations scrambling cannot be explained here. In order to account for languages with freer constituent order, the order in which items from the COMPS list are combined with their head is relaxed. For details see Müller (2016) or Müller (2015a).

is specified to be the empty list. This ensures that all complements are combined with the head first to form a $\overline{\text{N}}$ or a VP and the specifier(s) are combined with the respective projection after the combination with the complements.

Since specifiers always precede their heads, we could also have provided the schema with reference to DTRS rather than mentioning NON-HEAD-DTRS.

With the two schemata above and appropriately specified lexical items, we almost have everything that is needed for the analysis depicted in Figure 3. What has not been explained yet is how the part of speech information at the mother node of a tree is constrained. This is ensured by the Head Feature Principle, which was introduced in Section 4.5. Since the part of speech information is part of the HEAD value of signs (see (11)), it is passed up to mother nodes in syntactic structures. Depending on the part of speech, other information is passed up to the mother along the head path. For example, verbs project information about the form of the verb, that is, the maximal projection of a verb contains information on whether the verb inside the sentence is finite, a perfect participle or some other form of the verb. Information about the form of the preposition is also passed up in order to make it selectable by governing verbs. Similarly, for languages with case inflection at nouns the case information is projected to the NP node.

4.8 Adjunction

Our example sentence in (12) contains two adjuncts: the adverbial clause *after Mary introduced herself to the audience* is attached to the remainder of the sentence and *before* is attached to *met*. The basic technique to describe head-adjunct combinations is similar to what we saw about head-argument combinations. Adjuncts are functors and have a special feature (MODIFIED, MOD) whose value is a description of heads they can combine with. For instance, *before* can modify VPs. The MOD feature is a head feature, that is, it is projected along the head path. (20) gives the lexical item of *before*:

(20) CAT value of the adverb *before*:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{adv} \\ \text{MOD} \;\; \text{VP} \end{bmatrix} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix}$$

This lexical item for *before* should be contrasted with the preposition *before* as it is used in (21):

(21) *She met a man before the meeting.*

The *before* that is used in (21) selects for an NP and only after the combination of *before* and *the meeting* the complete phrase may modify a VP.

(22) CAT value of the preposition *before*:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{prep} \\ \text{MOD} \ \ \text{VP} \end{bmatrix} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle \ \text{NP} \ \rangle \end{bmatrix}$$

Since the information about the item that can be modified is part of the HEAD value, it is ensured that this information is also present at projections of *before*, that is, the PP *before the meeting* has the same MOD value as the preposition *before*.

Figure 5 shows the analysis of the example in (23):

(23) *She met him before.*

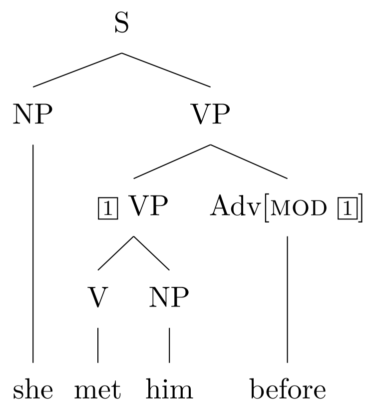

Figure 5: Analysis of sentence with adjunct

The MOD value of the adverb is identified with the head daughter. Head-adjunct structures are licensed by the following schema:

Schema 4 (Head-Adjunct Schema)
head-adjunct-phrase ⇒

$$\begin{bmatrix} \text{HEAD-DTR}|\text{SYNSEM} \ \boxed{1} \\ \text{NON-HEAD-DTRS} \left\langle \begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT} \begin{bmatrix} \text{HEAD}|\text{MOD} & \boxed{1} \\ \text{SPR} & \langle\rangle \\ \text{COMP} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

This schema enforces the identity between the MOD value of the non-head daughter with the SYNSEM value of the head daughter. The adjunct (the non-head daughter) has to be completely saturated. Without such a requirement the theory would admit strings like (24), in which the preposition *in* entered a head-adjunct structure without being completely saturated.

(24) * *Mary worked in.*

Heads cannot take (further) arguments if their valence features have the empty list (⟨⟩) as their value. Similarly, it has to be ensured that certain words or phrases cannot be used as modifiers. For example, a pronoun like *he* or a complete NP like *the man* does not modify anything. In order to make sure that such lexical items and phrases do not enter head-adjunct structures, their MOD value or rather the MOD value of their head is specified as *none*. Since *none* is incompatible with any *synsem* object, words like *he* and phrases like *the man* are incompatible with the requirements for non-head daughters in head-adjunct phrases.

4.9 Extraction: Modeling nonlocal dependencies as sequences of local dependencies

The analysis of unbounded dependencies is inherited from GPSG and originally due to Gazdar (1981). We want to explain it with reference to the examples in (25):

(25) a. *This man, she met before.*

b. *This man, I think that she met before.*

In what follows we assume an empty category-based analysis of nonlocal dependencies.[13] Figure 6 shows the analysis of (25a): the position of the object is taken by a trace. The trace is basically a joker that can fulfill whatever is required in a phrase. However, the trace has special properties. It passes up information about the missing element (NP in the example).[14] This is indicated by the slash ('/') in the figure. The information is passed on to higher nodes until it is finally bound off by some other element being compatible with the properties that are passed on. This element is called *filler*. A complete sentence is a verbal projection that has fully saturated valency lists and no element in the SLASH list.

The lexical item for a trace is given in (26).[15] Further NONLOC features are introduced below. Their values are the empty list, therefore they are not provided here to enhance readability.

(26) Lexical item for a trace (adapted from Pollard & Sag 1994: 164):

$$\begin{bmatrix} \text{PHON} & \langle\rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{NONLOC} & \begin{bmatrix} \text{INHER}|\text{SLASH} & \langle \boxed{1} \rangle \\ \text{TO-BIND}|\text{SLASH} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

[13] For traceless analyses of extraction see Bouma, Malouf & Sag (2001). A detailed discussion of this analysis can be found in Levine & Hukari (2006).

[14] Again the analysis is explained bottom up for explanatory purposes only.

[15] Pollard & Sag (1994) use sets as the value of nonlocal features. The mathematical formalization behind sets is very complicated (Pollard & Moshier 1990). We use lists in what follows.

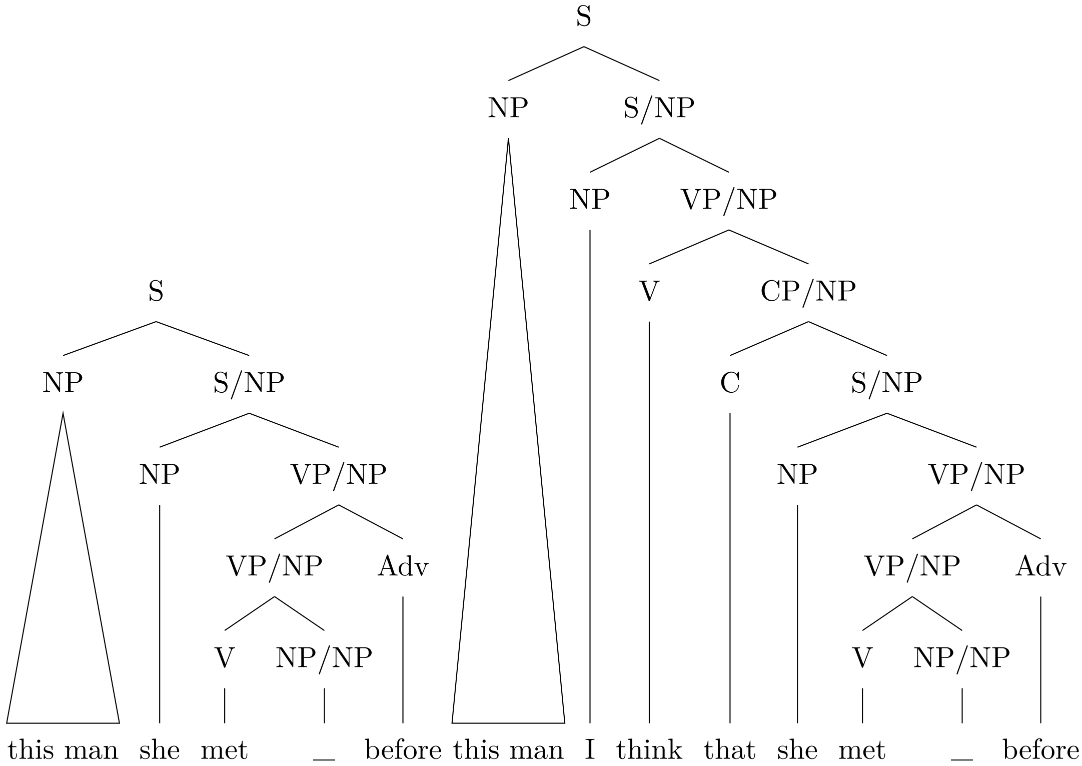

Figure 6: Analysis of nonlocal dependencies as local passing of information

This basically says: the word has no phonological material, i. e., nothing is pronounced. Whatever is locally required is compatible with the trace. The trace is a man without qualities, it does what it is told. But: whatever it is that is required by the syntactic context (NP, PP, an adjunct), the LOCAL value of the trace is identified with the element in the INHERITED|SLASH value ($\boxed{1}$). The SLASH value is passed up the tree by the Nonlocal Feature Principle, which states that the lists of nonlocal features at a mother node are the concatenation of the nonlocal features of the daughters minus those elements that have been bound off (those that are listed under TO-BIND in a daughter, usually in the head daughter). Figure 7 shows this in detail. The boxed number after a category symbol (for example the $\boxed{1}$ following the NP) refers to the LOCAL value of the respective linguistic object while boxes in front of category symbols as in earlier figures refer to SYNSEM values. All normal words have empty SLASH values, but the trace contributes a SLASH element. The respective lists are concatenated and since no word or schema has a specified TO-BIND|SLASH value, nothing is bound off until *she met before* is combined with the filler *this man*. This special filler-head combination is licensed by the following schema:

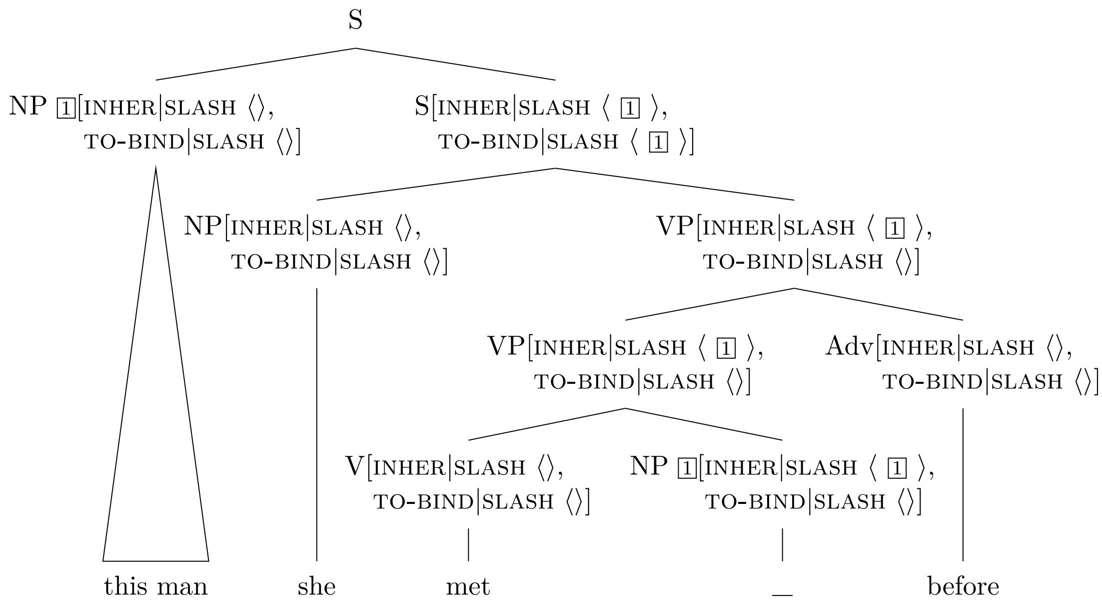

Figure 7: Analysis of nonlocal dependencies as local passing of information

Schema 5 (Head-Filler Schema)
head-filler-phrase ⇒

$$\begin{bmatrix} \text{HEAD-DTR } \boxed{1} \begin{bmatrix} \text{SYNSEM} \begin{bmatrix} \text{LOC|CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{verb} \\ \text{VFORM } \textit{fin} \end{bmatrix} \\ \text{SPR } \langle\rangle \\ \text{COMPS } \langle\rangle \end{bmatrix} \\ \text{NONLOC } \begin{bmatrix}\text{TO-BIND|SLASH } \langle \boxed{2} \rangle\end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{DTRS } \left\langle \begin{bmatrix} \text{SYNSEM} \begin{bmatrix} \text{LOC } \boxed{2} \\ \text{NONLOC|INHER|SLASH } \langle\rangle \end{bmatrix} \end{bmatrix}, \boxed{1} \right\rangle \end{bmatrix}$$

The LOCAL value that was passed up through the tree is identified with the LOCAL value of the non-head daughter. The specification of the INHER|SLASH value of the non-head daughter makes sure that nothing is extracted out of the filler. The head daughter is specified to be a finite clause with all arguments saturated.

There is a Nonlocal Feature Principle that ensures that the NONLOC values of the mother are the concatenation of the NONLOC values of the daughters minus the elements in TO-BIND. Since the TO-BIND|SLASH value of the head-daughter is non-empty, there has to be an INHER|SLASH value at the head daughter since something has to be bound off. The respective element in INHER|SLASH has to be [2]. INHER|SLASH of the head daughter may contain further elements, but [2] has to be in there and it will be bound off and not contained in the INHER|SLASH value of the mother node.

We used the empty element in (26) in the analysis of nonlocal dependencies and we want to close this subsection with some general comment on empty elements in HPSG. Empty elements are usually frowned upon within the HPSG community and some researchers do not use them at all or at least avoid them for certain kinds of phenomena (Sag & Fodor 1994; Bouma, Malouf & Sag 2001: Section 3.5), but there is no dogmatic ban on empty elements as for instance in Construction Grammar. For instance, suggestions to assume empty elements can be found in Bender (2000), Sag, Wasow & Bender (2003: 464), Borsley (1999, 2009) and Alqurashi & Borsley (2013). Personally, we agree with CxG views that there is a language acquisition problem with empty elements but we think this holds only for those empty elements that cannot be motivated by language-internal evidence. For instance, facts about object agreement in Basque are not accessible to learners of English and hence they could not learn an AgrO projection (as assumed for instance by Chomsky 1995: 7, 59–60). But learners of English have evidence that they can leave out determiners in plural noun phrases. As it has been shown elsewhere, there are situations in which grammars using empty elements capture the generalizations regarding omissible elements more directly than grammars without empty elements (Müller, 2014a; 2016: Chapter 19).

4.10 Roots, words, lexical rules

We showed in Section 4.3 how valence information is represented and in Section 4.7 how this valence information determines which kinds of trees are licensed. One important tool of HPSG has not been mentioned yet: lexical rules. Lexical rules are used to relate lexical objects. For languages with inflection, roots are the minimal objects described by linguists. These roots are related to inflected forms by lexical rules. Lexical rules are also used for derivational morphology. For example, the German adjective *lesbare* ‘readable’ is derived from the root *les-* ‘read’ by appending the suffix *-bar* ‘-able’ (Müller 2003). The resulting adjectival stem *lesbar* ‘readable’ is inflected by adding the suffix *-e*.

Lexical rules are also used for modeling valence alternations. The following simplified lexical rule accounts for the passive:

(27) $$\begin{bmatrix} \textit{stem} & \\ \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \langle \text{ NP } \rangle \oplus \boxed{1} \end{bmatrix} \mapsto \begin{bmatrix} \textit{word} & \\ \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{VFORM} & \textit{pas} \end{bmatrix} \\ \text{ARG-ST} & \boxed{1} \end{bmatrix}$$

The lexical rule in (27) basically maps a verb selecting for at least one NP to a participle passive that does not select for the subject but for all other arguments. The remaining arguments will be mapped to SPR and COMPS. English requires that $\boxed{1}$ starts with another NP or a sentential argument so that this NP (or sentential argument) can be put into the SPR list. Of course more has to be said about case assignment and so on but the basic explanation given above may be sufficient to understand the concept of a lexical rule. For details on passives in English see Pollard & Sag (1987) and on passives in Germanic languages in general see Müller & Ørsnes (2013a), Müller (2018b).

Lexical rules are doing some of the work that was done with transformations. Dowty (1978) called them lexically governed transformations. There is a crucial difference though: while transformations relate actual trees that have to be generated by the grammar before being able to function as inputs to transformations, lexical items license classes of trees. So lexical rules relate lexical items that license different classes of trees rather than relating trees directly. This difference is important when considering the psycholinguistic plausibility of theoretical models (Bresnan 1978).

4.11 Grammatical functions

HPSG refers to NPs, PPs, and so on but does not refer to subjects, objects, and other grammatical functions as descriptive primitives of the theory. An exception is the SUBJ feature, which refers to subjects, but this is needed for separating subjects from other arguments in SVO languages[16] and for allowing access to information about non-expressed subjects of non-finite verbs in control and raising structures (see also footnote 7).

As has been often mentioned in the literature, it is not trivial to define grammatical functions in a crosslinguistically valid way. For instance, in German, subject (as far as NPs are concerned) can be defined as equivalent with non-predicative nominative (Reis 1982). But in Icelandic, there are quirky case subjects and nominatives can be objects (Zaenen, Maling & Thráinsson 1985). The verb agrees with the nominative element independent of its grammatical function. Hence, the term *subject-verb agreement* is inappropriate for Icelandic. Rather one should talk about *nominative-verb agreement*. As already mentioned above, subjects in SVO languages are serialized differently from objects and hence there has to be a syntactic differentiation between subjects and objects. A further difference is controlability. Both of these criteria have been used in the grammar of Icelandic to identify subjects. Using the SUBJ feature for modeling these properties is one way to deal with the data, but it does not presuppose a unified concept of subject that applies to all subjects cross-linguistically. Nevertheless, the individual criteria that have been suggested for subjecthood may apply within languages and of course they are modeled in HPSG (position before the verb in SVO languages, controlability, agreement, omitability in imperatives, etc.). However, these criteria may cluster differently from language to language. This is entirely unproblematic since it is not necessary in HPSG to assign grammatical functions to constituents.

4.12 Levels of representation

HPSG does not assume a Deep Structure from which a Surface Structure is derived by transformations as in Government & Binding (Chomsky 1981), but there is something

[16] We use the SPR feature for subjects in this paper, but see Ginzburg & Sag (2000) for the use of SUBJ as valence feature. Some versions of HPSG do not distinguish between subjects and complements in the valence list of heads (Sag 2012; see Müller 2018c: Section 10.6.2.3 for discussion). These versions use a feature XARG to make one argument (usually the subject) accessible for control and also question tag formation (Bender & Flickinger 1999).

similar to Deep Structure: the ARGUMENT-STRUCTURE list (see Section 4.3). This list is a representation that contains all arguments of a head in a certain order.[17] This argument structure list can be used for linking valence requirements to semantic roles (Wechsler et al. 2019) as well as for Binding Theory (Pollard & Sag 1992, Branco 2019). So, things that are done in Deep Structure trees in GB, namely assigning theta roles to arguments at a certain position, are done within lexical items in HPSG.

4.13 Summary

This brief sketch of the formal foundation and basic tools mentioned many essential concepts that are used in HPSG. Of course a lot more could be and has been said about the properties of the formalism, but this introductory article is not the place to discuss them in detail. However, it cannot be emphasized enough that it is important that the formal details are worked out. The interested reader is referred to the work of Shieber (1986), Pollard & Sag (1987: Chapter 2), Johnson (1988), Carpenter (1992), King (1994, 1999), Pollard (1999) and Richter (2004, 2007). The work of King, Pollard, and Richter reflects current assumptions, that is, the model theoretic view on grammar that is assumed nowadays.

The following section deals with criteria for viewing an HPSG analysis as successful and after this section we fill in the missing pieces required for the analysis of the given example sentence.

5 Evaluation

Since HPSG is well-formalized it is always clear what a certain analysis stands for and what predictions it makes. It is therefore possible to test the analysis against collected data. This can be done in two ways: either by thinking about the consequences of an analysis or – more systematic and more reliable – by computer implementations that can be run against test suites (Oepen & Flickinger 1998). Test suites are data collections of either hand-made or otherwise available sentences or phrases. Hand-made test suites can contain ungrammatical strings, which are especially valuable since they can be used for testing for overgeneration of linguistic theories. Systematic testing is important since it is often the case that one believes to have found a real simplification of one's theory. Running the test suite can show us the one example out of several thousands not covered by the simpler grammar, or the additional examples getting unwanted structures/readings. In pencil and paper work, these examples could easily escape our attention. Another way to test implemented grammars is to let them generate strings for a given meaning. The results are often surprising. They reveal aspects of the grammar nobody ever thought about before.

During a workshop on *Progress in linguistics* at the Freie Universität Berlin in 2013 the first author suggested that proponents of linguistic theories should work out grammar fragments of reasonable size and provide lists of covered and rejected example sentences

[17] Butt (1995: 27) points out the LFG's f-structure corresponds to Deep Structure.

(see Müller & Ørsnes (2015) and Müller (2017) for examples). When theories are developed further it can be checked whether the amount of covered data stays constant or is getting bigger. This is a possible way to evaluate the success of theoretical work. While some of the frameworks currently available changed their fundamental assumptions frequently, this is not the case for HPSG. There were no radical breaks during the past 33 years. This ensures that the amount of data that is covered by HPSG theories steadily increases.

There are always many ways to write a grammar for a given data set. In order to decide which grammar is the best one, one can compare these grammars to grammars of other languages. Let's say there are two ways of describing phenomenon X. If we have one grammar that can account for the data and is compatible with what we know about another language or other languages then we choose this grammar over the other (Müller 2015b). Similarly, a simplicity metric can also be applied language internally: we chose the grammar that has to postulate fewer theoretical entities: fewer features, fewer empty elements. This is nothing special though. It is common scientific practice, also known as Occam's Razor. It should be noted though that the most compact description of linguistic knowledge is not necessarily the best one. There is ample evidence that a lot of linguistic knowledge is just stored as chunks in the human brain. So even though we could write a more compact grammar that derives chunks from their components rather than storing them, we do not necessarily do this since the empirical domain is not just a set of generated sentences. It is important how linguistic knowledge is represented in the brain and how it is used by speakers of the language.

6 Sample analysis

As was mentioned in Section 4.3, the shared task for authors of this volume is to analyze (12) – repeated here as (28) for convenience:

(28) *After Mary introduced herself to the audience, she turned to a man that she had met before.*

Some details concerning the analysis of (28) were already discussed in Section 4 by explaining the tools used in HPSG. But since some of the phenomena manifested in (28) are less foundational, we decided to put their analysis in a separate section.

In order to explain the HPSG analysis of the sentence, one has to explain how valence information is encoded and how it is linked to semantic representations (Section 4.3), and how the internal structure of basic sentences is licensed (Section 4.7). In order to account for the attachment of the adjunct *after Mary introduced herself to the audience* to the main clause and to explain how the relative clause *that she had met before* attaches to the noun it modifies, we have to explain how adjunction works (Section 4.8).

The analysis of relative clauses like *that she had met before* in the above example involves nonlocal dependencies. The respective schemata and mechanisms for establishing nonlocal dependencies are explained in Section 4.9. In Section 6.1, we provide a special schema for relative clauses, which is needed in addition.

Analyses of complex tenses are required to treat periphrastic forms like *had met*. Auxiliaries are raising verbs and hence we introduce the analysis of raising in Section 6.2. Case assignment and agreement play a role even in simple sentences. These phenomena are dealt with in Sections 6.3 and 6.4, respectively.

In order to account for the binding of the pronoun *herself* to *Mary*, something would have to be said about Binding Theory (Pollard & Sag 1992). Due to space limitations Binding Theory will not be covered in this overview.[18] Section 6.5 sketches the semantics that is used.

The analysis of (28) is implemented in the TRALE system (Meurers, Penn & Richter 2002, Penn 2004) as part of the BEngl grammar (Müller 2009, 2012, 2018b,a). This grammar is developed as part of the CoreGram project (Müller 2015b). The grammars that are developed in this project share a common core of grammatical constraints. There are grammars for German, Danish, Persian, Maltese, Mandarin Chinese, English, French, and some toy fragments of other languages. The detailed analysis of (28) including semantics is available at `https://hpsg.hu-berlin.de/~stefan/Pub/current-approaches-hpsg.html`. Due to space limitations we can not explain the analysis in full detail here, especially with respect to semantics. The reader is referred to other sources in what follows. Some of the analyses deviate from theory variants that may be more common in the literature. When this is the case, reasons for deviating will be mentioned or discussed in footnotes.

6.1 Relative clauses

The analysis of relative clauses builds on the analysis of nonlocal dependencies already introduced in Section 4.9. The type of relative clause that is relevant in the analysis of (12) consists of an extracted phrase containing a relative pronoun and a clause from which it is extracted. As the bracketing in (29) shows, the relative pronoun *who* can be realized far away from where it would be usually realized in a declarative sentence.

(29) *the man who*$_i$ [*I* [*believe* [*that* [*Mary* [*was* [*introduced* [*to* $\_\_i$]]]]]]]

Relative clauses are special in comparison to the nonlocal dependencies explained in Section 4.9 in that there are further conditions imposed on the fronted phrase: the fronted element has to contain a relative pronoun. The fronted element can be a relative pronoun as in our example and in (30a) or the relative pronoun may be embedded as in (30b–c):

(30) a. *who* (*she met before*)
b. *whose sister* (*she met before*)
c. *a friend of whose sister* (*she met before*)

[18] Binding Theory is assumed to operate on the list of all arguments of a head, the so-called ARG-ST list (see Section 4.3). An anaphor has to be coindexed with a less oblique element on the ARG-ST list if there is any. For the reflexive pronoun *herself* this means that it has to be coreferential with the subject of *introduce*. For further details see Pollard & Sag (1992) and Branco (2019).

This requirement for a relative pronoun to be present can be analyzed as another non-local dependency and we can use similar techniques to establish the dependency. The referential index of the relative pronoun is shared with an element in the INHER|REL list of the relative pronoun:

(31) SYNSEM value of the relative pronoun *that*:

$$\left[\begin{array}{ll} \text{LOC} & \left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{ll} \text{HEAD} & \textit{noun} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{array}\right] \\ \text{CONT} & \left[\text{IND } \boxed{1}\right] \end{array}\right] \\ \text{NONLOC} & \left[\text{INHER } \left[\text{REL } \langle \boxed{1} \rangle\right]\right] \end{array}\right]$$

In order to enhance readability, the INHER|SLASH and the TO-BIND values are not given, since their value is the empty list.

The lexical entry of the relative pronoun *whose* is rather similar, the only difference being the part of speech for *whose*, which is *det* rather than *noun*. Figure 8 shows the analysis of the noun phrase *a friend of whose sister*.

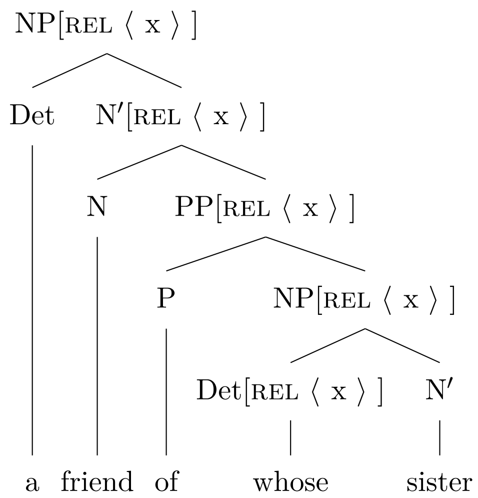

Figure 8: Representation of the referential index of a relative pronoun in the NON-LOC|INHER|REL list

A relative clause differs from an ordinary clause in that it can modify a noun. As the following examples show, ordinary clauses cannot do this:

(32) a. * *The man* [*she had met that before*] *laughs.*

b. * *The man* [*she had met him before*] *laughs.*

This means that the MOD value of verbs – which are the heads of clauses – has to be *none*. But for relative clauses we must have a representation where the relative clause has an $\overline{\mathrm{N}}$ as the MOD value. There are at least three ways to achieve this: one can assume that the MOD value of verbs is a default that can be overridden in the case of relative clauses (Sag 1997), one can assume that relative clauses are the projection of an empty relativizer (Pollard & Sag 1994: 213–217), or one can assume that relative clauses are unheaded (Müller 1999b,a). In the latter proposal a relative phrase is combined with a clause from which it is missing and the result is a relative clause with an appropriately specified MOD value. The latter proposal has the advantage that the special headless construction can also contribute the appropriate semantics. The relative clause *that she had met before* then behaves like other postnominal modifiers and can be adjoined to an $\overline{\mathrm{N}}$. As Sag (1997) points out, an analysis with the verb as head would require different meanings for the verb when used in ordinary clauses in comparison to verbs that are used in relative clauses. Rather than assuming verbs with nominal meanings for modification of $\overline{\mathrm{N}}$s, Sag assumes that relative clauses have verbal meaning and that there is a special schema that licenses the combination of $\overline{\mathrm{N}}$s and relative clauses. The approach suggested here does not need such an additional schema.

The Relative Clause Schema is given as Schema 6:

Schema 6 (Relative Clause Schema)
relative-clause ⇒

$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{rc} & \\ \text{MOD} & \overline{\text{N}}\ \boxed{1} \end{bmatrix} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SS} & \begin{bmatrix} \text{LOC} & \boxed{2} \\ \text{NLOC}|\text{INH} & \begin{bmatrix} \text{REL} & \langle \boxed{1} \rangle \\ \text{SLASH} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{SS} & \begin{bmatrix} \text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{VFORM} & \textit{fin} \end{bmatrix} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \\ \text{NLOC}|\text{TO-BIND} & \begin{bmatrix} \text{REL} & \langle \boxed{1} \rangle \\ \text{SLASH} & \langle \boxed{2} \rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

The schema combines a finite clause containing a gap ($\boxed{2}$) with an appropriate filler, that is, with a filler that has the same LOCAL value as the gap, namely $\boxed{2}$. The filler has to contain a relative pronoun, the referential index of which ($\boxed{1}$) is identified with the referential index of the $\overline{\mathrm{N}}$ that is modified by the relative clause. This schema is really similar to the Filler-Head Schema: the first daughter is the filler that binds off an unbounded dependency. One difference is that the Filler-Head Schema has one of the daughters designated as head daughter while the Relative Clause Schema is unheaded. Both schemata inherit from the same supertype, which states the constraints regarding a filler and the phrase from which it is extracted.

This schema does not include information about semantics but a treatment of semantics can be found in Müller (2013a: Section 11.2.2) and Müller (1999a: Section 2.7). For further details on semantics see also Section 6.5.

6.2 Raising

Our example sentence involves the auxiliary verb *have*. Auxiliaries are raising verbs (Pollard & Sag 1994: 143; Sag, Wasow & Bender 2003: 396; Sag et al. 2020): they do not care for the type of the subject of the embedded verb since they do not assign semantic roles to it. In languages that allow for subjectless constructions (as for instance German) auxiliaries may embed subjectless verbs (Kiss 1995: 87). The trick to describe raising predicates is that arguments may belong to several heads at the same time. So, a raising predicate takes as its subject whatever its complement requires as a subject. This is done with structure sharing (the tool that does most of the work in HPSG, see Section 4.1). (33) shows the lexical item for the finite form *had* of the auxiliary verb *have*:

(33) CAT value for *had*:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} \; \textit{fin} \end{bmatrix} \\ \text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} \; \textit{perf} \end{bmatrix} \\ \text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

The auxiliary selects for a VP (something with an empty COMPS list and one element in the SPR list). The element in the specifier list is identified with the element in the SPR list of the auxiliary.[19] Figure 9 shows the analysis of (34):

(34) *She had met him.*

The perfect participle *met* selects a subject via SPR ($\boxed{1}$) and an object via COMPS ($\boxed{3}$). It is combined with *him*, resulting in a VP. The auxiliary selects for a VP ($\boxed{2}$) and identifies its own element in SPR with the element in SPR of the VP ($\boxed{1}$). The result of the combination of auxiliary and VP is a VP that is still lacking a specifier. Since the specifier of the auxiliary was identified with the specifier of the VP headed by the perfect

[19] In reality this is more indirect: the element in the SPR list of the embedded verb is identified with the first element of the ARG-ST list of the auxiliary. The second element on ARG-ST is the selected VP. The elements on ARG-ST are mapped to SPR and COMPS, resulting in the values provided in (33). We did not include this in (33) for reasons of readability. For further details see Müller & Ørsnes (2013a). This paper also contains a generalized description of auxiliaries that works for both English and German, which allows for frontings of partial verbal projections.

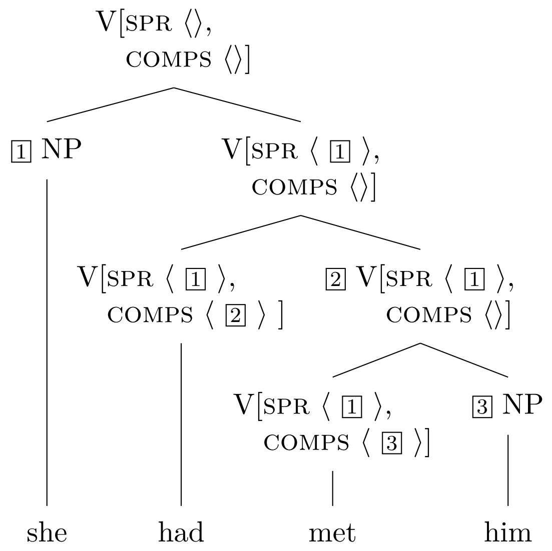

Figure 9: Analysis of sentence with auxiliary and subject raising

participle, the specifier of the complete VP is identical to the one of the VP headed by the perfect participle (1). This way *she* is both the subject of *had* and the subject of *met* although only the latter assigns a semantic role to it.

This kind of raising analysis also works for verbs like *seem* and verbs like *see* and *let* (also know as Exceptional Case Marking verbs). The same technique can be used for the analysis of complex predicates. For instance, for German and Persian it has been suggested that a predicate complex is formed. This can be modeled by not just raising the subject of the embedded predicate but by raising all arguments (Hinrichs & Nakazawa 1994, Müller 2010).

6.3 Case assignment

English has a relatively simple case system and until now we have ignored case in the example analysis. Pronouns do differ according to case and in order to rule out sequences like (35), one has to say something about case.

(35) * *Him likes he.*

Like other theories HPSG distinguishes between structural and lexical cases (Yip, Maling & Jackendoff 1987: 222). Structural cases are those that change according to syntactic environments, lexical cases stay constant. We assume that the nominal arguments of one and two place verbs in English have structural case. The first two arguments of three-place verbs with nominal arguments have structural case and the third one has lexical case since it cannot be promoted to subject and get nominative in passive sentences. The Case Principle (Meurers 1999b, Przepiórkowski 1999) says for verbal environments that the first element with structural case in the ARG-ST list gets nominative and all other elements with structural case get accusative. This works well for our lexical items repeated here as (36) with the case information specified:

(36) lexical items (active):

		ARG-ST
a.	meet	⟨ NP[*str*], NP[*str*] ⟩
b.	turn	⟨ NP[*str*], PP[*to*] ⟩
c.	introduce	⟨ NP[*str*], NP[*str*], PP[*to*] ⟩

The first element in the list is the subject. It has structural case and gets nominative. In cases in which the second element has structural case (see (36a) and (36c) but not (36b)), it gets accusative.

While case assignment by means of the Case Principle sounds straight forward, there are AcI verbs like *see* as in (37) that present a challenge for this version of the principle.

(37) *He saw her laugh.*

laugh is an intransitive verb with a subject argument. This argument is raised to the object of *see*. So it is simultaneously the subject of *laugh* and the object of *see*, which would result in a conflict if nothing special is said about such situations. Obviously, the higher verb should be given priority, so in order to avoid the conflict, the Case Principle has to be reformulated as follows: In verbal domains the first element of the ARG-ST list with structural case gets nominative provided it is not raised. In verbal domains, all other elements of the ARG-ST list with structural case get accusative provided they are not raised.[20] In order to be able to distinguish raised from non-raised elements a Boolean feature[21] RAISED is used. The elements of the ARG-ST list are not *synsem* objects but include *synsem* objects and have more internal structure:

(38) $\begin{bmatrix} \textit{arg} & \\ \text{ARG} & \textit{synsem} \\ \text{RAISED} & \textit{bool} \\ \text{REALIZED} & \textit{bool} \end{bmatrix}$

When an argument is raised from a predicate, the value of RAISED is instantiated as '+'. All those elements that are not raised are marked as RAISED−. Case is assigned to those elements that are RAISED−.

As the following example from Webelhuth (1985: 210) shows, German allows for the fronting of non-finite verb phrases that contain a nominative.

(39) [*Zwei Männer erschossen*] *wurden während des Wochenendes.*
two men.NOM shot were.PL during the weekend
'Two men were shot during the weekend.'

[20] This is basically a reformulation of Yip, Maling & Jackendoff's Case Principle (1987) without overriding case values. They applied their case theory to Icelandic and it comes as no surprise that the monotonic reformulation works for Icelandic as well.

[21] The possible (maximal specific) values of Boolean valued features are '+' and '−'.

The standard analysis of auxiliary-verb combinations in German assumes that auxiliaries form a verbal complex with the embedded verb (Hinrichs & Nakazawa 1994). Auxiliaries attract all arguments of their embedded verbal complements and hence they can assign them case. The problem with examples like (39) is that *zwei Männer erschossen* 'two men shot' forms a phrase and the argument *zwei Männer* is not represented at the mother node and hence, when *wurden* 'were' is combined with *zwei Männer erschossen* 'two men shot', *wurden* 'were' cannot attract the argument of *erschossen* 'shot'. This problem was solved by assuming that elements that get saturated stay in the valence list but are marked as realized (i.e., they have a REALIZED value of '+' rather than '−'). The realized elements are still around on the valence lists, which is the reason why they are called *spirits* (Meurers 1999b). Figure 10 shows the analysis of (34) with spirits (marked with checked off boxes). With this slight change in representation the

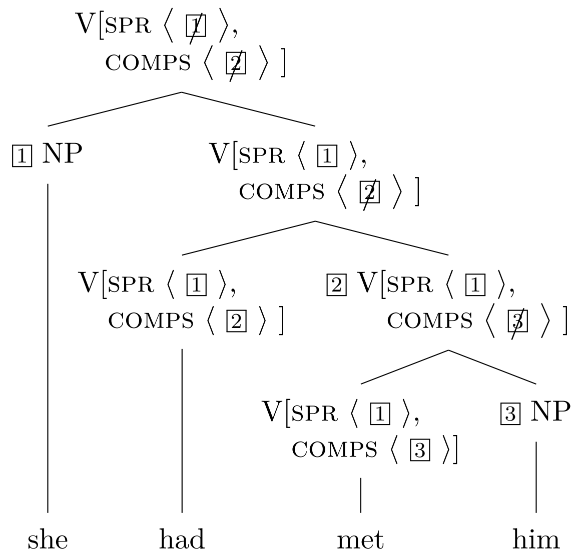

Figure 10: Analysis of sentence with auxiliary and subject raising

argument of *erschossen* 'shot' is still present at *zwei Männer erschossen* 'two men shot' and hence can be attracted by *wurden* 'were'. Since it is the first argument of *wurden* with structural case it is assigned nominative.

6.4 Agreement

A lot can be said about the many faces of agreement. A short overview article is not the right place to do this but the interested reader is referred to a great book by Wechsler & Zlatić (2003). Agreement in English is covered by Pollard & Sag (1994: Chapter 2). The verbs in our example sentence all are in the past tense so that no agreement is visible. But English has subject verb agreement. The subject is the first element of the argument structure list. A more general description of agreement is that the finite verb agrees with a nominative element. Since nominative is assigned to the first element of the ARG-ST list that has structural case (see Section 6.3), this element is the one that agrees with the finite verb. This characterization of verb-nominative agreement works for many languages including Icelandic, where we have quirky case subjects and objects in the nominative. These objects agree with the finite verb as is shown by examples like (40), which is taken from Zaenen, Maling & Thráinsson (1985: 451):

(40) *Hefur henni alltaf þótt Ólafur leibinlegur?*
has she.DAT always thought Olaf.NOM boring.NOM
'Has she always considered Olaf boring?'

Zaenen et al. (1985) show with several tests that *henni* 'she' should be regarded as the subject and *Ólafur* as the object. But agreement is with the nominative element and hence in sentences like (40) with the object. Here, it is assumed that quirky subjects have lexical case and hence an analysis assuming that the first NP with structural case (if there is any) agrees with the finite verb gets the facts right (Müller 2018b).

Again, like with case assignment agreement relations can exist between finite verbs and nominatives that are embedded in a separate phrase (Höhle 1997: 114):

(41) [*Die Hände gezittert*] *haben* / * *hat ihm diesmal nicht.*
the hands trembled have has him this.time not
'His hands did not tremble this time.'

Since the argument of *gezittert* 'trembled' is raised to the auxiliary *haben* 'have' as a spirit, it is accessible to *haben* and the agreement relation can be established. The same explanation works for the passive example in (39): since the nominative is an element of the ARG-ST of the auxiliary, the agreement relation can be established.

6.5 Semantics

Very little has been said about semantics so far since this book is about syntactic approaches. However, HPSG takes the integration of constraints on all linguistic levels seriously and therefore most HPSG analyses cover both syntactic and semantic aspects of the phenomena under consideration (Koenig & Richter 2019).

HPSG started out with Situation Semantics (Barwise & Perry 1983). Later, Minimal Recursion Semantics (Copestake, Flickinger, Pollard & Sag 2005) was developed, which is assumed by many researchers nowadays. We cannot explain the inner workings of MRS but we can point out some of its merits: MRS is a so-called underspecified semantics framework. Scope relations are represented in an underspecified way. So for many situations one gets one semantic representation which stands for several readings.[22] So far the linking between syntax and semantics in lexical items has been explained, see for instance example (14). The semantic representation is contained under CONT and it consists of an index (basically an event variable or a variable for an individual) and a list of relations. The index usually plays a role in these relations. The semantic contribution of a phrase is the concatenation of the RELS lists of its daughters plus a

[22] See for instance Egg (1999) for the three readings of *Max opened all windows again.*

special RELS list that is specified in phrasal schemata. For vanilla schemata like the Head-Complement Schema, the RELS list that is specified by the schema is the empty list. So the Head-Complement Schema is fully compositional in the sense that no extra information is added. However, phrasal schemata can contribute additional relations by this extra RELS list. It is a claim found in much of the Construction Grammar literature that certain phrasal configurations contribute their own meaning. This can be handled easily with HPSG's Semantics Principle, which states that the relations contributed by the mother node are the concatenation of the relations of the daughters plus the relations contributed by the phrasal schema (see for instance Copestake et al. 2005: Section 6.6).

The implemented grammar contains semantic constraints. So, the interested reader may inspect the analyzed example sentence[23] to see the semantic contributions of words and phrases. Click the top-most node in the tree and have a look at CONT, RELS, and HCONS. Click on boxes to inspect their content.

7 Conclusion

This paper introduced the framework of Head-Driven Phrase Structure Grammar (HPSG). We showed how HPSG captures morphology, constituent structure and relations between syntax and semantics by using feature value pairs, identity of values (structure sharing) and relational constraints. Issues of modularity and interfaces and psycholinguistic plausibility have been touched. Since the specification of features and values is purely declarative and no processing regime is associated with the linguistic knowledge it can be used in any direction, that is, it can be used for parsing and generation. This kind of linguistic knowledge is compatible with incremental processing that processes information from all descriptive levels including world knowledge.

The paper analyzes one example sentence in detail. The phenomena discussed involved basic constituent structure, valence, linking, case assignment, agreement, raising, extraction (nonlocal dependencies) and relative clause formation. It has been shown that all relations can be modeled as local relations. In some cases in which transformations/movement are/is suggested in other frameworks, identifying information via structure sharing is used. For instance in raising, arguments can be the arguments of several heads. Nonlocal dependencies are modeled by passing information up until it is bound off by a filler.

Acknowledgements

We thank Bob Borsley for comments on an earlier version of this paper and Bob Levine and Steve Wechsler for discussion. Special thanks go to the editors of this volume and Sam Featherston for detailed comments on earlier versions of this paper.

The students of the first author's MA class on grammar implementation also read this paper. We thank Jakob Wünsch for comments.

[23] https://hpsg.hu-berlin.de/~stefan/Pub/current-approaches-hpsg.html

References

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1. 1–27.

Alexopoulou, Theodora & Dimitra Kolliakou. 2002. On linkhood, topicalization and clitic left dislocation. *J. of Linguistics* 38(2). 193–245.

Alqurashi, Abdulrahman & Robert D. Borsley. 2013. Arabic relative clauses in HPSG. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 25–43. Stanford: CSLI Publications.

Bargmann, Sascha. 2015. *Syntactically Flexible VP-Idioms and the N-after-N Construction.* Poster presentation at the 5th General Meeting of PARSEME, Iasi, 23–24 September 2015.

Barwise, Jon & John Perry. 1983. *Situations and attitudes.* Cambridge, MA: MIT Press.

Bender, Emily M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence.* Stanford University dissertation.

Bender, Emily & Daniel P. Flickinger. 1999. Peripheral constructions and core phenomena: Agreement in tag questions. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 199–214. Stanford: CSLI Publications.

Bildhauer, Felix. 2008. *Representing information structure in an HPSG grammar of Spanish.* Universität Bremen Dissertation.

Bildhauer, Felix. 2011. Mehrfache Vorfeldbesetzung und Informationsstruktur: Eine Bestandsaufnahme. *Deutsche Sprache* 39(4). 362–379.

Bildhauer, Felix & Philippa Cook. 2010. German multiple fronting and expected topichood. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 68–79. Stanford: CSLI Publications.

Bird, Steven. 1995. *Computational phonology: A constraint-based approach* (Studies in Natural Language Processing). Cambridge, UK: Cambridge University Press.

Bird, Steven & Ewan Klein. 1994. Phonological analysis in typed feature systems. *Computational Linguistics* 20(3). 455–491.

Borsley, Robert D. 1987. *Subjects and Complements in HPSG.* Report No. CSLI-87-107. Stanford: Center for the Study of Language & Information.

Borsley, Robert D. 1989. An HPSG approach to Welsh. *J. of Linguistics* 25. 333–354.

Borsley, Robert D. 1999. Mutation and constituent structure in Welsh. *Lingua* 109(4). 267–300.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language and Linguistic Theory* 27(2). 225–265.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Branco, António. 2019. Binding theory. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). To appear. Berlin: Language Science Press.

Brants, Sabine, Stefanie Dipper, Peter Eisenberg, Silvia Hansen-Schirra, Esther König, Wolfgang Lezius, Christian Rohrer, George Smith & Hans Uszkoreit. 2004. TIGER: Linguistic interpretation of a German corpus. *Research on Language and Computation* 2(4). 597–620.

Bresnan, Joan. 1978. A realistic Transformational Grammar. In Morris Halle, Joan Bresnan & George A. Miller (eds.), *Linguistic theory and psychological reality*, 1–59. Cambridge, MA: MIT Press.

Bresnan, Joan & Ronald M. Kaplan. 1982. Introduction: grammars as mental representations of language. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), xvii–lii. Cambridge, MA/London: MIT Press.

Butt, Miriam. 1995. *The structure of complex predicates in Urdu* (Dissertations in Linguistics). Stanford: CSLI Publications.

Carpenter, Bob. 1992. *The logic of typed feature structures* (Tracts in Theoretical Computer Science). Cambridge: Cambridge University Press.

Chomsky, Noam. 1957. *Syntactic structures* (Janua Linguarum / Series Minor 4). The Hague/Paris: Mouton.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of mathematical psychology*, vol. 2, 269–321. New York: Wiley.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332.

Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics in information-based morphology. *Journal of Linguistics* 52(2). 311–374.

De Kuthy, Kordula. 2002. *Discontinuous NPs in German* (Studies in Constraint-Based Lexicalism 14). Stanford: CSLI Publications.

Dowty, David R. 1978. Governed transformations as lexical rules in a Montague Grammar. *Linguistic Inquiry* 9(3). 393–426.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Egg, Markus. 1999. Derivation and resolution of ambiguities in *wieder*-sentences. In Paul J. E. Dekker (ed.), *Proceedings of the 12th Amsterdam Colloquium*, 109–114.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Claire Grover & Enric Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science, vol. 12: Studies in HPSG*, chap. 1, 1–32. Edinburgh: Centre for Cognitive Science, University of Edinburgh.

Flickinger, Dan, Stephan Oepen & Gisle Ytrestøl. 2010. Wikiwoods: syntacto-semantic annotation for english wikipedia. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 1665–1671. Valletta, Malta: European Language Resources Association (ELRA).

Flickinger, Daniel P., Ann Copestake & Ivan A. Sag. 2000. HPSG analysis of English. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence), 254–263. Berlin: Springer Verlag.

Flickinger, Daniel P., Carl J. Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In William C. Mann (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 262–267. Chicago, IL.

Fodor, Jerry A., Thomas G. Bever & Merrill F. Garrett. 1974. *The psychology of language: An introduction to psycholinguistics and Generative Grammar*. New York: McGraw-Hill Book Co.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Ginzburg, Jonathan & Robin Cooper. 2004. Clarification, ellipsis, and the nature of contextual updates in dialogue. *Linguistics and philosophy* 27(3). 297–365.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford: CSLI Publications.

Goldberg, Adele E. 1996. Words by default: Optimizing constraints and the Persian complex predicate. *Berkeley Linguistic Society* 22. 132–146.

Goldberg, Adele E. 2003. Words by default: The Persian Complex Predicate Construction. In Elaine J. Francis & Laura A. Michaelis (eds.), *Mismatch: form-function incongruity and the architecture of grammar* (CSLI Lecture Notes 163), 117–146. Stanford: CSLI Publications.

Halliday, Michael A. K. 1970. *A course in spoken English: intonation*. Oxford University Press.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 11–38. Stanford: CSLI Publications.

Höhle, Tilman N. 1997. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Christa Dürscheid, Karl Heinz Ramers & Monika Schwarz (eds.), *Sprache im Fokus: Festschrift für Heinz Vater zum 65. Geburtstag*, 107–120. Republished as Höhle (2018b). Tübingen: Niemeyer.

Höhle, Tilman N. 1999. An architecture for phonology. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in Head-Driven Phrase Structure Grammar*, 61–90. Republished as Höhle (2018a). Stanford: CSLI Publications.

Höhle, Tilman N. 2018a. An architecture for phonology. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften* *von Tilman N. Höhle* (Classics in Linguistics 5), 571–607. Originally published as Höhle (1999). Berlin: Language Science Press.

Höhle, Tilman N. 2018b. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 417–433. First published as Höhle (1997). Berlin: Language Science Press.

Jackendoff, Ray S. 2008. Construction after Construction and its theoretical challenges. *Language* 84(1). 8–28.

Jacobson, Pauline. 1987. Review of Generalized Phrase Structure Grammar. *Linguistics and Philosophy* 10(3). 389–426.

Johnson, Mark. 1988. *Attribute-value logic and the theory of grammar* (CSLI Lecture Notes 16). Stanford: CSLI Publications.

Kathol, Andreas. 2001. Positional effects in a monostratal grammar of German. *J. of Linguistics* 37(1). 35–66.

Kay, Paul, Ivan A. Sag & Daniel P. Flickinger. 2015. *A Lexical Theory of Phrasal Idioms.* Ms. CSLI Stanford.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

King, Paul. 1994. *An Expanded Logical Formalism for Head-Driven Phrase Structure Grammar.* Arbeitspapiere des SFB 340 No. 59. Tübingen: Eberhard-Karls-Universität.

King, Paul. 1999. Towards truth in Head-Driven Phrase Structure Grammar. In Valia Kordoni (ed.), *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitsberichte des SFB 340 No. 132), 301–352. Tübingen: Universität Tübingen.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Niemeyer.

Klein, Ewan. 2000. A constraint-based approach to English prosodic constituents. In *38th Annual Meeting of the Association for Computational Linguistics: Proceedings of the conference*, 217–224. Hong Kong: ACL.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). To appear. Berlin: Language Science Press.

Kuhn, Jonas. 1996. An underspecified HPSG representation for information structure. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96. 16th International Conference on Computational Linguistics (COLING96). Copenhagen, Denmark, August 5–9, 1996*, 670–675. Copenhagen, Denmark: ACL.

Kuhn, Jonas. 2007. Interfaces in constraint-based theories of grammar. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 613–650. Oxford: Oxford University Press.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions* (CSLI Lecture Notes 166). Stanford: CSLI Publications.

Meurers, Walt Detmar. 1999a. German partial-VP fronting revisited. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of* *linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 129–144. Stanford: CSLI Publications.

Meurers, Walt Detmar. 1999b. Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43. 173–226.

Meurers, Walt Detmar & Stefan Müller. 2009. Corpora and syntax. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, vol. 29 (Handbücher zur Sprach- und Kommunikationswissenschaft), chap. 42, 920–933. Berlin: Mouton de Gruyter.

Meurers, Walt Detmar, Gerald Penn & Frank Richter. 2002. A web-based instructional platform for constraint-based grammar formalisms and parsing. In Dragomir Radev & Chris Brew (eds.), *Effective tools and methodologies for teaching NLP and CL*, 18–25. Proceedings of the Workshop held at 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, PA. ACL.

Müller, Stefan. 1999a. An HPSG-analysis for free relative clauses in German. *Grammars* 2(1). 53–105.

Müller, Stefan. 1999b. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Niemeyer.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford: CSLI Publications.

Müller, Stefan. 2003. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *J. of Linguistics* 39(2). 275–325.

Müller, Stefan. 2007. Qualitative Korpusanalyse für die Grammatiktheorie: Introspektion vs. Korpus. In Gisela Zifonun & Werner Kallmeyer (eds.), *Sprachkorpora – Datenmengen und Erkenntnisfortschritt* (Institut für Deutsche Sprache Jahrbuch 2006), 70–90. Berlin: Walter de Gruyter.

Müller, Stefan. 2009. On predication. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 213–233. Stanford: CSLI Publications.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655.

Müller, Stefan. 2012. *On the Copula, Specificational Constructions and Type Shifting*. Ms. Freie Universität Berlin.

Müller, Stefan. 2013a. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 3rd edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2013b. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950.

Müller, Stefan. 2014a. Elliptical constructions, multiple frontings, and surface-based syntax. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 91–109. Stanford: CSLI Publications.

Müller, Stefan. 2014b. Kernigkeit: Anmerkungen zur Kern-Peripherie-Unterscheidung. In Antonio Machicao y Priemer, Andreas Nolda & Athina Sioupi (eds.), *Zwischen Kern und Peripherie* (studia grammatica 76), 25–39. Berlin: de Gruyter.

Müller, Stefan. 2015a. HPSG – A synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), 937–973. Berlin: Walter de Gruyter.

Müller, Stefan. 2015b. The CoreGram project: Theoretical linguistics, theory development and verification. *J. of Language Modelling* 3(1). 21–86.

Müller, Stefan. 2016. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press.

Müller, Stefan. 2017. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Revise and resubmit. Berlin: Language Science Press.

Müller, Stefan. 2018a. *A lexicalist account of argument structure: Template-based phrasal LFG approaches and a lexical HPSG alternative* (Conceptual Foundations of Language Science 2). Berlin: Language Science Press.

Müller, Stefan. 2018b. *Germanic Syntax*. Ms. Humboldt Universität zu Berlin, to be submitted to Language Science Press. Berlin.

Müller, Stefan. 2018c. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 2nd edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press.

Müller, Stefan. 2019. Hpsg and construction grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). To appear. Berlin: Language Science Press.

Müller, Stefan, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.). 2019. *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). To appear. Berlin: Language Science Press.

Müller, Stefan & Walter Kasper. 2000. HPSG analysis of German. In Wolfgang Wahlster (ed.), *Verbmobil: foundations of speech-to-speech translation* (Artificial Intelligence), 238–253. Berlin: Springer Verlag.

Müller, Stefan & Bjarne Ørsnes. 2013a. Passive in Danish, English, and German. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 140–160. Stanford: CSLI Publications.

Müller, Stefan & Bjarne Ørsnes. 2013b. Towards an HPSG analysis of object shift in Danish. In Glyn Morrill & Mark-Jan Nederhof (eds.), *Formal Grammar: 17th and 18th International Conferences, FG 2012, Opole, Poland, August 2012, revised selected papers, FG 2013, Düsseldorf, Germany, August 2013: Proceedings* (Lecture Notes in Computer Science 8036), 69–89. Berlin: Springer Verlag.

Müller, Stefan & Bjarne Ørsnes. 2015. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). In Preparation. Berlin: Language Science Press.

Müller, Stefan, Marga Reis & Frank Richter (eds.). 2018. *Beiträge zur deutschen Grammatik: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5). Berlin: Language Science Press.

Oepen, Stephan & Daniel P. Flickinger. 1998. Towards systematic grammar profiling: Test suite technology ten years after. *Journal of Computer Speech and Language* 12(4). (Special Issue on Evaluation), 411–436.

Oepen, Stephan, Erik Velldal, Jan Tore Lønning, Paul Meurer, Victoria Rosén & Dan Flickinger. 2007. Towards hybrid quality-oriented machine translation: On linguistics and probabilities in MT. In *Proceedings of 11th conference on theoretical and methodological issues in machine translation*, 144–153. Skövde, Sweden.

Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology*. University of California, Berkeley dissertation.

Paggio, Patrizia. 2005. Representing information structure in a formal grammar of Danish. In Takashi Washio, Akito Sakurai, Katsuto Nakajima, Hideaki Takeda, Satoshi Tojo & Makoto Yokoo (eds.), *New frontiers in artificial intelligence: Joint JSAI 2005 Workshop post-proceedings* (Lecture Notes in Computer Science 4012), 93–102. Berlin: Springer Verlag.

Penn, Gerald. 2004. Balancing clarity and efficiency in typed feature logic through delaying. In Donia Scott (ed.), *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), main volume*, 239–246. Barcelona, Spain.

Pollard, Carl J. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and natural language*. Stanford University dissertation.

Pollard, Carl J. 1996. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Veröffentlichte Version eines Ms. von 1990. Berlin: Mouton de Gruyter.

Pollard, Carl J. 1999. Strong generative capacity in HPSG. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 281–298. Stanford: CSLI Publications.

Pollard, Carl J. & Andrew M. Moshier. 1990. Unifying partial descriptions of sets. In Philip P. Hanson (ed.), *Information, language and cognition* (Vancouver Studies in Cognitive Science 1), 285–322. Vancouver: University of British Columbia Press.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago, IL & London: University of Chicago Press.

Przepiórkowski, Adam. 1999. On case assignment and "adjuncts as complements". In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 231–245. Stanford: CSLI Publications.

Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between generative-enumerative and model-theoretic syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retor (eds.), *Logical Aspects of Computational Linguistics: 4th International Conference* (Lecture Notes in Computer Science 2099), 17–43. Berlin: Springer Verlag.

Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15), 171–211. Tübingen: Gunter Narr Verlag.

Richards, Marc. 2015. Minimalism. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42.3 (Handbooks of Linguistics and Communication Science), 803–839. Berlin: Mouton de Gruyter.

Richter, Frank. 2004. *A mathematical formalism for linguistic theories with an application in Head-Driven Phrase Structure Grammar.* Eberhard-Karls-Universität Tübingen Phil. Dissertation (2000).

Richter, Frank. 2007. Closer to the truth: A new model theory for HPSG. In James Rogers & Stephan Kepser (eds.), *Model-theoretic syntax at 10 – Proceedings of the ESSLLI 2007 MTS@10 Workshop, August 13–17*, 101–110. Dublin: Trinity College Dublin.

Richter, Frank & Manfred Sailer. 2009. Phraseological clauses as Constructions in HPSG. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 297–317. Stanford: CSLI Publications.

Riehemann, Susanne. 1997. Idiomatic constructions in HPSG. Paper presented at the 1997 HPSG conference, Ithaca.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford: CSLI Publications.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford: CSLI Publications.

Sag, Ivan A., Rui Chaves, Anne Abeillé, Bruno Estigarribia, Frank Van Eynde, Dan Flickinger, Paul Kay, Laura A. Michaelis-Cummings, Stefan Müller, Geoffrey K. Pullum & Tom Wasow. 2020. Lessons from the English auxiliary system. *J. of Linguistics* 56. Ahead of print (online).

Sag, Ivan A. & Janet Dean Fodor. 1994. Extraction without traces. In Raul Aranovich, William Byrne, Susanne Preuss & Martha Senturia (eds.), *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, 365–384. Stanford: CSLI Publications/SLA.

Sag, Ivan A., Philip Hofmeister & Neal Snider. 2007. Processing complexity in subjacency violations: The Complex Noun Phrase Constraint. In Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki & Suwon Yoon (eds.), *Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society*, 215–229. Chicago, IL: Chicago Linguistic Society.

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal* *and explicit models of grammar: A guide to current models*, 359–377. Oxford, UK/ Cambridge, MA: Blackwell Publishing Ltd.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI Lecture Notes 152). Stanford: CSLI Publications.

Schäfer, Roland & Felix Bildhauer. 2012. Building large corpora from the web using a new effcient tool chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, 486–493. Istanbul, Turkey: European Language Resources Association (ELRA).

Schlangen, David, Alex Lascarides & Ann Copestake. 2003. Resolving underspecification using discourse information. *Perspectives on Dialogue in the New Millennium* 114. 287.

Shieber, Stuart M. 1986. *An introduction to unification-based approaches to grammar* (CSLI Lecture Notes 4). Stanford: CSLI Publications.

Siegel, Melanie. 2000. HPSG analysis of Japanese. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence), 264–279. Berlin: Springer Verlag.

Soehn, Jan-Philipp & Manfred Sailer. 2008. At first blush on tenterhooks: About selectional restrictions imposed by nonheads. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2003, Vienna, Austria*, 149–161. Stanford: CSLI Publications.

Steedman, Mark J. & Jason Baldridge. 2006. Combinatory Categorial Grammar. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., 610–621. Oxford: Elsevier.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634.

Wahlster, Wolfgang (ed.). 2000. *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence). Berlin: Springer Verlag.

Wasow, Thomas. 2019. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). To appear. Berlin: Language Science Press.

Webelhuth, Gert. 1985. German is configurational. *The Linguistic Review* 4(3). 203–246.

Webelhuth, Gert, Jean-Pierre Koenig & Andreas Kathol (eds.). 1999. *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1). Stanford: CSLI Publications.

Wechsler, Stephen Mark & Larisa Zlatić. 2003. *The many faces of agreement* (Stanford Monographs in Linguistics). Stanford: CSLI Publications.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). To appear. Berlin: Language Science Press.

Wilcock, Graham. 2005. Information structure and Minimal Recursion Semantics. In Antti Arppe, Lauri Carlson, Krister Lindén, Jussi Piitulainen, Mickael Suominen, Martti Vainio, Hanna Westerlund & Anssi Yli-Jyrä (eds.), *Inquiries into words, constraints and contexts: Festschrift for Kimmo Koskenniemi on his 60th birthday* (CSLI Studies in Computational Linguistics ONLINE), 268–277. Stanford: CSLI Publications.

Wittenberg, Eva, Ray S. Jackendoff, Gina Kuperberg, Martin Paczynski, Jesse Snedeker & Heike Wiese. 2014. The processing and representation of light verb constructions. In Asaf Bachrach, Isabelle Roy & Linnaea Stockall (eds.), *Structuring the argument* (Language Faculty and Beyond 10), 61–80. Amsterdam: John Benjamins Publishing Co.

Wittenberg, Eva & Maria Mercedes Piñango. 2011. Processing light verb constructions. *The Mental Lexicon* 6(3). 393–413.

Yip, Moira, Joan Maling & Ray S. Jackendoff. 1987. Case in tiers. *Language* 63(2). 217–250.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3(4). 441–483.