Chapter 4

Headless in Berlin: Headless (nominal) structures in Head-Driven Phrase Structure Grammar

Stefan Müller ⓘ

Humboldt-Universität zu Berlin

This paper deals with the status of heads in Head-Driven Phrase Structure Grammar (HPSG). Firstly, background assumptions are presented: the lexical representation of valence at the head and projection of head features. Secondly, I discuss criteria for determining the head of a phrase. I use nominal structures as an example, since the DP/NP debate is still undecided across frameworks, and exploring the arguments from an HPSG perspective may be interesting for readers. Zwicky's (1985) criteria are discussed, and I show that most of them do not decide the issue for German nominal structures, but assignment of semantic roles by relational nouns and selectional relations in idioms (Osborne & Groß 2012, Bruening 2020) support NP structures. I discuss nominal structures with non-overt nouns and copulaless sentences in African American Vernacular English (AAVE) and argue for an empty nominal head. I show that empty elements can be eliminated from grammars but argue that they are nevertheless useful in nominal structures and copula constructions in AAVE, since they capture generalizations. However, there are other structures like Jackendoff's (2008) N-P-N construction that should be analyzed as unheaded. The paper closes with general considerations about the use of empty elements in grammars, arguing that they should be detectable in the input by systematic variation with overt material. This excludes the assumption of empty elements like AgrO or Topic in grammars of languages like German, since there is no overt material associated with these heads.

1 Introduction

Ulrike Freywald and Horst Simon asked proponents of various linguistic theories to take part in their workshop *Headedness and/or grammatical anarchy?* and ex-

plain the notion of head used in the respective theories. They asked the following questions:

- Are structural asymmetries a precondition for structure building?

- Or do "real" non-headed structures exist?

- If so, how are non-headed structures built?

- How does headedness/a headed structure work, if there is no head?

- Do we need the concept of "head" in grammatical theory?

The current paper addresses these questions. I start with an introduction of Head-Driven Phrase Structure Grammar (HPSG, Pollard & Sag 1987, 1994, Müller 2013) in Section 2; there, I explain how lexical heads determine the internal structure and external distribution of phrases. Following the introduction of the basic machinery in Section 2, I step back a bit and discuss more general, theory-neutral criteria for an element being a head in Section 3. Zwicky's (1985) criteria for being a head are applied in the notorious DP/NP debate. German data shows that most of the criteria deliver inconclusive results, but some seem to argue for N as the head. After comparing the complexity of NP and DP structures and discussing the assignment of semantic roles in nominalizations, selection, and idioms, I argue for assuming N as the head in nominal structures.

Section 4 deals with the question of how to deal with structures in which there is no visible head. Again, I discuss nominal structures and show how nounless nominal structures can be described by assuming an empty nominal head. Furthermore, I explain the analysis of predicative structures in African American Vernacular English and why the assumption of an empty head was suggested in Sag et al. (2003: Section 15.3.5).

Frameworks like Construction Grammar reject empty elements dogmatically (Goldberg 2003: 219, 2006: 10, Hoffmann & Trousdale 2013: 3, Fillmore 2013: 112, Michaelis 2013: 134) since they are said to be unacquirable. I show in Section 5 that grammars with empty elements may be transformed into grammars without empty elements, and I argue that the NP grammar with empty nominal heads is in fact easier to learn than the grammar without empty elements, since it captures the facts about omissible elements directly.

Apart from nominal structures in which we have a head but it is invisible, there are other structures in which it is impossible to identify one central element that determines the structure of the whole unit and where the stipulation of an empty head cannot be motivated by anything theory-external. Section 6 shows how such phrases can be analyzed and why they are unproblematic for HPSG even

though the theory has "head-driven" in its name, which seems to suggest that all structures have to have a head.

Section 7 discusses language acquisition and provides criteria for when the assumption of empty elements is appropriate. Section 8 provides a summary of the paper.

2 Heads and HPSG

The notion of head is crucial for Head-Driven Phrase Structure Grammar: most phrases in grammars have a central element that is responsible for the internal structure of the phrase and for its distribution. For example, prepositions determine the case of the NP they combine with:

(1) a. zu diesem   Termin
         to this.DAT appointment

     b. * zu diesen   Termin
           to this.ACC appointment

The form of the preposition in prepositional objects is important, since it is responsible for the external distribution of the whole phrase: while an *auf* PP can function as the object of *warten* 'to wait', an *an* PP cannot:

(2) a. Ich warte auf den Mann.
         I   wait on the man
         'I am waiting for the man.'

     b. * Ich warte an den Mann.
           I   wait at the man

It is clear for prepositions, verbs, and adjectives that they have valence and that their form and/or inflectional properties are responsible for the distribution of the whole phrase. The problematic cases (determiner and noun) and respective criteria for head status are discussed in Section 3.

I turn now to the foundational assumptions of HPSG (treatment of valence and percolation of head information) in order to be able to explain headless structures with reference to these more common structures.

In HPSG, valence information is expressed by means of lists. For example, valence lists of two-place verbs contain two elements (in German).[1] One of these arguments gets the nominative and the other one the accusative case.

---

[1]It is commonly assumed that finite verbs in OV languages select all their arguments in one valence list (Pollard 1996: 295–296, Kiss 1995: Section 3.1.1, Müller 2002), while there are two valence lists for SVO languages like English: one list for preverbal arguments (specifiers) and one for post-verbal arguments (complements; Sag, Wasow & Bender 2003: Section 4.3; Müller 2022: Section 4.3).

I assume binary branching structures for German, as most authors working on German in HPSG do (see for instance Hinrichs & Nakazawa 1994, Kiss 1995, Meurers 1999, Müller 1999b, Kathol 2000, Holler-Feldhaus 2001). Which argument is combined with the head is not constrained, so the head can combine with the nominative or with the accusative first. The argument that is not combined with the head is passed up in the tree. Figure 1 shows this for the example in (3):

(3) [dass] niemand ihn kennt
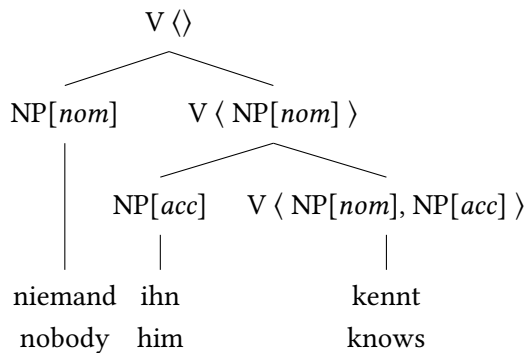   that   nobody   him knows
     'that nobody knows him'

$$V \langle \rangle$$

NP[*nom*]   V ⟨ NP[*nom*] ⟩

NP[*acc*]   V ⟨ NP[*nom*], NP[*acc*] ⟩

niemand   ihn   kennt
nobody   him   knows

Figure 1: Analysis of *niemand ihn kennt* 'nobody knows him'

The verb *kennt* 'knows' requires one NP in the nominative and one in the accusative. After combining *kennt* 'knows' with the accusative object *ihn* 'him', one gets a linguistic object that requires an NP in the nominative. If this linguistic object is combined with the nominative, an element with an empty valence list results. Since the head of this linguistic object is a verb, the whole linguistic object is a sentence.

Many theoretical papers discuss tree structures without providing the rules that actually license the trees. HPSG uses abstract dominance schemata to license linguistic objects. A representation of such a schema is shown in Figure 2.
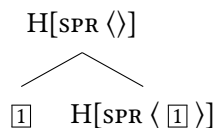
H[SPR ⟨⟩]

1   H[SPR ⟨ 1 ⟩]

Figure 2: Visualization of the Specifier-Head Schema

This treelet shows how heads can be combined with an element of their specifier list (SPR stands for specifier). Usually the SPR list of a head contains exactly one element (the subject of SVO languages and the determiner in NP structures[2], ☐ in Figure 2), which means that the SPR list of the mother node is the empty list.[3]

Heads are marked by H in the figures. This is supposed to indicate that all head information, that is, information that is relevant for the distribution of the phrase, is present at both the head daughter and the mother. For sentences, this would be the information that the part of speech of the head is *verb* and whether the verb is finite, a participle, or an infinitive with or without *to*. Of course other information about required arguments, extracted arguments and adjuncts, and

---

[2]Some authors assume an additional valence feature for subjects, namely SUBJ (Borsley 1987, Pollard & Sag 1994: Chapter 9). I assume a head feature SUBJ for control and raising. Subjects in SVO languages like English and the Scandinavian languages are treated as specifiers (Müller 2022). There are analyses of the NP assuming that the determiner is a marker of the head rather than a dependent selected via valence features (Van Eynde 2006, Allegranza 2007, Sag 2012). See also Van Eynde (2021) for an overview of alternative approaches to nominal structures within HPSG. The marker-based approaches provide a simple analysis of determinerless nominal structures (Van Eynde 2006: 167, 174–175), but the syntactic simplicity of syntactic structures comes at a price: the resulting structures are missing the quantifier usually contributed by the determiner. The only solution to the problem I am aware of was suggested by Allegranza (1998). Allegranza (1998) suggests an analysis in which nouns that may appear without determiner (plurals and mass nouns) introduce a quantifier lexically. If these nouns are used without a determiner or a quantifier, the lexically introduced quantifier is used. In all other cases the lexically introduced quantifier is removed (p. 103). The solution involves disjunctions and subtraction operations over sets and is rather complex. Furthermore, Allegranza's account fails on examples like *alledged water*, since in his setup the quantifier scopes over *water* directly. While this probably can be fixed, any imaginable solution is probably not simpler than what was suggested so far and hence I prefer the approach described in this paper.

In addition, marker-based approaches do not capture the parallelism between verbs and nouns in nominalizations like (i):

(i) a. Caesar destroyed the city.
     b. Caesar's destruction of the city

In the approach assumed here, both the subject *Caesar* and the determiner phrase *Caesar's* will be selected by the verb and by the noun derived from the verb, respectively. The heads will assign semantic roles to the selected element (Machicao y Priemer & Müller 2021). In the marker-based analysis *Caesar's* would select *destruction of the city* despite the fact that it fills a semantic role of *destruction* (Frank Van Eynde, p.c. 2019). I prefer the more uniform analysis of the examples in (i). Nominalizations will be discussed further in Section 3.2.3.

[3]See Müller & Ørsnes (2013) for an analysis of object shift in Danish in which objects appearing to the left of the verb (like subjects) are treated as specifiers. In this analysis the SPR list may contain more than one element. See also Ng (1997: Sections 5.3, 6.3) for an analysis of nominal structures in English and Chinese with multiple specifiers and Wang & Liu (2007) for such an analysis of Chinese nominal structures.

relative pronouns within a phrase, among other things, are also relevant for the external distribution of a phrase. This information is part of the complex categories that are assumed in HPSG. The head information is the information that is directly shared between lexical heads and intermediate and maximal projections of the lexical head.

The trees we saw so far are convenient for visualization, but HPSG uses typed feature value structures to model all aspects of linguistic objects; even the internal configuration of complex syntactic objects is represented by feature value pairs. (4) shows how the Specifier-Head Schema can be described with feature value pairs.[4]

(4)   Specifier-Head Schema:
       specifier-head-phrase ⇒
       $$\begin{bmatrix} \text{SPR} & \langle \rangle \\ \text{HEAD-DTR} & \boxed{1} \left[\text{SPR} \langle \boxed{2} \rangle \right] \\ \text{DTRS} & \langle \boxed{1}, \boxed{2} \rangle \end{bmatrix}$$

The symbol ⇒ stands for a logical implication: if a feature structure is of type *specifier-head-phrase*, the restrictions on the right side of the implication have to hold. Types and implicational constraints are discussed further below. The daughters in a tree are represented in a list, which is the value of the DAUGHTERS feature (Ginzburg & Sag 2000: 30). In the case at hand, we have two daughters: $\boxed{1}$ and $\boxed{2}$. The daughter $\boxed{1}$ is the head daughter. In addition to being in the DTRS list, it is identified with the value of the feature HEAD-DTR. The SPR list contains a description of the other daughter ($\boxed{2}$). Figure 3 shows the tree representation for structures licensed by the schema in (4). *Mannes* 'man' selects a determiner.

Verbs project information about part of speech and inflection. The part of speech and case information determines the distribution of nominal projections. HPSG groups information that belongs together into one attribute value matrix (AVM). Part of speech information and case information form the value of HEAD in (5).[5] *noun* is the type of the feature description. Feature structures of type *noun*

---

[4]Even though the schema is more formal than the little treelets, it is still a simplification in that not all feature-value paths are fully specified as they would have to be according to the theory. (4) leaves out the paths leading to SPR and a path in the list of daughters. For details see Ginzburg & Sag (2000: 30) and Müller & Machicao y Priemer (2019).

[5]Of course other properties like number, gender, and declension class are relevant for the distribution as well. Some authors bundle case, person, number, gender, and declension class as agreement features inside of HEAD (Kathol 1999: 262), and others refer to the number and gender information contained in the semantic index contributed by nouns (Pollard & Sag 1994: Section 2.5.1, Müller 2007b: Section 13.2). I omit declension class here, since it is not relevant for the current discussion.
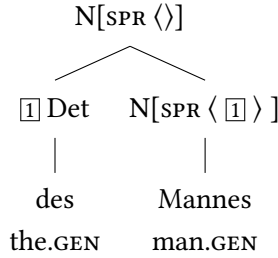
N[SPR ⟨⟩]

⟨1⟩ Det   N[SPR ⟨ ⟨1⟩ ⟩]

des   Mannes

the.GEN   man.GEN

Figure 3: Specifier-Head structure

always have a CASE feature. This feature may have the values *nom, gen, dat,* or *acc* in German. For *Frau* 'woman', we may leave the value underspecified, since *Frau* is compatible with any of the four cases in German, but for *Mannes* 'man', which is in the genitive, it has to be *gen*.

$$
(5) \quad
\begin{bmatrix}
\text{PHON} & \langle\ Mannes\ \rangle \\
\text{HEAD} & \begin{bmatrix} noun \\ \text{CASE}\ gen \end{bmatrix} \\
\text{SPR} & \langle\ \text{Det}\ \rangle \\
\text{COMPS} & \langle\ \rangle
\end{bmatrix}
$$

Since HPSG allows values of features to be internally complex, features that have to be projected from lexical items can be grouped together and the projection of head features can be set up in a general way: all information that is present under HEAD is shared between head daughter and mother, that is, the information in the description of the head daughter is identical to the respective information at the mother. Figure 4 shows the analysis of the nominal phrase *des Mannes* 'the.GEN man.GEN'.

$$\begin{bmatrix} \text{HEAD}\ ⟨1⟩ \end{bmatrix}$$

Det $\begin{bmatrix} \text{HEAD}\ ⟨1⟩\ \begin{bmatrix} noun \\ \text{CASE}\ gen \end{bmatrix} \end{bmatrix}$

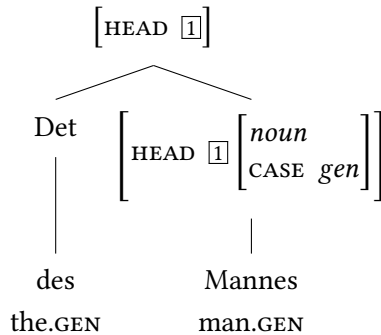des   Mannes

the.GEN   man.GEN

Figure 4: Specifier-Head structure

The information concerning part of speech and case of the noun is shared with the respective information for the whole phrase.

Apart from Specifier-Head structures, there are also Head-Complement structures. These play a role in the combination of relational nouns with their complements. Figure 5 gives an example.

$$
\begin{bmatrix} \text{HEAD } \boxed{1} \\ \text{COMPS } \langle \rangle \end{bmatrix}
$$

$$
\begin{bmatrix} \text{HEAD } \boxed{1}\ noun \\ \text{COMPS } \langle\ \boxed{2}\ \rangle \end{bmatrix} \qquad \boxed{2}\ \text{NP}[gen]
$$

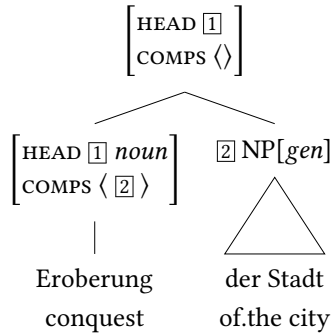|	
Eroberung	der Stadt
conquest	of.the city

Figure 5: Head-Complement structure

Head-Complement phrases are parallel to Specifier-Head phrases. The only difference is that complements are selected via another valence feature (COMPS rather than SPR). The schema for head-complement combinations that is parallel to Figure 2 is shown in Figure 6. As with Specifier-Head phrases, the valence
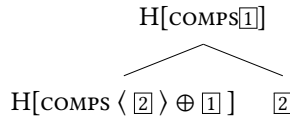
$$
\text{H}[\text{COMPS}\boxed{1}]
$$

$$
\text{H}[\text{COMPS } \langle\ \boxed{2}\ \rangle \oplus \boxed{1}\ ] \qquad \boxed{2}
$$

Figure 6: Visualization of the Head-Complement Schema

list is split into two parts: one list with exactly one element $\langle\ \boxed{2}\ \rangle$ and another list with the rest ($\boxed{1}$). $\boxed{2}$ is identified with the other daughter, and $\boxed{1}$, the list containing the rest, is identified with the COMPS value of the mother node. The combination of a head with its complements as it is given in Figure 6 combines the head with the first element in the valence list. This is exactly what we find in SOV languages. For SOV languages and languages with scrambling see (Müller 2020: Sections 9.1.1, 9.4, 2021a: Sections 3, 4). The order of combination differs from the one in Specifier-Head structures, in which the head is combined with the last element in the SPR list first. Again see Müller & Ørsnes (2013) for details.

(6) shows the schema that licenses Head-Complement phrases. Apart from the schemata introduced so far, HPSG has schemata for head-adjunct combinations

(6) Head-Complement Schema:
     head-complement-phrase ⇒

$$\begin{bmatrix} \text{COMPS} & \boxed{1} \\ \text{HEAD-DTR} & \boxed{2}\begin{bmatrix} \text{COMPS } \boxed{1} \oplus \langle\,\boxed{3}\,\rangle \end{bmatrix} \\ \text{DTRS} & \langle\,\boxed{2},\boxed{3}\,\rangle \end{bmatrix}$$

and for nonlocal dependencies (for less general schemata see Section 6). The two schemata above are sufficient to be able to explain the assumptions about heads and headedness made in HPSG.[6]

All feature structures in HPSG have to be of a certain type. These types are organized in hierarchies. All feature structures modeling linguistic signs are of type *sign*. Linguistic signs are divided into phrases and words. For these objects, we have the types *phrase* and *word*. Phrases can be categorized into phrases that have a head (*headed-phrase*) and phrases without a head (*non-headed-phrase*). *specifier-head-phrase* and *head-complement-phrase* are subtypes of the type *headed-phrase*.

We want to say the following about structures: if there is a head in the structure, then the head features of the head daughter have to be identical to the head features of the mother. HPSG allows for an elegant expression of this fact using an implicational constraint:[7]

---

[6]A reviewer asked how agreement between determiners, adjectives, and nouns can be accounted for in HPSG. These items agree in case, number, gender, and match in declension class. Agreement is usually analyzed using structure sharing of features of items that select others/are selected by others. Since nouns select their determiners, agreement between nouns and determiners can be assured. Similarly, adjectives select nouns, and hence adjective-noun agreement can be taken care of. Since nouns agree with their determiners, the agreement between all three elements is accounted for. See Pollard & Sag (1994: Section 2.5.1) and Müller (2007b: Section 13.2) for worked-out proposals for agreement in German noun phrases and Wechsler & Zlatić (2003) on agreement in HPSG in general.

[7]An alternative formulation of the Head Feature Principle, the so-called Generalized Head Feature Principle, is suggested by Ginzburg & Sag (2000: 33). They suggest that all syntactic and semantic information of the head daughter is shared with the information of the mother by default. As (4) and (6) show, the valence information at mother nodes differs from the valence information at the head daughter in Specifier-Head phrases and in Head-Complement phrases. The same is true for semantic information: usually the semantic information at the mother node differs from the information at the head daughter, since the mother node has a collection of the semantic contributions of all daughters. This is captured by Ginzburg & Sag because sharing all information is a default that is overwritten in subtypes of *headed-phrase*. Defaults are often used in linguistics to describe unmarked cases, but what the Generalized Head Feature Principle sets as a default never actually holds. In fact, there is not a single structure in any HPSG theory I am aware of in which all syntactic and semantic features of head daughter and mother are identical. So the Generalized Head Feature Principle is not a generalization. It is never true, and hence I do not use it.

$$(7) \quad \textit{headed-phrase} \Rightarrow \begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{head-dtr} & \left[\textsc{head} \ \boxed{1}\right] \end{bmatrix}$$

(7) specifies a constraint that holds for all feature structures of type *headed-phrase*, including those that are subtypes of *headed-phrase*. The constraint identifies the head features of the head daughter with the head features of the mother.

The fact that we have an implication in (7) cannot be emphasized enough. This means that the conclusion has to hold only if the antecedent is true. If the antecedent is false, nothing is said about the presence of head daughters or the values of head features. This means that one can assume headless structures in HPSG, and there are plenty of examples of headless constructions in the literature (Müller 1999a, 1999b: Chapter 10). Hence it would be wrong to claim that HPSG assumes that all structures must be headed. I will return to headless constructions in Section 6.

Before turning to such truly headless constructions, in the following section I want to discuss nominal structures, which are interesting for two reasons. For one, researchers still disagree as to which element in a nominal structure is the head. And for another, both determiner and noun may be omitted in German, which means that nominal structures could be problematic for linguistic theories in general.

3 Nominal structures

Since the 1970s there have been proposals to treat the determiner as the head of nominal structures. Such proposals became popular within the framework of GB (Chomsky 1981) but are entertained in other frameworks as well: there are proposals in Categorial Grammar, LFG, HPSG, and Dependency Grammar that assume the determiner to be the head. See Ajdukiewicz (1935: 6), Vennemann & Harlow (1977), Brame (1982), Hudson (1984: 90–92), Hellan (1986), Abney (1987), Netter (1994, 1998), Van Langendonck (1994), Salzmann (2020, 2022) for DP proposals and Pollard & Sag (1994), Demske (2001: 49), Müller (2007b: Section 6.6.1), Bruening (2009, 2020) for NP proposals in various frameworks. Hudson (2004) working in Word Grammar, a version of Dependency Grammar, suggests mutual dependency between determiner and noun.[8]

---

[8]Due to space limitations, it is not possible to go into the details of a comparison, but such mutual dependencies are also assumed in HPSG: the noun selects the determiner via the valence feature SPR and the determiner selects the noun via the feature SPECIFIED (Pollard & Sag 1994: 50).

3.1 Tests for head status

I talked about prepositions, verbs, and adjectives at the beginning of the previous section. It is clear that these categories are heads of their respective phrasal units. This begs the question whether there are criteria for headedness that could help deciding the question for nominal structures. Zwicky (1985) looked at tests for headhood in the 80s more carefully. The tests will be repeated below and it will be examined whether they are useful in the DP/NP debate.

3.1.1 The subcategorizand

Zwicky (1985: Section 2.1.2) states that the subcategorizand is likely to be the head. The subcategorizand is the lexical element, in contrast to the phrasal one(s), and it may appear in certain configurations. For instance, the verb *give* can appear with two NP arguments (e.g. in *give her a book*) or with an NP and a PP as arguments (e.g. in *give a book to her*). On the other hand, *donate* is restricted to NP and PP. In both cases, the lexical element (the verb) is the head of the respective phrase. For nominal structures, Zwicky argues that Det must be the subcategorizand, since the Det is the sole lexical element in Det-$\overline{\text{N}}$ combinations, and hence the determiner is the only plausible candidate for a lexical head. Unfortunately, he missed the fact that the determiner may be complex both in English and other languages as well:[9]

(8) a. the Queen of England's son
  b. unter des Körpersportlers   Haut[10]
      below the body.sportsman's skin
      'below the body builder's skin'

Since both the determiner and the $\overline{\text{N}}$ can be phrasal, this test does not really help here.

3.1.2 The morphosyntactic locus

A further test discussed by Zwicky (1985: Section 2.1.3) is the test for the morphosyntactic locus. The inflectional features are located at the noun in English:

(9) a. the child
  b. the children

---

[9]See Pollard & Sag (1994: 51–54) and Ginzburg & Sag (2000: 193) for analyses of *'s* as determiner and of complex prenominal phrases as determiner phrases.
[10]taz [German newspaper], 1995-01-04, p. 15, quoted from Müller (1999b: 59)

However, this test does not help for German, since determiners are inflected as well:

(10) a. das   / dieses   Kind
          the.SG.N   this.SG.N child(N)

      b. die   / diese   Kinder
          the.PL   these.PL children.PL

3.1.3 Determinant of concord

Zwicky (1985: Section 2.2.2) looks at the element that determines concord within a phrase. Sometimes it is claimed that the determiner is responsible for the inflection class of the adjective. The determiners in (11) have a fixed inflection class and the other elements in the nominal structure have to be appropriate for the respective class with it:

(11) a. ein kluger Mann
          a   smart man

      b. der kluge Mann
          the smart man

However, it is equally possible to argue the other way round, and Zwicky (1985: 9) does exactly this: gender is an inherent property of most nouns and the determiner has to match the gender of the noun:

(12) a. der   Mann
          the.M man(M)

      b. die   Frau
          the.F woman(F)

This suggests that the noun is the head in nominal structures. Therefore we can conclude that this test fails as well for German: sometimes the determiner, sometimes the noun determines concord.

3.1.4 Semantic functor

A further criterion suggested by Zwicky (1985: Section 2.1.1) is the one of the semantic selector. Unfortunately, this criterion does not really decide the issue either. It is true – as Zwicky notes on page 4 – that for instance the universal quantifier selects the semantic contribution of the nominal part and incorporates it into the complete formula. The nominal part *Frauen* 'women' corresponds to the Q in (13b):

(13)   a. alle Frauen
            all   women

       b. $\lambda Q(\lambda P(\forall x(Q(x) \to P(x))))$

But on the other hand, we have relational nouns like *conquest* whose arguments may be realized in the position of the determiner. The meaning representation of (14a) has to contain (14b) somewhere:

(14)   a. Peters Eroberung der   Stadt
            Peter's conquest   of.the town

       b. conquest(Peter, town)

This means that this criterion is not reliable either. The determiner embeds the semantic contribution of the remaining nominal group, and the remaining nominal group may embed parts contributed by the determiner.

3.1.5 The distributional equivalent

Zwicky (1985: 12) states that the noun is the distributional equivalent of the whole phrase, including the determiner. Proper names like *Kim* and plural nouns like *penguins* can be used instead of *the penguins*.

As a reviewer pointed out, the criterion is a rather odd one since it could not be applied to all heads that obligatorily require arguments. Examples are prepositions in German and verbs like *devour* in English. Since a single preposition cannot be used anywhere without its NP argument, the preposition is not distributionally equivalent to the PP and hence would not qualify as the head. Clearly an unwanted result.

3.1.6 Obligatoriness

Both the determiner (15a) and the noun (15b) may be omitted in German. It is even possible to omit both of them, as (15c) shows:[11]

(15)   a. Er hilft Frauen.
            he helps women
            'He helps women.'

       b. Er hilft den klugen.
            he helps the smart
            'He helps the smart ones.'

---

[11]The adjectives in (15) could be nominalizations, but I am talking about elliptical constructions here. Nominalizations would be written with capital letters.

    c. Er hilft   klugen.
         he helps smart
         'He helps smart ones.'

As the translation of the examples shows, the pronoun *one* is used in the parallel English structures. However, English also permits nominal structures without a visible noun (Zwicky 1985, Arnold & Spencer 2015: 13). Zwicky notes that structures with omitted noun are always elliptical. This means that nouns are obligatorily present, and if they are missing, their omission is due to ellipsis. So, if this criterion is accepted, it decides in favor of N as the head.

A reviewer pointed out that there are certain cases in which the nominal part is optional, but when it does not occur, this is not due to ellipsis. Examples are *this*, *that*, *we*, and *you*:

(16)   a. this man

        b. we sailors

I think that a double categorization of these elements as determiners and full NPs is justified. So *we* would be a full DP/NP if it is used without other material and something different in constructions like *we sailors*. See Section 3.1.8.1 for more on pronoun-noun combinations. In any case, examples like the ones just mentioned show that the criterion cannot be applied without further qualifications.

3.1.7 Language acquisition, uniformity, and Poverty of the Stimulus

Abney's dissertation (Abney 1987) made the DP analysis very popular within the generative world. Abney argued for a treatment of English nominal structures that is parallel to the structures assumed for the sentential domain. Many authors assume an IP/VP analysis for English, not just those working in Mainstream Generative Grammar (MGG), but also in LFG (Bresnan, Asudeh, Toivonen & Wechsler 2016: 102). In such analyses, there is a functional I projection in addition to verbal projections. An advantage of the DP analysis is that one has parts of the theory that are similar, and hence one can claim to have found deeper laws. Apart from this, language acquisition is used as an argument in the DP/NP discussion: Chomsky still believes that language cannot be learned from input alone (Berwick, Pietroski, Yankama & Chomsky 2011).[12] Since – according to Chomsky – language is acquired despite this Poverty of the Stimulus, there must be innate language-specific knowledge which helps us to acquire language from the input

---

[12]The authors discuss auxiliary inversion. See Müller (2020: Section 13.8.2.4) and Sag et al. (2020: Section 1) for a critical discussion of these claims.

that is available. The claim was that children can acquire language because all phrases have the same internal structure and knowledge about this structure is innate and therefore helps to acquire language (Haegeman 1994: 106).[13] In particular, nominal structures have a DP structure which is parallel to the IP/VP structures.

Fodor (2001: 739) points out that the situation is not as simple if movement to places in otherwise invisible structure is possible. For example, it is not obvious in some cases whether verbs are in V, I, or C, whether a language is V2 or not, or whether we have an SVO or an SOV language, as the following three examples from English, Danish, and German illustrate (Müller 2022: Section 6.2.2, Figure 6.11):

(17) a. Conny reads a book.   −V2, SVO

      b. Conny læser en bog.   +V2, SVO

      c. Conny liest ein Buch.   +V2, SOV

So, having simple transitive sentences in the input is not enough to decide. Sentences with auxiliaries would help the linguist to decide between SOV and SVO languages, but this wouldn't help to distinguish between −V2 and +V2; for this the linguist would need examples with fronted objects, rather than subjects as in (17). As Fodor points out, there are many questions concerning how language acquisition is supposed to work in a Principles & Parameters setting. It is unclear how a child can determine which way to set the parameters. Fodor suggests a model assuming innate treelets that can be used in analyses of utterances and shows that this avoids problems of alternative approaches. While this seems to be the most plausible approach with the Principles & Parameters framework, there are still serious issues (discussed by Fodor herself), and of course the overall question is how information about treelets distinguishing between V2 and non-V2 languages are supposed to make it into our genome (Hauser, Chomsky & Fitch 2002). Assuming data-driven approaches without a rich UG (Freudenthal et al. 2007) seems to be preferable.

But let us assume for the sake of the argument that uniformity of basic $\overline{\overline{\text{X}}}$ structure would help in language acquisition. Even with this assumption, there remains a problem with this argument, namely that many researchers (from different frameworks) believe that the assumption of an IP structure is not plausible

---

[13]Haegeman (1994: 106) states that "the principles of X′ theory will be part of UG, they are innate. The ordering constraints found in natural languages vary cross-linguistically and thus have to be learned by the child through exposure. Very little data will suffice to allow the child to fix the ordering constraints of the language he is learning. A child learning English will only need to be exposed to a couple of transitive sentences to realize that in English verbs precede their complements."

for German (Bayer & Kornfilt 1990, Haider 1993, Berman 2003). If German does not have an IP, it is not reasonable to assume that the DP is parallel to the sentential domain and that constraints on both domains are part of our innate linguistic knowledge.[14] Hence, uniformity of nominal and sentential domain is not an argument for the DP analysis.

I want to close this section with a somewhat ironic remark. Although I do not believe in the "parallelism in structure helps language acquisition" argument, I want to point out that the NP analysis suggested here is parallel to the analysis of the sentential domain assumed in HPSG: auxiliaries are treated as verbs, not as Is or Ts (Sag et al. 2020). The subject of verbs in SVO languages are treated as specifiers, and so are determiners in nominal structures (Müller 2022: Sections 4.3–4.4), so we have arrived at parallelism all the same.

3.1.8 Pronoun-noun combinations, selection, agreement, and idioms

Zwicky's (1985) criteria for head status discussed so far are theory-neutral, as far as this is possible. The DP analysis is considered the standard in Mainstream Generative Grammar and authors usually refer to Abney (1987) for a thorough argumentation for the DP analysis. However, Salzmann (2022), this volume, points out that all previous arguments for the DP analysis depend on theory-internal assumptions. If these assumptions are not made, the arguments collapse. Since MGG changed considerably since the 80s, none of the original arguments holds any longer. Salzmann suggests a new argument based on agreement data from Bosnian/Croatian/Serbian (BCS). In what follows, I want to discuss three phenomena that seem to argue for DP analyses (pronoun-noun combinations, selection, and agreement) and one controversial phenomenon that is very interesting in the DP/NP debate: idioms.

3.1.8.1 Pronoun-noun combinations

Let us start with pronoun-noun combinations:

(18) a. Ich Idiot habe mich   gefreut.
          I   idiot have myself been.glad
          'I idiot was glad.'

---

[14]Haider (1992) assumes a functional head in the German clause that is not IP. In his DP approach he assumes a parallel between his FP (Functional Projection) and the DP. While there is a parallel in the functional/lexical structure of FP and DP, the makeup of the respective phrases in German is quite different. The head of the FP is the place for the finite verb or a complementizer and SpecCP is the target for fronted constituents in V2 clauses.

b. Du Idiot hast dich   gefreut.
   you idiot have yourself been.glad
   'You idiot were glad.'

c. Wir Idioten haben uns   gefreut.
   we idiots   have   ourselves been.glad
   'We idiots were glad.'

d. * Er Idiot hat sich   gefreut.
   he idiot has himself been.glad
   Intended: 'He was glad and he is an idiot.'

The examples in (18) show that the pronoun agrees with the verb in person and number, while the noun together with a determiner is always third person. (18d) shows that third person pronouns are not possible in this construction, so there is something idiosyncratic about it. Simon (2003: 139–140) and references cited there see data like this as evidence for the DP analysis, but I think this construction should not be treated as an instance of the normal NP or DP construction. Note also that some languages have this construction and combine the pronoun with a full nominal projection (Höhn 2016: Section 5.3). In these languages, one would not say that a D head selects an NP, but the pronoun would have to select a full DP. This is actually the solution suggested by Höhn (2016: 568): he assumes a PersP with the pronoun as head selecting a DP.[15] So, it seems reasonable to treat the pronoun as a head, but the whole construction should not be decisive in the DP/NP discussion.

3.1.8.2 Selection

Salzmann (2022: Section 3.2.2.2), this volume, mentions the fact that incorporation seems to require selection of nominal structures without determiners (NPs), while otherwise, verbs select nominal structures with determiners (DPs). For this to have any force as an argument for DP, one needs the assumption that only maximal projections can be selected. However, this assumption is not made in HPSG. For example, partial verb phrase fronting is explained by assuming that non-maximal verbal projections may be combined with governing heads (Müller 1996b; 2002: Section 2.2.2; Meurers 1999). And once non-maximal projections can be combined with heads, we can have heads combining with bare nouns, $\overline{\text{N}}$s, and NPs, and hence there is no argument for DP. See for example Müller (2010: 632) for the suggestion that light verbs in Persian may combine with lexical nouns.

---

[15]Note that this begs the question why governing heads selecting for DPs can take PersPs as well.

Salzmann also states that it is impossible to "select the absence of structure" (Salzmann 2020: 28, 32, 2022: Section 3.2.3), but this does not apply to HPSG. Since $\overline{\text{N}}$ is defined as a nominal projection without a specifier, one can select for something with absent structure. Furthermore, material that is combined contributes to the properties of a complex category. The respective contributions can be selected for. This is independent of the question at which projection level the respective combination takes place. See for example the use of the marking feature in Pollard & Sag (1994: 45–46) or Van Eynde (2006).

3.1.8.3 Agreement

Salzmann (2020: Section 4.3) discusses agreement patterns from Bosnian/Croatian/Serbian (BCS) and argues that they show that there has to be a DP layer over an NP layer to get the facts right. While he is very careful to show that Abney's (1987) arguments for the DP analysis are theory-internal, the same holds for Salzmann's new argument: agreement is established via the Probe/Goal mechanism of Minimalism (Chomsky 2001). This crucially relies on c-command and the proximity of agreement source and agreement target. In comparison to this, agreement is dealt with differently in frameworks like HPSG: the main expressive tool is structure sharing. It is not assumed that there is an agreement source and an agreement target (Pollard & Sag 1994: Section 2.2), but instead, both exponents of morpho-syntactic features are treated alike and the information on both sides is simply identified (Pollard & Sag 1994: Chapter 2, Kathol 1999, Wechsler & Zlatić 2003, Van Eynde 2021). Salzmann points out that there are two types of agreement in BCS: adjectives and determiners may agree in grammatical and semantic gender. If a certain head agrees in semantic gender, the next higher head cannot go back to grammatical gender. Salzmann concludes from this that there must be a DP layer, since in terms of c-command the determiner would be the highest head and its features would be decisive for agreement with elements outside of the nominal structure (p. 38). But this conclusion is dependent on many theory-internal assumptions. Bruening (2020: Section 4), working in the same framework and assuming an NP approach, developed an alternative theory of the agreement facts.

So again, the argument that Salzmann suggests instead of Abney's arguments is also a theory-internal one.[16]

---

[16]Salzmann's puzzle is not solved for HPSG yet. Van Eynde (2020) discusses agreement in BCS but does not solve the problem of inaccessibility of one of the two agreement options. In his account both agreement options are always available. So further research and modification of the general theory of agreement is needed but it is not necessary to assume an agreement theory based on c-command.

3.1.8.4 Idioms

A very interesting argument comes from Bruening (2020: Section 5). Bruening argues that idioms make reference to dependency chains. This was also suggested within Dependency Grammar (O'Grady 1998, Osborne & Groß 2012: Section 4.2). Osborne & Groß (2012: Section 4.2) argue for the importance of dependency relations in linguistic descriptions and explicitly claim that all idioms are based on dependency chains. They assume that determiners are dependents of nouns and explicitly state that idioms with fixed verbs, free nouns, and fixed determiners do not exist (p. 180).

If the Catena claim is correct, this is 100% compatible with the NP analysis suggested here. Sag (2007) and Kay, Sag & Flickinger (2015) developed a local theory of idioms that is based on selection. This is compatible with the claims made by Bruening (2020) and Osborne & Groß (2012: Section 4.2).

Salzmann (2020: 31) argues that examples like the ones in (19) are counterexamples to the dependency chain claim:

(19)   She plays the piano/trombone/flute.

In the specific collocation at hand one has to use the definite determiner and the actual instrument is open. Salzmann sees this as a data point that could be used to argue against the dependency chain claim. I would argue, however, that the fact that an instrument has to be inserted shows that *piano*/*trombone*/*flute* are part of the idiom. The material that is fixed in idioms varies to a great degree. Sometimes case is fixed, sometimes it is not. Sometimes idioms can be passivized or used in relative clauses, sometimes they cannot (Nunberg et al. 1994). In the case at hand, the semantic properties of the noun slot are specified: *She plays the volleyball* is not possible. Hence the nouns are part of the collocation and the determiner depends on a collocation element as predicted by the theory.

3.1.9 Summary

Summing up the discussion of Zwicky's (1985) tests for headedness and their application to the DP/NP issue, it can be said that these tests deliver inconclusive results. Further arguments for either NP or DP are either theory-internal or pro-NP (the idiom data).

3.2 The DP analysis

Having discussed criteria for head status in the previous subsection, I now turn to the DP analysis and show why an NP analysis should be preferred.

3.2.1 Personal pronouns

Figure 7 shows the analysis of personal pronouns in the DP analysis. Personal pronouns are complete and stand for a full nominal structure. Hence they are D^0 rather than N^0. D^0 is projected to the maximal level, that is, to DP. This begs the question how languages without determiners are analyzed (e.g., Slavic languages; Zlatić 2014). Since there are no determiners, maximal projections within nominal structures have to be NPs. Since personal pronouns are placeholders for the whole structure, they should be NPs as well. Hence we had languages in which pronouns are DP and others in which pronouns are NPs, which would be somewhat unsatisfying.

3.2.2 Possessive pronouns

The next question concerns possessive pronouns. Possessive pronouns and possessor phrases in general will play a major role in my argument for an NP analysis in HPSG, which is the reason why the proposals to treat them in a DP approach are discussed here. There are proposals to analyze possessives like determiners, that is, as D^0 (left figure in Figure 8). Since possessives may be complex and since possessive pronouns alternate with such possessive phrases, both should occupy the same position. This is the reason for G. Müller (2007a: 18) to analyze possessive pronouns as specifiers. The middle figure in Figure 8 shows a GB rendering of this analysis. Finally, Olsen (1991: 52) observes that a DP like *seine Stadt* 'his town' is third person but *seine* 'his' is first person, and that this is evidence against the possessive pronoun being the head. Therefore, she assumes that *-e* is the D head and *mein-* is a DP functioning as the specifier. While this seems to be convincing at first, one could assume that *seine* has a first person referen-

DP
|
D′
|
D^0
|
er

Figure 7: Personal pronouns in the DP analysis

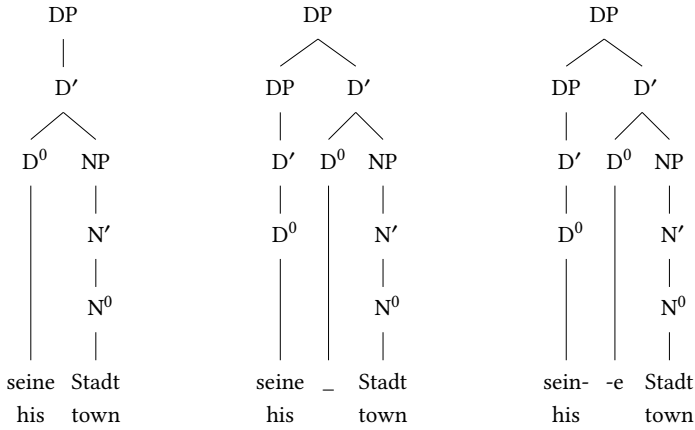Figure 8: Possible analyses for possessive pronouns: left as D^0, middle as specifier of an empty D^0, right as specifier with inflection in D^0 following Olsen (1991: 53)

tial index but syntactic features for third person for DP-internal agreement. The rightmost analysis in Figure 8, therefore, is not the only possibility. One could assume the one in the middle as well. The analysis to the right would not be an option within HPSG anyway, since usually, fully inflected words are inserted into syntax rather than bound morphemes like *-e*.

The assignment of thematic roles by relational nouns also plays a role in the analysis of possessives. These are discussed in the following section.

3.2.3 Relational nouns and assignment of semantic roles

If possessives were analyzed as D heads as in Figure 8 (left) and in Figure 9, relational nouns would have to assign a semantic role to a head position that is higher up in the tree (Olsen 1991: 51). This is prohibited since, according to Chomsky (1981: 47), semantic roles may be assigned to argument positions (A positions) only. Chomsky explicitly does not count head positions among these. One could claim that the possessive pronoun is a D^0 and that the agent gets its semantic role within the NP and is then moved out of the NP into the head position, as shown in the left figure in Figure 10. In such a setting, the movement out of the NP would have to target a head position, which is also prohibited (see e.g., Radford (2004: Sections 5.1, 6.1, 7.1) on various types of movement). This means that possessive pronouns have to be placed into SpecDP. It follows that the semantic role assigned by the relational noun is either assigned nonlocally, that
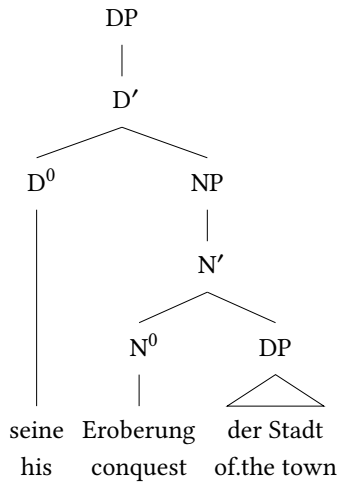
Figure 9: Possessive pronouns as D^0 and assignment of a semantic role by a relational noun
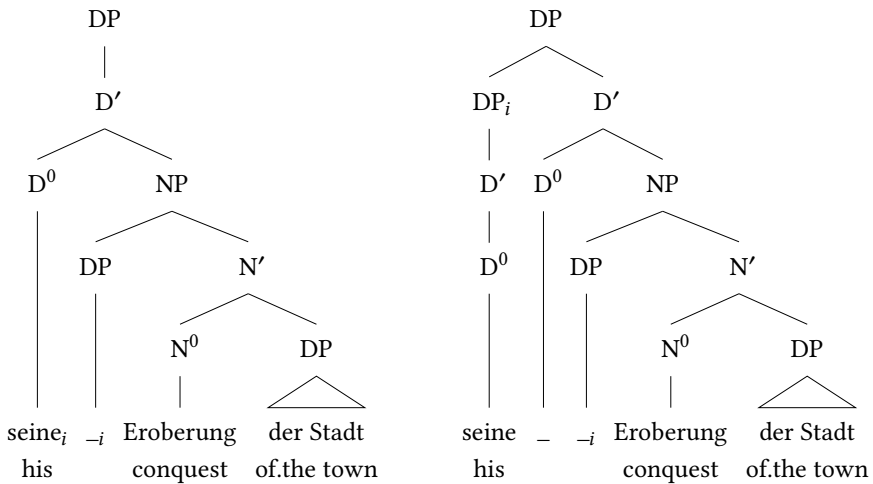

Figure 10: Assignment of semantic role to SpecNP and successive movement from a phrase position into a head position or a specifier position, respectively

is, from within the NP to a specifier position of the DP, or that the assignment is local within the NP and the receiving element is moved into the specifier position of the DP (see Figure 10 (right)).[17]

Given what was said in this and the previous section, it follows that possessive pronouns have to be in SpecDP in the DP system. Figure 11 shows the structures of the DP and NP analysis in a fully fledged $\overline{\text{X}}$ system.
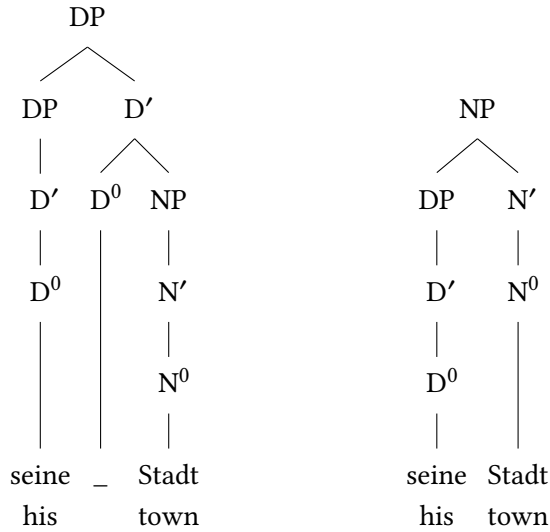

Figure 11: Comparison of the construction with possessive pronouns in a fully fledged $\overline{\text{X}}$ system in the DP and in the NP analysis

It is obvious that the NP analysis is much simpler.

3.2.4 The DP analysis in Minimalism and the NP analysis in HPSG

The HPSG analysis with an NP structure is really minimal: the lexical noun is combined directly with the determiner/possessive pronoun. Since the noun does not require anything but the determiner, it has the category $\overline{\text{N}}$.[18] The possessive pronoun is complete as well and need not be combined with other elements.

---

[17]HPSG does not refer to movement within NPs, but, as a reviewer pointed out, similar effects can be obtained by assuming that D embeds an $\overline{\text{N}}$ and that the specifier of the embedded $\overline{\text{N}}$ is shared with the specifier of the determiner (see Hinrichs & Nakazawa (1994) for argument attraction in general).

[18]Categories are feature bundles. $\overline{\text{N}}$ is an abbreviation for a nominal object selecting a determiner. (5) is an example.

Hence it may be combined with the noun, directly resulting in a fully saturated nominal projection: an NP. The left figure in Figure 12 shows the HPSG analysis. The figure on the right-hand side of Figure 12 shows the respective analysis in
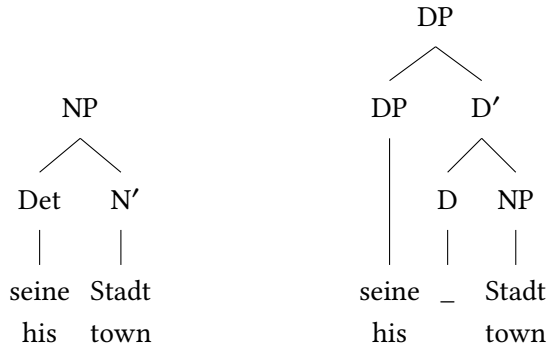

Figure 12: NP and DP analysis in HPSG and Minimalism

Minimalism. Although the problem with unnecessary unary branching nodes does not exist in the framework of Bare Phrase Structure (Chomsky 1995a),[19] the problem with role assignment remains. In contrast, in the NP analysis, possessives are in the specifier position of the noun and can receive their semantic role there. If one assumes a DP analysis, one would have to assign the semantic role to the governing head (in HPSG) or to an even higher element – the specifier of the governing head. Within Minimalism, one could – or rather, had to – assume non-local role assignment across phrase boundaries or movement of the possessive pronoun out of the NP into the dominating DP (Salzmann 2020: 18).

In conclusion, one can say that there is almost no theory-external evidence for a DP or NP analysis. The criteria for headedness are inconclusive in the DP/NP area. Only few tests clearly decide the issue, and these are in favour of an NP analysis. Theory-internal considerations show, however, that the NP analysis must be preferred in non-transformational approaches.

---

[19]In Bare Phrase Structure Grammar, unary projections of determiners or nouns do not exist. Linguistic objects are combined with Merge and the category of the result is determined by Labeling. The Label is basically part of speech information, bar levels are not used. A noun and its dependent form a phrase and if the noun does not require any further arguments, the result of the combination will be a complete nominal object, which corresponds to the classical NP. Assuming a DP analysis of *the house*, the noun like *house* is categorized as NP right away. So in Bare Phrase Structure, lexical items can be both minimal and maximal at the same time (Chomsky 1995b: 64). This is parallel to what Categorial Grammar assumed since Ajdukiewicz (1935) and what is assumed in the analysis of nominal structures in HPSG as well. See also Muysken (1982) for an early suggestions to collapse bar-levels in the GB framework. See Müller (2020: Section 4.6.2) for further discussion of Bare Phrase Structure Grammar.

The next section deals with invisible heads, and nominal structures will play an important role in this section as well.

4 Invisible heads

Wunderlich (1987: 37) writes the following on empty elements in syntax:

> Eine sinnfällige Sprachtheorie sollte die Prinzipien der Sprache so nahe wie möglich nachzeichnen und nicht Repräsentationen für Positionen vorsehen, die aus funktionalen Gründen gar nicht erscheinen. Dem erwähnten sprachinhärenten Prinzip möchte ich daher das methodologische Prinzip „Vermeide leere Kategorien" zur Seite stellen. (Wunderlich 1987: 37)[20]

While early HPSG used empty elements in nonlocal dependencies (traces, Pollard & Sag 1994: 164) and empty heads for the analysis of relative clauses (Pollard & Sag 1994: 216), later publications tried to avoid empty elements (Sag & Fodor 1995, Sag 1997, Bouma, Malouf & Sag 2001). This section discusses two examples of empty heads: Subsection 4.1 deals with nominal structures again and suggests an empty nominal head, and Subsection 4.2 deals with copula constructions in African American Vernacular English (AAVE), for which an empty verbal head was suggested.

Another empty verbal head is assumed in the analysis of German by almost all HPSG theoreticians working on German. I followed an approach without such an empty verbal head from 1993 until 2003 (see Müller 2002: Section 1.9), but I am now convinced that the assumption of the empty verbal head is necessary to account for apparent multiple frontings in German (Müller 2003, 2005). The discussion of the arguments for an analysis with an empty verbal head cannot be included here due to space limitations, but the reader is referred to a book-length discussion of the data, the analysis, and its alternatives in Müller (2021b).

See also Borsley (1999, 2009, 2013) for further explicit suggestions of analyses with empty heads.

4.1 Nominal heads

In the previous section, I argued for an NP analysis, that is, for an analysis in which the noun is the head. This begs the question how to analyze phrases that distributionally behave like NPs but do not contain a noun. The phrases in (20f–k) may appear in places in which the NPs in (20a–e) may appear:

---

[20]Language theory should model the principles of language as closely as possible. It should not assume representations for positions that do not appear for functional reasons. In addition to the principle inherent to language mentioned already [Avoid Pronoun], I would like to add the methodological principle Avoid Empty Categories. [my translation, St.M.]

(20)   a. die kluge Frau
          the smart woman

       b. die Frau   aus Hamburg
          the woman from Hamburg

       c. die kluge Frau   aus Hamburg
          the smart woman from Hamburg

       d. die kluge Frau,   die wir kennen
          the smart woman who we know

       e. die kluge Frau   aus Hamburg, die wir kennen
          the smart woman from Hamburg who we know

       f. die kluge
          the smart
          'the smart one'

       g. die aus   Hamburg
          the from Hamburg
          'the one from Hamburg'

       h. die kluge aus   Hamburg
          the smart from Hamburg
          'the smart one from Hamburg'

       i. die, die   wir kennen
          the who we know
          'the one who we know'

       j. die kluge, die   wir kennen
          the smart who we know
          'the smart one who we know'

       k. die kluge aus   Hamburg, die   wir kennen
          the smart from Hamburg who we know
          'the smart one from Hamburg who we know'

For instance, all phrases in (20) may function as the subject of the verb *lacht* 'laughs'. Therefore it is appropriate to categorize all these phrases with the same label, rather than to assume that those in (20a–e) are NPs and those in (20f–k) DPs, say. If we want to analyze (20f–k) as NPs, we either have to assume an empty nominal head or we have to formulate rules for NPs that say that an NP may consist of a determiner and one or several adjectives, or of a determiner and PPs, or relative clauses, or some variation of these elements. The set of rules would grow and the generalizations would not be captured (see Section 5.1). Instead of

this, one can simply assume an empty nominal head. The advantage of this is that all phrases in (20) can be analyzed with the same set of rules and that they have the same structure. Figure 13 shows the analysis of (20h). A simple trick to get rid of the empty element in Figure 13 is to assume a unary branching rule that projects the adjective to $\overline{\text{N}}$ (Wunderlich 1987). Note though that this unary branching rule does not account for (20g). We will come back to this in Section 5.
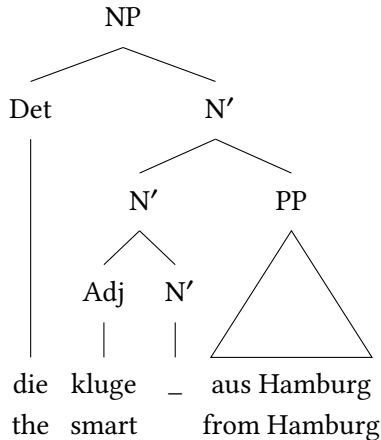
Figure 13: Analysis of *die kluge aus Hamburg* 'the smart one from Hamburg' with empty nominal head

As was noted in Section 3.1.6, determiners can be omitted as well. This is possible for all nouns in the plural:

(21) a. Frauen
       women

   b. Frauen, die wir kennen
      women who we know

   c. kluge Frauen
      smart women

   d. kluge Frauen, die wir kennen
      smart women who we know

Mass nouns may be used without a determiner in the singular as well:

(22) a. Getreide
       grain

   b. Getreide, das gerade gemahlen wurde
      grain   that just   ground   was
      'grain that was just ground'

    c. frisches Getreide
       fresh   grain

    d. frisches Getreide, das gerade gemahlen wurde
       fresh   grain   that just   ground   was
       'fresh grain that was just ground'

As I did for structures without a noun, one may assume an empty determiner (Pollard & Sag 1994: 90). The analysis of (21a) is shown in Figure 14.
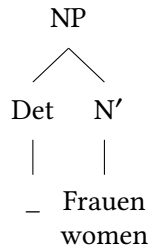

Figure 14: Analysis of nominal structures without a determiner

Interestingly, both determiner and noun may be omitted in a phrase, resulting in phrases consisting of one or several adjectives and possibly PPs and relative clauses:

(23)   a. Ich helfe klugen.
        I   help smart
        'I help smart ones.'

    b. Dort drüben steht frisches, das gerade gemahlen wurde.
        there over   stands fresh   that just   ground   was
        'Over there is fresh [grain] that was just ground.'

The structures for (23a) and a similar NP including a modifying PP are shown in Figure 15.

Instead of an empty determiner, one can assume a unary branching rule projecting an $\overline{\text{N}}$ to NP (Müller 2007b: 88).[21] This and other alternatives to empty elements will be discussed in Section 5. But before turning to alternatives to empty heads, I want to discuss empty verbal heads in the next subsection.

---

[21]The computational implementation (Müller 1996a) of the grammar described in Müller (1999b) did not contain empty elements. I used a unary branching rule for structures without a determiner. A lexical rule, as suggested by Michaelis (2006: 80), is not an option. See Müller (2007b: Section 6.6.2) on this point.
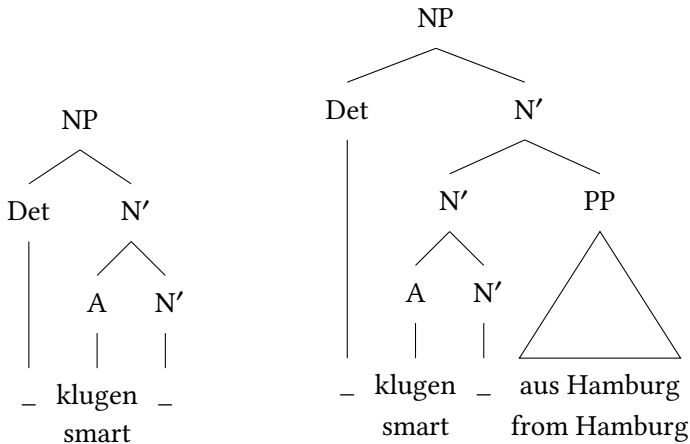
Figure 15: Analysis of nominal structures lacking both determiner and noun: *klugen* 'smart ones' and *klugen aus Hamburg* 'smart ones from Hamburg'

4.2 Verbal heads

Bender (2001) discusses data from African American Vernacular English (AAVE), in which the copula can be omitted, resulting in sentences like (24), taken from Sag, Wasow & Bender (2003: 457):

(24)   a. Chris at home.
       b. We angry with you.
       c. You a genius.
       d. They askin for help.

Sag, Wasow & Bender (2003: Section 15.3.4) discuss a phrasal schema that combines a predicate selecting for an NP in a certain form with this NP directly.

(25)   S → NP Pred

While this provides an account for examples like (24), it fails on examples like (26), also taken from Sag et al. (2003: 463):

(26)   a. How old they say his baby ϕ?
       b. Tha's the man they say ϕ in love.

The interesting fact about these examples is that the predicate is extracted in (26a) and the subject is extracted in (26b). The rule in (25) cannot apply in the analysis of (26), since Sag et al. do not assume traces for extraction, and hence one would

need a special rule for the case in (26a) for combining a subject with an extracted predicate and a special rule for the case in (26b) for combining a predicate with an extracted subject.[22] Rather than assuming three unrelated rules, Sag et al. follow Bender (2001) in assuming an empty head for the copula. This is an interesting twist in the discussion of empty elements, since the need to assume this empty copula was caused by eliminating empty elements in the analysis of extraction phenomena by Bouma, Malouf & Sag (2001).

5 Grammar conversion

In the previous section, I suggested using empty elements in the analysis of nominal structures. Current grammatical theories have different views regarding empty elements. There are hundreds of empty elements of various categories in Minimalism (Webelhuth 1995: 76; Newmeyer 2004: 194; 2005: 82; Müller 2020: Section 4.6.1.1), in most Construction Grammar variants there is not a single one, and in other frameworks it depends on the author whether empty elements are assumed and, if so, which ones.[23] Apart from the two empty elements mentioned already, I am using only two further empty elements in my grammars: one for nonlocal movement and one for head movement (Müller 2013, 2021b).

I show in this section which formal means used in various frameworks correspond to each other and how grammars using empty elements can be transformed into grammars without them. This may help to objectify the discussion, which is a bit emotional sometimes.

5.1 Phrase structure grammars

It was shown as early as the 1960s that grammars with empty elements can be transformed into grammars without them by inserting the empty elements into the grammar rules. This results in new grammar rules in which the respective symbols do not appear any longer (Bar-Hillel et al. 1961, Müller 2004).

Let's take the grammar in (27a) as an example. We can eliminate the empty element for $\overline{\text{N}}$ by adding new rules for all rules where $\overline{\text{N}}$ appears on the right-hand side of the rule. The result of such a transformation is shown in (27b):

---

[22]Note that this is basically the same problem as the one I pointed out in the discussion of phrasal approaches in Construction Grammar (Müller 2006: 854).

[23]Müller (2020: Section 19.1) gives an overview of approaches with and without empty elements in Categorial Grammar, GPSG, LFG, TAG, Dependency Grammar, and HPSG.

(27) a. NP → Det $\overline{\text{N}}$ b. NP → Det $\overline{\text{N}}$

NP → Det

  $\overline{\text{N}}$ → Adj $\overline{\text{N}}$ $\overline{\text{N}}$ → Adj $\overline{\text{N}}$

$\overline{\text{N}}$ → Adj

  $\overline{\text{N}}$ → $\overline{\text{N}}$ PP $\overline{\text{N}}$ → $\overline{\text{N}}$ PP

$\overline{\text{N}}$ → PP

  $\overline{\text{N}}$ → _

  Det → die   Det → die
  Adj → klugen   Adj → klugen
  $\overline{\text{N}}$ → Frauen $\overline{\text{N}}$ → Frauen

When we insert empty elements into rules, it may happen that all elements on the right-hand side are deleted, which has the effect of creating new empty elements. Hence the step of inserting empty elements into rules has to be applied until it converges, no new empty elements are produced, and all empty elements are eliminated from the grammar.

As is demonstrated by the simple example in (27), the elimination of empty elements may result in an increase of the number of rules.[24] One rule from grammar (27a) (the one for the lexical item of the empty element) was removed and we got three new rules in (27b) instead. The generalization that nouns may be omitted in German is not captured directly in the new grammar any longer. Instead we have a largish number of descriptions of constituents that can form an NP or an $\overline{\text{N}}$, respectively.

It is often argued that there cannot be empty elements since these would be invisible and hence not learnable. The acquisition problem seems to argue for empty elements in the nominal structures at hand, though, since what has to be acquired is the fact that the noun can be omitted in elliptical constructions.

The empty determiner is not part of the example in (27), but it is clear that its elimination by the techniques described above results in a unary branching rule that projects an $\overline{\text{N}}$ to an NP. See Zlatić (2014: 31) for the assumption of such a unary projection.

5.2 Lexical rules

As was mentioned above, the number of empty heads that are assumed in Minimalist work is significant. Some contribute semantic information and are important for valence alternations like the one in (28):

---

[24]Wunderlich (1987: 38), discussing a proposal with an empty head by Olsen (1987), suggests a rule projecting nouns from adjectives, but does not mention cases like (20g), in which no adjective is present.

(28)   a.   Er   bäckt einen Kuchen.
            he.NOM bakes a.ACC cake

       b.   Er   bäckt ihr   einen Kuchen.
            he.NOM bakes her.DAT a.ACC cake

(28a) shows the transitive verb *backen* 'to bake' with a nominative and an ac-
cusative argument. (28b) has a dative argument in addition.

Almost all linguistic theories handle such alternations without empty elements.
LFG, HPSG, and SBCG analyze such valence alternations via lexical rules instead
(Toivonen 2013, Müller 2018, Sag, Boas & Kay 2012). Figure 16 is a comparison
of the two analyses. The left-hand side shows a lexical rule-based analysis relat-

$$V[\langle\, NP_x, NP_z, NP_y\, \rangle]$$ $$BenefactiveP[\langle\, NP_x, NP_z, NP_y\, \rangle]$$

$$V[\langle\, NP_x, NP_y\, \rangle]$$ $$V[\langle\, NP_x, NP_y\, \rangle]$$ Benefactive
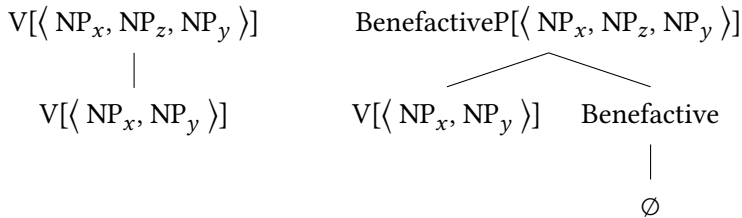
$$\varnothing$$

Figure 16: Comparison of the analysis of valence alternations by lexical rule and
empty head

ing a word with two elements in the valence list to a word with three elements
in the valence list. The right-hand side shows an analysis with an empty head:
The benefactive head selects a verb with two arguments, and the result of the
combination is a projection that takes three arguments.[25]

Lexical rules in HPSG are basically unary branching rules (Briscoe & Copes-
take 1999, Meurers 2001) and hence it does not come as a big surprise that they
correspond to constructions with an empty head.

5.3 Recategorization of phrases

The previous section dealt with lexical rules. Lexical rules relate words or stems
to other words or stems. One way to model lexical rules is parallel to unary
branching rules. But nothing prevents one from relating phrases to one another.

---

[25]Analyses suggested in the literature usually combine a VP with one further head, that is, the
verbal head is combined with its arguments first and the result of this combination is then
combined with the benefactive head (Bosse & Bruening 2011: 75).

For instance, Partee (1986) and, following her, Müller (2009) suggest recategorizing NPs like *ein guter Lehrer* 'a good teacher' as in (29a) via a unary projection into an NP that can be used predicatively as in (29b). (The lexical rule-based analysis by Ginzburg & Sag (2000: 409) has scope problems, since lexical rules can be applied to single lexical elements only, and other elements that can appear in NPs (adjectives for example) cannot be part of the input of the lexical rule (Kasper 1997, Müller 2012).)

(29)   a. Ein guter Lehrer lobt.
            a   good teacher praises
            'A good teacher praises.'

        b. Er ist ein guter Lehrer.
            he is   a   good teacher
            'He is a good teacher.'

Figure 17 shows the analysis with a unary branching syntactic rule. The rule projects a NP[PRD−] to NP[PRD+], and of course the semantic type of the NP is adapted as well. This is not shown in the figure, since I do not have the space to introduce semantic representations in this paper. What is missing from the figure is that both valence and semantics of the dominating NP are different from the one of the dominated NP. The predicative NP selects for a subject and introduces a respective relation that relates the subject to the predicative noun.

$$\text{NP[PRD+]}$$
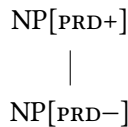$$|$$
$$\text{NP[PRD−]}$$

Figure 17: Semantic type raising via unary rule

The same effect can be reached by assuming an empty nominal head selecting for a PRD− phrase and projecting a PRD+ one.[26]

Another interesting case are free relative clauses. Free relative clauses have the form of relative clauses but function like NP or PP arguments or adjuncts

---

[26]Proposals in MGG sometimes use an empty Pred head projecting a PredP (Bowers 1993). This is not entirely equivalent to what is suggested here, since all categories are projected as PredP, and this makes it impossible for governing heads to select the syntactic category of the predicative element they combine with. As Pollard & Sag (1994: 105–106) pointed out, verbs like *grow*, *get*, *turn out*, *become*, and *end up* select different kinds of predicative elements. In the analysis in Figure 17, an NP is projected into an NP. The syntactic category remains selectable.

(Bausewein 1991, Müller 1999a). For example, in (30) the relative clause *wem er vertraut* 'who he trusts' fills the slot of the dative object of *helfen* 'to help':

(30)   *Wem* er vertraut, hilft   er auch.[27]
       who he trusts   helps he too
       'He helps those he trusts.'

Groos & van Riemsdijk (1981: Section 2) suggest an analysis with an empty head (XP) that is modified by a relative clause and the properties of the relative phrase are related to the ones of the empty head. This analysis as sketched in (31a) is interesting since it is parallel to normal relative clause structures containing an overt pronoun as in (31b):

(31)   a.   [$_{NP[dat]}$ _NP[dat] [Wem er vertraut]], hilft er auch.
       b.   [$_{NP[dat]}$ Denen$_{NP[dat]}$, [denen er vertraut]], hilft   er auch.
                         those   who he trusts   helps he too
            'He helps those he trusts.'

However, the problem is that relative clauses are adjuncts and modification by adjuncts is optional. To maintain an analysis like the one by Groos & van Riemsdijk, one would have to assume that modification of the empty head by a relative clause is obligatory, since otherwise one would have complete XPs in the grammar that could function as arguments in other areas of the grammar (Müller 1999a: 97). For example, one could derive sentences with ditransitive verbs and all of the arguments could be saturated by the empty element:

(32)   * dass _NP[nom] _NP[dat] _NP[acc] gibt
         that   gives
         As in: 'that she gives it to her'

Usually adjunction is not obligatory however. While empty elements and unary projections are equivalent in most cases, we have a clear difference here. If one analyzes free relatives using a unary branching rule mapping a free relative clause to an XP, it is clear that this rule can only apply if there is a free relative clause, while nothing ensures the presence of an adjunct in the analysis using an empty head.

5.4 Summary

I have used nominal structures without overt nouns to further support the point that empty elements may help capturing generalizations in some cases. They

---

[27]Engel (1977: 234)

can be avoided in other parts of grammars, and unary branching projections in the lexicon or in the syntax may be assumed instead. Sometimes unary projections have an advantage over empty heads since they can ensure the obligatory presence of constituents that would be treated as adjuncts to empty elements.

6 Headless structures

Section 4 dealt with structures in which we usually find certain elements, and it was argued that it is reasonable to use empty elements in the places in which the heads would appear when realized overtly. That is, it was assumed that there is a head even though it is invisible. The argumentation for empty elements is based on the fact that the respective positions are usually filled. In Section 5.3, I argued that the assumption of an empty head in the analysis of free relative clauses would permit ill-formed structures and argued for an analysis without head. However, Jackendoff (2008) pointed out that there are sequences like those in (33), called N-P-N expressions, where it is not reasonable to assume that one of the involved elements is the head or that there is some kind of bigger structure from which an element is missing:[28]

(33)   student after student

Such sequences can be used in NP positions within larger structures, but they do not have the structure of NPs internally. For instance, there is no determiner and there is the restriction that the second N has to be identical with the first one. The meaning of N-P-N expressions cannot be determined compositionally from the meaning of the parts: N *after* N roughly means *many Ns in succession* (Jackendoff 2008: 26).

---

[28]See also Jacobs (2008) for further examples from German and Müller (2020: Section 21.10.1) for discussion. Another example of a construction that is usually treated as headless in HPSG is coordination. Borsley (2005) shows that the Minimalist analysis of coordination as ConjP suggested by Kayne (1994: Chapter 6), Johannessen (1998: Chapter 3), and others fails in several respects. Categorial Grammar also assumes a functor-based approach with the conjunction as the head. However, the result of the combination of two X with a conjunction categorized as (X/X)\X is an X, which gets the external distribution right as far as the main category is concerned, while this remains a puzzle in the Minimalist proposals (governing heads select a DP, not a ConjP). Nevertheless, there are cases in which the last conjunct determines the properties of the complete phrase, as Borsley has shown. So additional mechanisms seem to be needed to get the headed analysis right. I will not take a stand on this issue here but point the reader to Abeillé & Chaves (2021) for a general discussion and an overview of approaches to coordination in HPSG.

All theories assuming that all structures have a head/functor (Minimalism, Dependency Grammar, Categorial Grammar) have a problem. The previous section showed that one can charm away empty heads if one does not like them. Similarly, one can conjure up empty heads if one needs them. Figure 18 shows a hypothetical Dependency Grammar analysis. This analysis assumes an empty
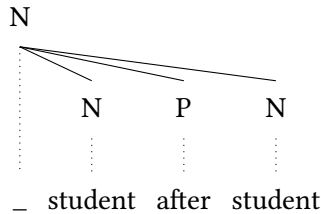

Figure 18: Analysis of the N-P-N Construction with empty head

head that selects the two Ns and a P.

Since Minimalism allows for binary branching only, one would need two empty heads to model the N-P-N construction with empty heads.[29]

Like for Constructional Grammar, Jackendoff's examples are entirely without any problem for HPSG: a special schema combines N, P, and N (and possibly further Ps and Ns). Figure 19 shows the respective analysis.
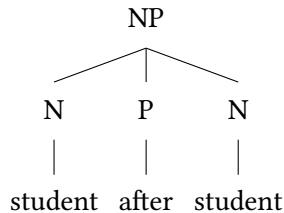

Figure 19: Construction-based analysis of N-P-N structures

I can hear the reproaches: "But the assumption of a special schema is a stipula-

---

[29]G. Müller (2011) suggests an analysis in which the preposition selects a noun and bears a feature REDUP, which triggers a reduplication of the noun. It remains unclear why a structure with a preposition as the main element should project an NP and how sequences of the type N-P-N-P-N (*student after student after student*) should be analyzed. See Jackendoff (2008: Section 4.4) and Bargmann (2015) on sequences of the latter type. Note also that G. Müller stated that his analysis predicts that adjectival modifiers of the N in N-P-N constructions are not permitted in German, a claim that is empirically false (Müller 2021c: Section 4.1).

tion!" This is true, but an empty head would be a stipulation as well. The N-P-N schema captures everything that can be and must be said about the construction: three or more elements (see Bargmann 2015) are combined idiosyncratically, resulting in an idiosyncratic meaning.

Having shown that empty elements can be removed from grammars (Section 5) and can be added if theories require heads (Figure 18), I now turn to the question of whether there are limits/style guides for the assumption of empty heads.

7 Good and bad empty elements

As mentioned on page 97, HPSG follows Wunderlich in assuming that syntactic theories should avoid the stipulation of empty elements because of methodological considerations. I have demonstrated that sometimes the assumption of empty elements is warranted (empty nominal and verbal heads) and sometimes it is not and should therefore be avoided (N-P-N construction). This section tries to give a more general answer to the question when it is legitimate to assume an empty element and when it is not. In general, it has to be explained how syntactic structures that are suggested can be acquired by language learners. If one assumes a lean Universal Grammar as Hauser, Chomsky & Fitch (2002) or none at all (Tomasello 2003, Goldberg 2006), then there must be language-particular evidence for empty elements. Analyses that are solely theory-internally motivated like, for instance, the analysis of PPs by Radford (1997: 452) are not legitimate. Radford assumes that case can be checked in specifier positions only. In addition, he assumes five empty elements and complicated movement processes. His analysis is shown in Figure 20. The necessity for these empty elements follows from the theoretical apparatus that is assumed. Since there is no independent evidence for the apparatus, it is not acquirable and hence has to be innate. This contradicts Minimalist assumptions (Hauser, Chomsky & Fitch 2002). A precondition to detect absence of elements is that the positions of the prospective elements can be filled in principle (Müller 2015: 40). It is not legitimate to argue cross-linguistically for empty elements, since the cross-linguistic evidence is available to us as linguists but not to those who acquire the language.

Summing up, one can say that empty determiners, nouns, and verbs can be acquired from linguistic evidence from within the language that is acquired, but categories like AgrO, Topic, Pred, etc. that are motivated with reference to other languages cannot. The respective categories should be assumed for the languages in which they have visible forms (e.g., AgrO for Basque and Topic for Japanese) but not for languages without any morphological reflexes.
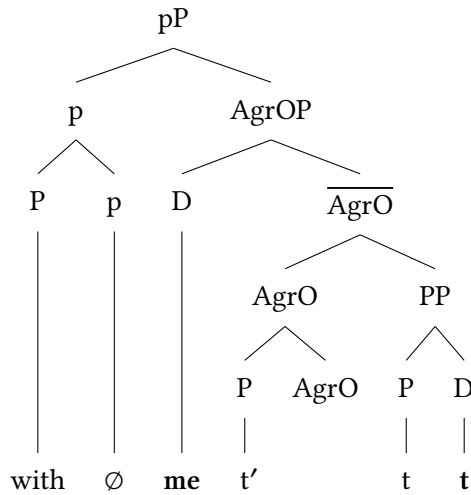
Figure 20: Theory-internally motivated analysis of a PP following Radford

8 Summary

This paper discussed the role of heads in syntactic structures (within the framework of HPSG). Heads project morpho-syntactic features (part of speech, case, verb form, and so on), and they have valence specifications determining the structure of phrases. While it is clear for most types of phrases which element is the head, theory-neutral criteria for determining heads often fail to decide the question of whether N or D is the head in German nominal structures. I used thematic role assignment and selection in idioms to argue for an NP analysis. Apart from discussing the notorious DP/NP issue, I discussed two cases in which empty heads were assumed (again nominal structures in German and copula constructions in AAVE). These empty heads filled slots in which overt material can be realized. For the N-P-N construction, an empty head could be stipulated as well, in order to save claims made in other frameworks that all structures have to have a head. Since HPSG does not make this claim, I argued for a headless construction instead. I have shown that grammars assuming empty elements can be converted into ones without empty elements in a straightforward way. Nevertheless there are conditions on the use of empty elements in grammatical theories: the elements should be recoverable from input in the language under consideration. There has to be language-internal evidence for assuming an empty element, that is the position that is assumed should be filled in some situations.

Acknowledgments

References

Abeillé, Anne & Rui P. Chaves. 2021. Coordination. 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 9), 725–776. Berlin: Language Science Press. DOI: 10.5281/zenodo.5599848.

Abney, Steven Paul. 1987. *The English noun phrase in its sentential aspect.* Cambridge, MA: MIT. (Doctoral dissertation). http://www.vinartus.net/spa/87a.pdf (10 February, 2021).

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

Allegranza, Valerio. 1998. Determiners as functors: NP structure in Italian. In Sergio Balari & Luca Dini (eds.), *Romance in HPSG* (CSLI Lecture Notes 75), 55–107. Stanford, CA: CSLI Publications.

Allegranza, Valerio. 2007. *The signs of determination: Constraint-based modelling across languages* (Sabest. Saarbrücker Beiträge zur Sprach- und Translationswissenschaft 16). Bern: Peter Lang.

Arnold, Doug & Andrew Spencer. 2015. A constructional analysis for the skeptical. In Stefan Müller (ed.), *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar, Nanyang Technological University (NTU), Singapore*, 41–60. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2015.3.

Bar-Hillel, Yehoshua, Micha A. Perles & Eliahu Shamir. 1961. On formal properties of simple phrase-structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 14(2). 143–172. DOI: 10.1524/stuf.1961.14.14.143.

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.

Bausewein, Karin. 1991. Haben kopflose Relativsätze tatsächlich keine Köpfe? In Gisbert Fanselow & Sascha W. Felix (eds.), *Strukturen und Merkmale syntaktischer Kategorien* (Studien zur deutschen Grammatik 39), 144–158. Tübingen: Gunter Narr Verlag.

Bayer, Josef & Jaklin Kornfilt. 1990. Restructuring effects in German. In Elisabeth Engdahl (ed.), *Parametric variation in Germanic and Romance* (Edinburgh Working Papers in Cognitive Science 6), 21–42. University of Edinburgh: Edinburgh University, Center of Cognitive Science.

Bender, Emily M. 2001. *Syntactic variation and linguistic competence: The case of AAVE copula absence.* Stanford University. (Doctoral dissertation).

Berman, Judith. 2003. *Clausal syntax of German* (Studies in Constraint-Based Lexicalism 12). Stanford, CA: CSLI Publications.

Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the Stimulus revisited. *Cognitive Science* 35(7). 1207–1242. DOI: 10.1111/j.1551-6709.2011.01189.x.

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

Borsley, Robert D. 1999. Mutation and constituent structure in Welsh. *Lingua* 109(4). 267–300. DOI: 10.1016/S0024-3841(99)00019-4.

Borsley, Robert D. 2005. Against ConjP. *Lingua* 115(4). 461–482. DOI: 10.1016/j.lingua.2003.09.011.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language & Linguistic Theory* 27(2). 225–265. DOI: 10.1007/s11049-009-9067-3.

Borsley, Robert D. 2013. On the nature of Welsh unbounded dependencies. *Lingua* 133. 1–29. DOI: 10.1016/j.lingua.2013.03.005.

Bosse, Solveig & Benjamin Bruening. 2011. Benefactive versus experiencer datives. In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer & Barbara Tomaszewicz (eds.), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, 69–77. Somerville, MA: Cascadilla Press.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language & Linguistic Theory* 19(1). 1–65. DOI: 10.1023/A:1006473306778.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24(4). 591–656.

Brame, Michael. 1982. The head-selector theory of lexical specifications and the non-existence of coarse categories. *Linguistic Analysis* 10(4). 321–325.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-functional syntax*. 2nd edn. (Blackwell Textbooks in Linguistics 16). Oxford: Wiley-Blackwell. DOI: 10.1002/9781119105664.

Briscoe, Ted J. & Ann Copestake. 1999. Lexical rules in constraint-based grammar. *Computational Linguistics* 25(4). 487–526. http://www.aclweb.org/anthology/J99-4002 (10 February, 2021).

Bruening, Benjamin. 2009. Selectional asymmetries between CP and DP suggest that the DP hypothesis is wrong. In Laurel MacKenzie (ed.), *Proceedings of the 32th Annual Penn Linguistics Colloquium* (Penn Working Papers in Linguistics 15.1), 26–35. Philadelphia, PA: University of Pennsylvania.

Bruening, Benjamin. 2020. The head of the nominal is N, not D: N-to-D movement, hybrid agreement, and conventionalized expressions. *Glossa: a journal of general linguistics* 5(1). 15.1–19. DOI: 10.5334/gjgl.1031.

Chomsky, Noam. 1981. *Lectures on government and binding* (Studies in Generative Grammar 9). Dordrecht: Foris Publications. DOI: 10.1515/9783110884166.

Chomsky, Noam. 1995a. Bare phrase structure. In Héctor Campos & Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero* (Georgetown Studies in Romance Linguistics), 51–109. Washington, DC: Georgetown U Press.

Chomsky, Noam. 1995b. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/9780262527347.001.0001.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language* (Current Studies in Linguistics 36), 1–52. Cambridge, MA: MIT Press.

Demske, Ulrike. 2001. *Merkmale und Relationen: Diachrone Studien zur Nominalphrase des Deutschen* (Studia Linguistica Germanica 56). Berlin: Walter de Gruyter. DOI: 10.1515/9783110811353.

Engel, Ulrich. 1977. *Syntax der deutschen Gegenwartssprache* (Grundlagen der Germanistik 22). Berlin: Erich Schmidt Verlag.

Fillmore, Charles J. 2013. Berkeley Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar* (Oxford Handbooks), 110–132. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195396683.013.0007.

Fodor, Janet Dean. 2001. Setting syntactic parameters. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory* (Blackwell Handbooks in Linguistics), 730–767. Oxford: Blackwell Publishers Ltd. DOI: 10.1002/9780470756416.ch23.

Freudenthal, Daniel, Julian M. Pine, Javier Aguado-Orea & Fernand Gobet. 2007. Modeling the developmental patterning of finiteness marking in English, Dutch, German, and Spanish using MOSAIC. *Cognitive Science* 31(2). 311–341. DOI: 10.1080/15326900701221454.

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

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224. DOI: 10.1016/S1364-6613(03)00080-9.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language* (Oxford Linguistics). Oxford: Oxford University Press.

Groos, Anneke & Henk van Riemsdijk. 1981. Matching effects in free relatives: A parameter of core grammar. In Adriana Belletti, Luciana Brandi & Luigi Rizzi (eds.), *Theory of markedness in Generative Grammar: Proceedings of the IVth GLOW conference*, 171–216. Pisa: Scuola Normale Superiore.

Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory.* 2nd edn. (Blackwell Textbooks in Linguistics 1). Oxford: Blackwell Publishers Ltd.

Haider, Hubert. 1992. Die Struktur der deutschen Nominalphrase: Lexikalische und funktionale Strukturen. In Ludger Hoffmann (ed.), *Deutsche Syntax: Ansichten und Aussichten* (Jahrbuch des Instituts für deutsche Sprache 1991), 304–333. Berlin: de Gruyter. DOI: 10.1515/9783110622447-013.

Haider, Hubert. 1993. *Deutsche Syntax – generativ: Vorstudien zur Theorie einer projektiven Grammatik* (Tübinger Beiträge zur Linguistik 325). Tübingen: Gunter Narr Verlag.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579. DOI: 10.1126/science.298.5598.1569.

Hellan, Lars. 1986. The headedness of NPs in Norwegian. In Peter Muysken & Henk van Riemsdijk (eds.), *Features and projections* (Studies in Generative Grammar 25), 89–122. Dordrecht: Foris Publications. DOI: 10.1515/9783110871661-005.

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

Hoffmann, Thomas & Graeme Trousdale. 2013. Construction Grammar: Introduction. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of*

Construction Grammar (Oxford Handbooks), 1–13. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195396683.013.0001.

Höhn, Georg F. K. 2016. Unagreement is an illusion: Apparent person mismatches and nominal structure. *Natural Language & Linguistic Theory* 34(2). 543–592. DOI: 10.1007/s11049-015-9311-y.

Holler-Feldhaus, Anke. 2001. An HPSG analysis of German Wh-interrogatives. In W. Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7), 77–109. Stanford, CA: CSLI Publications.

Hudson, Richard. 1984. *Word grammar.* Oxford: Basil Blackwell.

Hudson, Richard. 2004. Are determiners heads? *Functions of Language* 11(1). 7–42. DOI: 10.1075/fol.11.1.03hud.

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

Jacobs, Joachim. 2008. Wozu Konstruktionen? *Linguistische Berichte* 213. 3–44.

Johannessen, Janne Bondi. 1998. *Coordination* (Oxford Studies in Comparative Syntax). Oxford: Oxford University Press.

Kasper, Robert T. 1997. *The semantics of recursive modification.* Ms. Ohio State University. http://www.essex.ac.uk/linguistics/external/clmt/papers/hpsg/modification.ps (26 February, 2020).

Kathol, Andreas. 1999. Agreement and the syntax-morphology interface in HPSG. In Robert D. Levine & Georgia M. Green (eds.), *Studies in contemporary Phrase Structure Grammar*, 223–274. Cambridge, UK: Cambridge University Press.

Kathol, Andreas. 2000. *Linear syntax* (Oxford Linguistics). Oxford: Oxford University Press.

Kay, Paul, Ivan A. Sag & Daniel P. Flickinger. 2015. *A lexical theory of phrasal idioms.* Ms. CSLI Stanford.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783110934670.

Machicao y Priemer, Antonio & Stefan Müller. 2021. NPs in German: Locality, theta roles, and prenominal genitives. *Glossa: a journal of general linguistics* 6(1). 1–38. DOI: 10.5334/gjgl.1128.

Meurers, W. Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2). 161–217. DOI: 10.1080/033258601753358605.

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

Michaelis, Laura A. 2006. Construction Grammar. In Keith Brown (ed.), *The ency-clopedia of language and linguistics*, 2nd edn., 73–84. Oxford: Elsevier Science Publisher B.V. (North-Holland). DOI: 10.1016/B0-08-044854-2/02031-9.

Michaelis, Laura A. 2013. Sign-Based Construction Grammar. In Thomas Hoff-mann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Gram-mar* (Oxford Handbooks), 133–152. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195396683.013.0008.

Müller, Gereon. 2007a. *Einführung in die Syntax: Funktionale Kategorien II: DP.* Handout University of Leipzig. http://home.uni-leipzig.de/muellerg/mu609.pdf (9 August, 2022).

Müller, Gereon. 2011. Regeln oder Konstruktionen? Von verblosen Direktiven zur sequentiellen Nominalreduplikation. In Stefan Engelberg, Anke Holler & Kris-tel Proost (eds.), *Sprachliches Wissen zwischen Lexikon und Grammatik* (Insti-tut für Deutsche Sprache, Jahrbuch 2010), 211–249. Berlin: de Gruyter. DOI: 10.1515/9783110262339.211.

Müller, Stefan. 1996a. The Babel-System: An HPSG fragment for German, a parser, and a dialogue component. In Peter Reintjes (ed.), *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, 263–277. London: Practical Application Company.

Müller, Stefan. 1996b. Yet another paper about partial verb phrase fronting in German. In Jun-ichi Tsujii (ed.), *Proceedings of COLING-96: 16th International Conference on Computational Linguistics (COLING96). Copenhagen, Denmark, August 5–9, 1996*, 800–805. Copenhagen, Denmark: Association for Computa-tional Linguistics.

Müller, Stefan. 1999a. An HPSG-analysis for free relative clauses in German. *Grammars* 2(1). 53–105. DOI: 10.1023/A:1004564801304.

Müller, Stefan. 1999b. *Deutsche Syntax deklarativ: Head-Driven Phrase Struc-ture Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783110915990.

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

Müller, Stefan. 2003. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62.

Müller, Stefan. 2004. Elliptical constructions, multiple frontings, and surface-based syntax. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 91–109. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/FG/2004/ (3 February, 2021).

Müller, Stefan. 2005. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203. 297–330.

Müller, Stefan. 2006. Phrasal or lexical constructions? *Language* 82(4). 850–883. DOI: 10.1353/lan.2006.0213.

Müller, Stefan. 2007b. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 1st edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

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, CA: CSLI Publications. DOI: 10.21248/hpsg.2009.11.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655. DOI: 10.1017/S0022226709990284.

Müller, Stefan. 2012. *On the copula, specificational constructions and type shifting*. Ms. Freie Universität Berlin.

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

Müller, Stefan. 2015. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI: 10.15398/jlm.v3i1.91.

Müller, Stefan. 2018. *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. DOI: 10.5281/zenodo.1441351.

Müller, Stefan. 2020. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 4th edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI: 10.5281/zenodo.3992307.

Müller, Stefan. 2021a. Constituent order. 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 9), 369–417. Berlin: Language Science Press. DOI: 10.5281/zenodo.5599836.

Müller, Stefan. 2021b. *German clause 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. 2021c. 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 9), 1497–1553. Berlin: Language Science Press. DOI: 10.5281/zenodo.5599882.

Müller, Stefan. 2022. *Germanic syntax* (Textbooks in Language Sciences). To appear. Language Science Press. https://hpsg.hu-berlin.de/~stefan/Pub/germanic.html (23 March, 2022).

Müller, Stefan & Antonio Machicao y Priemer. 2019. Head-Driven Phrase Structure Grammar. In András Kertész, Edith Moravcsik & Csilla Rákosi (eds.), *Current approaches to syntax: A comparative handbook* (Comparative Handbooks of Linguistics 3), 317–359. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110540253-012.

Müller, Stefan & Bjarne Ørsnes. 2013. 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. DOI: 10.1007/978-3-642-39998-5_5.

Muysken, Pieter. 1982. Parametrizing the notion of "head". *Journal of Linguistic Research* 2. 57–75.

Netter, Klaus. 1994. Towards a theory of functional heads: German nominal phrases. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 297–340. Stanford, CA: CSLI Publications.

Netter, Klaus. 1998. *Functional categories in an HPSG for German* (Saarbrücken Dissertations in Computational Linguisticsand Language Technology 3). Saarbrücken: German Research Center for Artificial Intelligence (DFKI) & University of the Saarland.

Newmeyer, Frederick J. 2004. Against a parameter-setting approach to typological variation. *Linguistic Variation Yearbook* 4. 181–234. DOI: 10.1075/livy.4.06new.

Newmeyer, Frederick J. 2005. *Possible and probable languages: A Generative perspective on linguistic typology* (Oxford Linguistics). Oxford: Oxford University Press.

Ng, Say Kiat. 1997. *A double-specifier account of Chinese NPs using Head-Driven Phrase Structure Grammar*. University of Edinburgh, Department of Linguistics. (MSc Speech and Language Processing).

Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538. DOI: 10.2307/416483.

O'Grady, William. 1998. The syntax of idioms. *Natural Language & Linguistic Theory* 16(2). 279–312. DOI: 10.1023/A:1005932710202.

Olsen, Susan. 1987. Zum sogenannten substantivierten Adjektiv im Deutschen: Deutsch als eine pro-drop Sprache. *Studium Linguistik* 21. 1–35.

Olsen, Susan. 1991. Die deutsche Nominalphrase als „Determinansphrase". In Susan Olsen & Gisbert Fanselow (eds.), *DET, COMP und INFL: Zur Syntax funktionaler Kategorien und grammatischer Funktionen* (Linguistische Arbeiten 263), 35–56. Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783111353838.35.

Osborne, Timothy & Thomas M. Groß. 2012. Constructions are catenae: Construction Grammar meets Dependency Grammar. *Cognitive Linguistics* 23(1). 165–216. DOI: 10.1515/cog-2012-0006.

Partee, Barbara H. 1986. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk, Dick de Jongh & Martin Stokhof (eds.), *Studies in Discourse Representation Theory and the theory of generalized quantifiers* (Groningen-Amsterdam Studies in Semantics 8), 115–143. Dordrecht: Foris Publications. Reprint as: Noun Phrase Interpretation and Type-Shifting Principles. In Barbara H. Partee (ed.), *Compositionality in formal semantics: Selected papers by Barbara H. Partee* (Explorations in Semantics 1), 203–230. Dordrecht: Blackwell Publishers Ltd, 2004. DOI: 10.1002/9780470751305.ch10.

Pollard, Carl J. 1996. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Published version of a Ms. dated January 1990. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110873467.279.

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

Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics 4). Chicago, IL: The University of Chicago Press.

Radford, Andrew. 1997. *Syntactic theory and the structure of English: A Minimalist approach* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9781139166706.

Radford, Andrew. 2004. *English syntax: An introduction.* Cambridge, UK: Cambridge University Press.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–483. DOI: 10.1017/S002222679700652X.

Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford Department of Linguistics and CSLI's LinGO Lab*, 394–414. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2007.23.

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, CA: CSLI Publications.

Sag, Ivan A., Hans C. Boas & Paul Kay. 2012. Introducing Sign-Based Construction Grammar. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 1–29. Stanford, CA: 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 & Thomas Wasow. 2020. Lessons from the English auxiliary system. *Journal of Linguistics* 56(1). 87–155. DOI: 10.1017/S002222671800052X.

Sag, Ivan A. & Janet Dean Fodor. 1995. Extraction without traces. In Raul Aranovich, William Byrne, Susanne Preuss & Martha Senturia (eds.), *WCCFL 13: The Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, 365–384. Stanford, CA: CSLI Publications/SLA.

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

Salzmann, Martin. 2020. The NP vs. DP debate. Why previous arguments are inconclusive and what a good argument could look like. Evidence from agreement with hybrid nouns. *Glossa: a journal of general linguistics* 5(1). 83. 1–46. DOI: 10.5334/gjgl.1123.

Salzmann, Martin. 2022. The NP vs. DP-debate and notions of headedness. In Ulrike Freywald, Horst J. Simon & Stefan Müller (eds.), *Headedness and/or grammatical anarchy?* 55–71. Berlin: Language Science Press. DOI: 10.5281/zenodo.7142664.

Simon, Horst J. 2003. *Für eine grammatische Kategorie 'Respekt' im Deutschen: Synchronie, Diachronie und Typologie der deutschen Anredepronomina* (Linguistische Arbeiten 474). Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783110949513.

Toivonen, Ida. 2013. English benefactive NPs. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '13 conference, University of Debrecen*, 503–

523. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/18/papers/lfg13toivonen.pdf (10 February, 2021).

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Van Eynde, Frank. 2006. NP-Internal agreement and the structure of the noun phrase. *Journal of Linguistics* 42(1). 139–186. DOI: 10.1017/S0022226705003713.

Van Eynde, Frank. 2020. Agreement, disagreement and the NP vs. DP debate. *Glossa: a journal of general linguistics* 5(1). 65. 1–23. DOI: 10.5334/gjgl.1119.

Van Eynde, Frank. 2021. Nominal structures. 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 9), 275–313. Berlin: Language Science Press. DOI: 10.5281/zenodo.5599832.

Van Langendonck, Willy. 1994. Determiners as heads? *Cognitive Linguistics* 5(3). 243–259. DOI: 10.1515/cogl.1994.5.3.243.

Vennemann, Theo & Ray Harlow. 1977. Categorial Grammar and consistent basic VX serialization. *Theoretical Linguistics* 4(1–3). 227–254. DOI: 10.1515/thli.1977.4.1-3.227.

Wang, Lulu & Haitao Liu. 2007. A description of Chinese NPs using Head-Driven Phrase Structure Grammar. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford Department of Linguistics and CSLI's LinGO Lab*, 287–305. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2007.17.

Webelhuth, Gert. 1995. X-bar Theory and Case Theory. In Gert Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program: Principles and Parameters in syntactic theory* (Generative Syntax 1), 15–95. Oxford: Blackwell Publishers Ltd.

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

Wunderlich, Dieter. 1987. Vermeide Pronomen – Vermeide leere Kategorien. *Studium Linguistik* 21. 36–44.

Zlatić, Larisa. 2014. Definiteness and structure of NPs in Slavic. In Lilia Schürcks, Anastasia Giannakidou & Urtzi Etxeberria (eds.), *The nominal structure in Slavic and beyond* (Studies in Generative Grammar 116), 17–38. Berlin: De Gruyter Mouton. DOI: 10.1515/9781614512790.17.

Zwicky, Arnold M. 1985. Heads. *Journal of Linguistics* 21(1). 1–29. DOI: 10.1017/S0022226700010008.