Chapter 33

HPSG and Minimalism

Robert D. Borsley

University of Essex

Stefan Müller

Humboldt-Universität zu Berlin

This paper compares work done under the heading *Minimalist Program* with work done in Head-Driven Phrase Structure Grammar. We discuss differences in the respective approaches and the outlook of theories. We have a look on the procedural/constraint-based views on grammar and discuss the differences in complexity of the structures that are assumed. We also address psycholinguistic issues like processing and language acquisition.

1 Introduction

The Minimalist framework, which was first outlined by Chomsky in the early 1990s (Chomsky 1993; 1995b), still seems to be the dominant approach to syntax. It is important, therefore, to consider how HPSG compares with this framework. In a sense both frameworks are descendants of the transformation-generative approach to syntax, which Chomsky introduced in the 1950s. HPSG is a result of the questioning of transformational analyses that emerged in the late 1970s. This led to Lexical Functional Grammar (Bresnan & Kaplan 1982) and Generalized Phrases Structure Grammar (Gazdar et al. 1985), and then in the mid 1980s to HPSG (Pollard & Sag 1987, see Flickinger, Pollard & Wasow (2018), Chapter 2 of this volume for more on the origins of HPSG). Minimalism in contrast remains committed to transformational, i.e. movement, analyses. It is simpler in some respects than the earlier Government and Binding framework (Chomsky 1981), but as we will see below, it involves a variety of complexities.

The relation between the two frameworks is clouded by the rhetoric that surrounds Minimalism. At one time "virtual conceptual necessity" was said to be its guiding principle. A little later, it was said to be concerned with the "perfection of language", with "how closely human language approaches an optimal solution to design conditions that the system must meet to be usable at all" (Chomsky 2002: 58). Much of this rhetoric seems designed to suggest that Minimalism is quite different from other approaches and should not be assessed in the same way. In the words of Postal (2003: 19), it looks like "an attempt to provide certain views with a sort of privileged status, with the goal of placing them at least rhetorically beyond the demands of serious argument or evidence". However, the two frameworks have enough in common to allow meaningful comparisons.

Both frameworks seek to provide an account of what is and is not possible both in specific languages and in language in general. Moreover, both are concerned not just with local relations such as that between a head and its complement or complements but also with non-local relations such as those in the following:

(1) a. The student knows the answer.
     b. It seems to be raining.
     c. Which student do you think knows the answer?

In (1a), *the student* is subject of *knows* and is responsible for the fact that *knows* is a third person singular form, but *the student* and *knows* are not sisters if they form a VP. In (1b) the subject is *it* because the complement of *be* is *raining* and *raining* requires an expletive subject, but *it* and *raining* are obviously not sisters. Finally, in (1c), *which student* is understood as the subject of *knows* and is responsible for the fact that it is third person singular, but again the two elements are structurally quite far apart. Both frameworks provide analyses for these and other central syntactic phenomena, and it is quite reasonable to compare them and ask which is the more satisfactory.[1]

Although HPSG and Minimalism have enough in common to permit comparisons, there are obviously many differences. Some are more important than others, and some relate to the basic approach and outlook, while others concern the nature of grammatical systems and syntactic structures. In this chapter we will explore the full range of differences.

The chapter is organized as follows: in Section 2, we look at differences of approach between the two frameworks. Then in Section 3, we consider the quite different views of grammar that the two frameworks espouse, and in Section 4,

---

[1]As noted below, comparison is complicated somewhat by the fact that Minimalists typically provides only sketches of analyses in which various details are left quite vague.

we look at the very different syntactic structures which result. Finally, in Section 5, we consider how the two frameworks relate to psycholinguistic issues, namely processing and language acquisition.

2 Differences of approach and outlook

> add brief intro

2.1 Sociological differences

A difference between MGG and HPSG that should not be underestimated is the fact that MGG has one very important key figure, the by far most-cited linguist (and scientist in general) Noam Chomsky. This is how Carl Pollard put it:

> There is no Chomsky-like figure who is always assumed to be basically on the right track no matter what s/he proposes. HPSG research is normal science: the testing of hypotheses that appear plausible given accepted assumptions. The goal is not to fill in the details of a vague theory which is assumed to be basically right, but to successively replace empirical hypothesis with ones that make better predictions. (Pollard 1997: Section 1.2.1)

While Ivan Sag is perceived as the leading figure by outsiders, the fact that Ivan Sag's recent switch to Sign-Based Construction Grammar (Sag 2012) is not unanimously adopted by the HPSG community (see Müller (2018), Chapter 37 of this volume and Müller (2018b: Section 10.6.2) for discussion), that empty elements are assumed by various researchers (Bender 2000; Müller 2014; Borsley 1999; 2009; 2013) even though Ivan Sag argued against them since 1994 (Sag & Fodor 1994) and that defaults are not used by all researchers even though they play a prominent role in Ivan Sag's work (e.g., in Ginzburg & Sag 2000) shows that Pollard's claim is correct. Of course Ivan Sag is the most influential HPSG grammarian but this influence is based on properly argued empirically interesting research.

> one Borsley paper is enough. which one?

2.2 Formalization and exhaustivity

As many of the chapters in this volume emphasize, HPSG is a framework which places considerable emphasis on detailed formal analyses of the kind that one

might expect within Generative Grammar.[2] Thus, it is not uncommon to find lengthy appendices setting out formal analyses. See, for example, Sag's (1997) paper on English relative clauses, Van Eynde's (2015) book on predicative constructions and especially Ginzburg & Sag (2000), which has a 50 page appendix. One consequence of this is that HPSG has had considerable influence in computational linguistics. Sometimes theoretical work comes paired with computer implementations, which show that the analyses are consistent and complete (e.g., all publications coming out of the CoreGram project (Müller 2015b) and the HPSG textbook for German that comes with implementations corresponding to the individual chapters of the book (Müller 2007b)). For more on the relation of HPSG and computational linguistics see Bender & Emerson (2018), Chapter 29 of this volume.

In Minimalism things are very different. Detailed formal analyses are virtually non-existent. There appear to be no appendices like those in Sag (1997) and Ginzburg & Sag (2000). In fact the importance of formalization has long been downplayed in Chomskyan work (e.g., by Chomsky in an interview with Huybregts & Riemsdijk (1982: 73) and in discussions between Pullum (1989) and Chomsky (1990: 146)), and this view seems fairly standard within Minimalism (see also the discussion in Müller (2016a: Section 3.6.2)). Chomsky & Lasnik (1995: 28) attempt to justify the absence of detailed analyses when they suggest that providing a rule system from which some set of phenomena can be derived is not "a real result" since "it is often possible to devise one that will more or less work". Instead, they say, "the task is now to show how the phenomena [...] can be deduced from the invariant principles of UG with parameters set in one of the permissible ways". In other words, providing detailed analyses is a job for unambitious drudges, and real linguists pursue a more ambitious agenda. Postal (2004: 5) comments that what we see here is "the fantastic and unsupported notion that descriptive success is not really that hard and so not of much importance". He points out that if this were true, one would expect successful descriptions to be abundant

---

[2]We follow Ginzburg & Sag (2000: 2) in counting HPSG among Generative Grammar in the sense it was defined by Chomsky (1965: 4), namely as a framework that provides an explicit characterization of the theories developed within it. When we refer to work in Government & Binding or Minimalism, we follow Culicover & Jackendoff (2005: 3) in using the term *Mainstream Generative Grammar* (MGG). It should be kept in mind that there is another meaning associated with the term *generative*. A generative grammar in the latter sense generates a set (Chomsky 1957: 13. HPSG is not generative in this sense but rather model-theoretic. See Pullum & Scholz (2001) for differences between generative-enumerative and model theoretic approaches. See also Richter (2018), Chapter 3 of this volume and Wasow (2018), Chapter 28 of this volume.

within transformational frameworks. However, he suggests that "the actual descriptions in these frameworks so far are not only not successful but so bad as to hardly merit being taken seriously". Postal does much to justify this assessment with detailed discussions of Chomskyan work on strong crossover phenomena and passives in Chapters 7 and 8 of his book.

There has also been a strong tendency within Minimalism to focus on just a subset of the facts in whatever domain is being investigated. As Culicover & Jackendoff (2005: 535) note, "much of the fine detail of traditional constructions has ceased to garner attention". This tendency has sometimes been buttressed by a distinction between core grammar, which is supposedly a fairly straightforward reflection of the language faculty, and a periphery of marked constructions, which are of no great importance and which can reasonably be ignored. However, as Culicover (1999) and others have argued, there is no evidence for a clear cut distinction between core and periphery. It follows that a satisfactory approach to grammar needs to account both for such core phenomena as *wh*-interrogatives, relative clauses, and passives but also with more peripheral phenomena such as the following:

(2)   a. It's amazing the people you see here.
      b. The more I read, the more I understand.
      c. Chris lied his way into the meeting.

These exemplify the nominal extraposition construction (Michaelis & Lambrecht 1996), the comparative correlative construction (Culicover & Jackendoff 1999; Borsley 2011), and the *X's Way* construction (Salkoff 1988; Sag 2012). As has been emphasized in other chapters, the HPSG system of types and constraints is able to accommodate broad linguistic generalizations and highly idiosyncratic facts and everything in between.

The general absence in Minimalism of detailed formal analyses is quite important. It means that Minimalists may not be fully aware of the complexity of the structures they are committed to and this allows them to sidestep the question whether it is really justified. It also allows them to avoid the question of whether the very simple conception of grammar that they favour is really satisfactory. Finally, it may be that they are unaware of how many phenomena remain unaccounted for. These are all important matters.

The general absence of detailed formal analyses has also led to Minimalism having little impact on computational linguistics. There has been some work that has sought to implement Minimalist ideas (Stabler 2001; Fong & Ginsburg 2012; Fong 2014), but Minimalism has not had anything like the productive relation

with computational work that HPSG has enjoyed. Existing Minimalist implementations are rather toy grammars analyzing very simple sentences, some are not faithful to the theories they are claimed to be implementing,[3] and some do not even parse natural language but require pre-segmented, pre-formatted input. For example, Stabler's test sentences have the form as in (3).

(3)   a. the king will -s eat
          b. the king have -s eat -en
          c. the king be -s eat -ing
          d. the king -s will -s have been eat -ing the pie

See Müller (2016a: Section 4.7.2) for discussion.

2.3 Empirical quality

There are, then, issues about the quantity of data that is considered in Minimalist work. There are also issues about its quality. Research in HPSG is typically quite careful about data and often makes use of corpus and experimental data (see for example An & Abeillé 2017; Müller 1999b; 2002; Bildhauer & Cook 2010; Müller, Bildhauer & Cook 2012; Van Eynde 2015: Chapter 7 for examples of work with attested examples and for experimental work). This use of corpus data and attested examples is based on the insight that introspection alone is not sufficient, given that an enormous amount of time is spent to work out analyses and it would be unfortunate if these analyses were built on a shaky empirical basis. See Müller (2007a) and Meurers & Müller (2009) for the discussion of introspection vs. corpus data. Research in Minimalism is often rather less careful.[4] In a review of a collection of Minimalist papers, Bender (2002: 434) comments that: "In these papers, the data appears to be collected in an off-hand, unsystematic way, with unconfirmed questionable judgments often used at crucial points in the argumentation". She goes on to suggest that the framework encourages "lack of concern for the data, above and beyond what is unfortunately already the norm in formal syntax, because the connection between analysis and data is allowed to be

[add more references]

---

[3]Fong's grammars are simple Definite Clause Grammars, that is, context-free phrase structure grammars, and hence nowhere near an implementation of Minimalism, contrary to claims by Berwick, Pietroski, Yankama & Chomsky (2011: 1221).

[4]We hasten to say that we do not claim this to be true for all Minimalist work. There are researchers working with corpora or at least with attested examples (Wurmbrand 2003) and there is experimental work. Especially in Germany there were several large scale Collaborative Research Centers with a strong empirical focus which also fed back into theoretical work, including Minimalist work. The fact that we point out here is that there is work, including work by prominent Minimalists, that is rather sloppy as far as data is concerned.

remote.". Similar things could be said about a variety of Minimalist work. Consider, for example, Aoun & Li (2003), who argue for quite different analyses of *that*-relatives and *wh*-relatives on the basis of the following (supposed) contrasts, which appear to represent nothing more than their own judgements (p. 110–112):

(4) a. The headway that Mel made was impressive.

   b. ?? The headway which Mel made was impressive.

(5) a. We admired the picture of himself that John painted in art class.

   b. * We admired the picture of himself which John painted in art class.

(6) a. The picture of himself that John painted in art class is impressive.

   b. *? The picture of himself which John painted in art class is impressive.

None of the native speakers we have consulted find significant contrasts here which could support different analyses. The example in (7a) with a *which* relative clause referring to *headway* can be found in Cole et al. (1982), Williams (1989: 437) and Falk (2010: 221) have examples with a reflexive coreferential with a noun in a relative clause introduced by *which* as in William's (7b) and corpus examples like (7c,d) can be found as well:

(7) a. The headway which we made was satisfactory.

   b. the picture of himself which John took

   c. The words had the effect of lending an additional clarity and firmness of outline to the picture of himself which Bill had already drawn in his mind–of a soulless creature sunk in hoggish slumber.[5]

   d. She refused to allow the picture of himself, which he had sent her, to be hung, and it was reported that she ordered all her portraits and busts of him to be put in the lumber attics.[6]

Given that it is relatively easy to come up with counter examples it is surprising that authors do not do a quick check before working out rather complex analyses.

   Another quite important example of ignoring relevant data is Chomsky's work on Labeling.

---

[5]Wodehouse, P.G. 1917. *Uneasy Money*, London: Methuen & Co., p. 186, http://www.literaturepage.com/read.php?titleid=uneasymoney&abspage=186, 2018-09-18.

[6]Jerrold, Clare. 1913. *The married life of Queen Victoria*, London: G. Bells & Sons, Ltd. https://archive.org/stream/marriedlifeofque00jerruoft/marriedlifeofque00jerruoft_djvu.txt, 2018-09-19.

As I say elsewhere I think any detailed discussion of labelling should be in section 3 though it seems reasonable to make the point here that the discussion of these issues has been marred by a failure to notice some fairly obvious examples.

Labeling is concerned with the question of the category that results when two items are combined. As is discussed in more detail below, Chomsky assumes that the combination of two items with the labels X and Y has to be either X or Y. Chomsky defines which label is used in two different cases: the first case defines that the label is the label of the head if the head is a lexical item and the second case defines the label to be the label of the category from which something is extracted (Chomsky 2008: 145). As Chomsky notes, these rules are not unproblematic since the label is not uniquely determined in all cases. An example is the combination of two lexical elements since in such cases both elements can be the label of the resulting structure. Chomsky notices that this could result in deviant structures, but claims that this concern is unproblematic and ignores it. This means that rather fundamental notions in a grammar theory were ill-defined. A solution to this problem was provided in his 2013 paper (published five years later). However, this paper is inconsistent (Müller 2016a: Section 4.6.2) and even insiders find it incomprehensible. But this is not the point we want to focus on here. Rather we want to show one more time that empirical standards are not met. Chomsky uses underdetermination in his labeling rules to account for two possible structures in (8), an approach going back to Donati (2006):

(8)   what [C [you wrote *t*]]

(8) can be an interrogative clause as in *I wonder what you wrote.* or a free relative clause as in *I will read what you wrote.*. According to the labeling rule that accounts for sentences from which an item is extracted, the label will be CP since the label is taken from the clause. However, since *what* is a lexical item, *what* can determine the label as well. If this labeling rule is applied *what you wrote* is assigned DP as a label and hence the clause can function as a DP argument of *read.*

Chomsky's proposal is interesting but it does not extend to cases involving free relative clauses with complex *wh*-phrases (so-called pied-piping) as they are attested in examples like (9):

(9)   I'll read [whichever book] you give me.

The example in (9) is from one of the standard references on free relative clauses: Bresnan & Grimshaw (1978: 333), which is also cited in other mainstream generative work as for example (Groos & van Riemsdijk 1981).

Draft of 26th December 2018, 19:51

Apart from the fact that complex *wh*-phrases are possible there is even more challenging data in the area of free relative clauses: the examples in (10) and (11) show that there are non-matching free relative clauses:

(10) Sie kocht, worauf   sie Appetit hat.[7]   (German)
      she cooks where.on she appetite has
      'She cooks what she feels like eating.'

(11) a. Worauf   man sich mit einer Pro-form beziehen kann, […] ist
          where.upon one self with a   Pro-form refer   can   is
          eine Konstituente.[8]
          a   constituent
          'If you can refer to something with a Pro-form, […] it is a constituent.'

      b. [Aus wem] noch etwas   herausgequetscht werden kann, ist
           out who   yet something out.squeezed   be   can   is
          sozial dazu   verpflichtet, es abzuliefern; …[9]
          socially there.to obliged   it to.deliver
          'Those who have not yet been bled dry are socially compelled to hand
          over their last drop.'

In (10) a relative clause with a PP relative phrase functions as an accusative object. In (11) the relative clauses function as subjects. (11b) is another example of a relative clause with a complex *w* phrase. See Bausewein (1990) and Müller (1999a) for further discussion of free relative clauses and attested data.

According to Donati (2006: Section 5), pied piping does not exist in free relatives (see also Citko (2008: 930–932) for a rejection of this claim). Given how much attention the issue of Labeling got and how central this is to Minimalist analyses this situation is quite surprising: an empirically wrong claim made in 2002/2003[10] is the basis for foundational work from 2002 until 2013 even though the facts are common knowledge in the field. Ott (2011) develops an analysis in which the category of the relative phrase is projected, but he does not have a solution for nonmatching free relative clauses, as he admits in a footnote on

---

[7]Bausewein (1990: 154).

[8]From the main text of: Günther Grewendorf, *Aspekte der deutschen Syntax. Eine Rektions-Bindungs-Analyse.* Studien zur deutschen Grammatik, number 33. Tübingen: Gunter Narr Verlag. 1988, p. 16, quoted from Müller (1999a: 61).

[9]Wiglaf Droste, taz, 01.08.97, p. 16, quoted from Müller (1999a: 61).

[10]Versions of Donati's paper were presented at *Going Romance* in 2002 and the most important Mainstream Generative Grammar conference in Europe GLOW in 2003.

page 187. The same is true for Citko's analysis (2008), in which the extracted XP can provide the label. So, even though the data is known for decades, it is ignored by authors and reviewers and foundational work is build on shaky empirical grounds. See Müller (2016a: Section 4.6.2) for a more detailed discussion of labeling.

Note that we are not just taking one bad example of Minimalist work. It is probably the case that papers with dubious judgments can be found in any framework if only this is due to the repetitions of unwarranted claims made by others. The point is that Aoun & Li are influential (quoted by 455 other publications as of 2018-09-14). Others rely on these judgments or the analyses that were motivated by them. New conclusions are derived from analyses since theories make predictions. If this process continues for a while an elaborated theoretical building results that is not empirically supported. Note furthermore that the criticism raised here is not the sqabble of two authors working in an alternative framework. This criticism also comes from practitioners of Mainstream Generative Grammar. For example, Wolfgang Sternefeld and Hubert Haider, both very prominent figures in the German Generative Grammar school criticized the scientific standards in Minimalism heavily (Sternefeld & Richter 2012; Haider 2018). As Sternefeld & Richter (2012: 266–268) point out it is not just the case that Minimalist publications are based on empirically problematic foundations, it is even worse: researchers like Epstein & Seely (2006: Section 1.1) publish statements about research methodology that can only be understood as an immunization strategy. The authors discuss the two curves in Figure 1: one has a sinus-like shape and hits three out of seven of the data points on a straight line and the other one is parallel to the line with the data points and hence hits none of the data points. The second approach gets none of the data right. Nevertheless the authors argue that one should prefer this theory since it was closer to the truth and more illuminating: "Clearly, Theory 1 is 'empirically preferable' by a 'winning score' of 3–0. The point [...] is that Theory 2, despite getting none of the data correct, [...] is 'closer to the truth' or 'more illuminating' than the empirically preferable Theory 1. Hence, we believe, Theory 2 is a better working or guiding hypothesis upon which to base future research." (Epstein & Seely 2006: 2)

2.4 Argumentation for invisible entities and the assumption of innate linguistic knowledge

There are also differences in the kind of arguments that the two frameworks find acceptable. It is common within Minimalism to assume that some phenomenon which cannot be readily observed in some languages must be part of their gram-
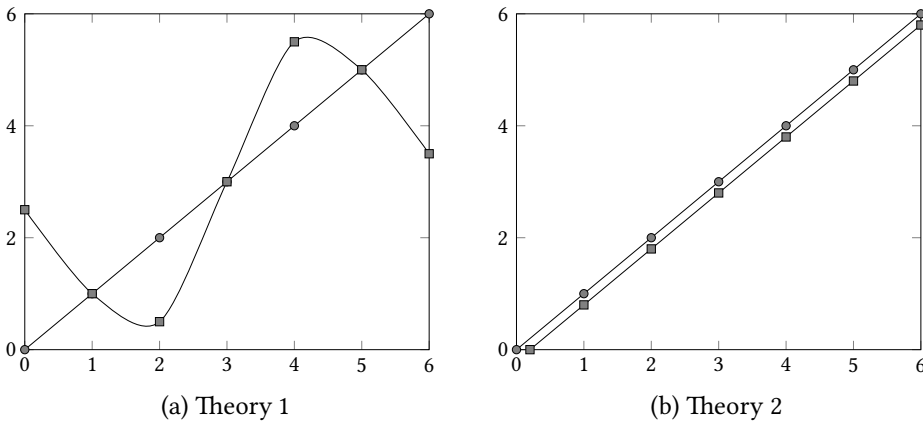
Figure 1: Immunization in Minimalist work: According to Epstein & Seely (2006: Section 1.1) Theory 2, which gets none of the data points, is preferable to Theory 1 with three data points since Theory 2 is "more illuminating" and "closer to to the truth"

matical system because it is clearly present in other languages. Notable examples would be case (Li 2008) or (object) agreement (Meinunger 2000: Chapter 4), which are assumed to play a role even though there are no visible manifestations within some languages (e.g., Mandarin Chinese and German, respectively). This stems from the longstanding Chomskyan assumption that language is the realization of a complex innate language faculty. From this perspective, there is much in any grammatical system that is a reflection of the language faculty and not in any simple way of the observable phenomena of the language in question. If some phenomenon plays an important role in many languages it is viewed as a reflection of the language faculty, and hence it must be a feature of all grammatical systems even those in which it is hard to see any evidence for it. An example – taken from a textbook on Minimalism (Hornstein, Nunes & Grohmann 2005: 124) – is an analysis of prepositional phrases in English. Figure 2 shows the analysis.[11] Due to theory internal assumptions the case requirement of the preposition cannot be checked in the P-DP combination. According to the version of the theory adopted by the authors, case has to be checked in specifier positions. Therefore it was assumed that the preposition moves to an Agr head and the DP moves to the specifier position of this Agr head. The problem is of course that DP and

---

[11]This analysis is actually a much simpler variant of the PP analysis which appeared in an earlier textbook by Radford (1997: 452). For discussion of this analysis see Sternefeld (2006: 549–550) and Müller (2016a: Section 4.6.1.2).

Figure 2: Minimalist analysis of a PP according to (Hornstein, Nunes & Grohmann 2005: 124) and the analysis assumed in HPSG and all other phrase structure-based frameworks

P are in the wrong order now. However, the authors argue that this is the order that is manifested in Hungarian and that Hungarian is a language which has postpositions and these are agreeing with their nominal dependent. The authors assume that Hungarian postpositions are prepositions underlyingly and that the DP following the preposition moves to the left because of a movement process that is triggered by agreement. It is claimed that this movement exists both in Hungarian and in English but that the movement is covert (that is, invisible) in the latter language.

This line of argument would be reasonable if a complex innate language faculty was an established fact, but it isn't, and since Hauser, Chomsky & Fitch (2002), it seems to have been rejected within Minimalism. It follows that ideas about an innate language faculty should not be used to guide research on individual languages. Rather, as Müller (2015b: 25) puts it, "grammars should be motivated on a language-specific basis". Does this mean that other languages are irrelevant when investigating a specific language? Clearly not. As Müller also puts it, "In situations where more than one analysis would be compatible with a given dataset for language X, the evidence from language Y with similar constructs is most welcome and can be used as evidence in favor of one of the two analyses for language X." (2015b: 43). In practice, any linguist working on a new language will use apparently similar phenomena in other languages as a starting point.

It is important, however, to recognize that apparently similar phenomena may turn out on careful investigation to be significantly different.[12]

3 Different views of grammar

We turn now to more substantive differences between HPSG and Minimalism, differences in their conceptions of grammar, especially syntax, and differences in their views of syntactic structure. As we will see, these differences are related. In this section we consider the former, and in the next we will look at the latter.

3.1 Constraint-based vs. derivational and generative-enumerative approaches

As is emphasized throughout this volume, HPSG assumes a declarative or constraint-based view of grammar. It also assumes that the grammar involves a complex systems of types and constraints. Finally, it assumes that syntactic analyses are complemented by separate semantic and morphological analyses. In each of these areas, Minimalism is different. It assumes a procedural view of grammar. It assumes that grammar involves just a few general operations. Finally, it assumes that semantics and morphology are simple reflections of syntax. We comment on each of these matters in the following subsections.

Whereas HPSG is a declarative or constraint-based approach, Minimalism seems to be firmly committed to a procedural approach. Chomsky (1995b: 219) remarks that: "We take L [a particular language] to be a generative procedure that constructs pairs (π, λ) that are interpreted at the articulatory-perceptual (A-P) and conceptual-intentional (C-I) interfaces, respectively, as 'instructions' to the performance systems.". Various arguments have been presented within HPSG for a declarative view, but no argument seems to be offered within Minimalism for a procedural view. Obviously, speakers and hearers do construct representations and must have procedures that enable them to do so, but this is a matter of performance, and there is no reason to think that the knowledge that is used in performance has a procedural character (see Section 5.1 on processing). Rather, the fact that it is used in both production and comprehension suggests that it should be neutral between the two and hence declarative.

---

[12]Equally, of course, apparently rather different phenomena may turn out on careful investigation to be quite similar. For further discussion of HPSG and comparative syntax, see Borsley (2018).

Another difference between constraint-based and generative-enumerative approaches is that the first type of proposals provides a way to get graded acceptability into the picture (Pullum & Scholz 2001: Section 3.1). Since HPSG grammars are basically feature-value pairs with equality (or other relations) between values, it is possible to weigh constraints and it is possible to admit constraint violations and work with structures with violated constraints. So looking at the sentences in (12) we see that more and more constraints are violated:

> Add foot-note on cumulative constraint violation.

(12) a.   I am the chair of my department.

 b. * I are the chair of my department.

 c. * Me are the chair of my department.

 d. * Me are the chair of me's department.

 e. * Me are chair the of me's department.

 f. * Me are chair the me's department of.

In comparison to this a generative-enumerative grammar enumerates a set and a sequence either is in the set or it is not.[13]

For further discussion of the issues, see Section 5.1 of this paper and e. g. Pullum & Scholz (2001), Postal (2003), Sag & Wasow (2011; 2015), and Wasow (2018), Chapter 28 of this volume.

3.2 Underspecification

> work this out

Minimalism: Only what is explicitly ruled in works. HPSG: Everything that is not ruled out works.

3.3 Types and constraints vs. general operations

The declarative-procedural contrast is an important one, but the contrast between the complex systems of types and constraints that are assumed within HPSG and the few general operations that form a Minimalist grammar is arguably more important.[14] Much work in Minimalism has three main operations Merge, Agree, and Move or Internal Merge. Merge combines two expressions, either words or phrases, to form a larger expression with the same label as one of

---

[13]For a discussion of Chomsky's (1964; 1975: Chapter 5) proposals to deal with different degrees of acceptability see Pullum & Scholz (2001: 29).

[14]A procedural approach doesn't necessarily involve a very simple grammatical system. The Standard Theory of Transformational Grammar (Chomsky 1965) is procedural but has many different rules, both phrase structure rules and transformations.

the expressions (Chomsky 1995b: 244; 2008: 140). Its operation can be presented as shown in Figure 3. In the case of English, the first alternative is represented
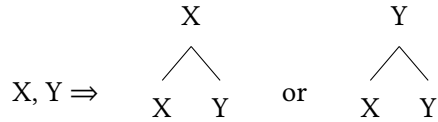

Figure 3: Merge

by situations where a lexical head combines with a complement while the second is represented by situations where a specifier combines with a phrasal head. Chomsky (2008: 146) calls items merged with the first variant of Merge *first-merged* and those merged with the second variant *later-merged*.

Agree, as one might suppose, offers an approach to various kinds of agreement phenomena. It involves a probe, which is a feature or features of some kind on head, and a goal, which the head c-commands.[15] At least normally, the probe is a linguistic object with an uninterpretable feature or features with no value and the goal has a matching interpretable feature or features with appropriate values (Chomsky 2001: 3–5). Agree values the uninterpretable feature or features and they are ultimately deleted, commonly after they have triggered some morphological effect. Agree can be represented as in Figure 4 (where the "*u*" prefix identifies a feature as uninterpretable). Unsurprisingly subject-verb agreement

Figure 4: Agree

is one manifestation of Agree, where X is T(ense) and Y is a nominal phase, for

---

[15]It is not necessary to state the c-command restriction since it sis assumed that linguistic objects are built bottom-up and when a head is combined with its dependent the head can "see" the dependent but not the context in which the resulting object is embedded. Information about dominating material will be available only later in the derivation. (Guido Mensching, p. c. 2018)

Minimalism a DP, inside the complement of T. Here, and elsewhere, Agree is a non-local relation involving elements which are not sisters. This contrasts with the situation in HPSG, in which subject-verb agreement is a consequence of the subject and its VP sister sharing features and further feature sharing between the VP and the V that heads it.

Finally, Move, also called Internal Merge, is an operation which makes a copy of a constituent of some expression and merges it with the expression (Chomsky 1995b: Section 4.4; 2008: 140). The original element that is copied normally undergoes deletion. The process can be presented as in Figure 5. This covers both the

Figure 5: Move

A′-movement process assumed for unbounded dependency constructions such as *wh*-interrogatives and the A-movement process assumed for raising sentences and passives. A question arises about so-called head-movement, where a head moves to a higher head position. This appears to mean that it must be possible for the copy to be merged with the head of the expression that contains it. However, this is incompatible with the widely assumed extension condition, which requires Merge to produce a larger structure. One response is the idea espoused in Chomsky (1995a: 368; 2001: 37) that head-movement takes place not in the syntax but in the PF component, which maps syntactic representations to phonetic representations. It seems that its status is currently rather unclear.

The three operations just outlined interact with lexical items to provide syntactic analyses. It follows that the properties of constructions must largely derive from the lexical items that they contain. Hence, the properties of lexical items are absolutely central to Minimalism. Oddly, the obvious implication – that the lexicon should be a major focus of research – seems to be ignored. As Newmeyer (2005: 95, fn. 9) comments:

> […] in no framework ever proposed by Chomsky has the lexicon been as important as it is in the MP [Minimalist Program]. Yet in no framework

proposed by Chomsky have the properties of the lexicon been as poorly investigated. (Newmeyer 2005: 95, fn. 9)

Sometimes it is difficult to derive the properties of constructions from the properties of visible lexical elements. But there is a simple solution: postulate an invisible element. The result is a large set of invisible functional heads. As we will see in the next section, these heads do the work in Minimalism that is done by phrase types and the constraints on them in HPSG.

Although Minimalism is a procedural approach and HPSG a declarative approach, there are some similarities between Minimalism and early HPSG, the approach presented in Pollard & Sag (1987; 1994). In much the same way as Minimalism has just a few general mechanisms, early HPSG had just a few general phrase types. Research in HPSG in the 1990s led to the conclusion that this is too simple and that a more complex system of phrase types is needed to accommodate the full complexity of natural language syntax. Nothing like this happened within Minimalism, almost certainly because there was little attempt within this approach to deal with the full complexity of natural language syntax. As noted above, the approach has rarely been applied in detailed formal analyses. It looks too simple and it appears problematic in various ways. It is also a major source of the complexity that is characteristic of Minimalist syntactic structures, as we will see in the next section.

3.4 Feature deletion and "crashing at the interfaces"

We already mentioned uninterpretable features in the previous subsection. It is assumed in Minimalism that features like case features are not interpretable and hence have to be consumed since otherwise the derivation would "crash at the interfaces". This means that Minimalism claims that a case marked NP like *der Mann* 'the man' is not interpretable unless it is somehow stripped of its case information.So in Minimalism *der Mann* needs something on top of the DP that consumes the case features. While this seems cumbersome to most working outside Minimalism, there are actually deeper problems connected to the consumption of case features. There are situations in which you need case features more than once. An example for this are free relative clauses as the one in (13b):

(13) a. der   Mann
          the.NOM man

      b. Ich treffe, wen   ich treffen will.
          I   meet   who.ACC I   meet   want.to
          'I meet whoever I like to meet.'

wen is the accusative object in the relative clause. Since it is an object its case feature will be checked by the governing verb *treffen* 'meet'. *wen* will then be a DP without any case information. However, the case of the relative phrase in free relative clauses is not arbitrary. It is important for the integration of the free relative clause in the matrix clause. The case of *wer* 'who' in a complete relative clause has to be known since it is important for the external distribution of the free relative clause, as the examples in (14) show:

(14) a. Wer   mich treffen will,   kann vorbeikommen.
          who.NOM me   meet   wants.to may over.come
          'He who wants to meet me may come over.'

      b. * Ich treffe, wer   mich treffen will.
            I   meet who.NOM me   meet   wants.to
          'I meet whoever wants to meet me.'

HPSG also consumes resources in a way: items in valence representations are not projected up the tree once the requirement is saturated, but the difference is that objects with a certain structure and with certain features are not modified. A case-marked NP is not deprived of this case information for whatever reasons. We think that this is the right way to deal with morphological markings and with feature specifications in general.

3.5 Some implications

We will look in detail at the implications for syntactic structure of this machinery in the next section. However, we will note some implications in the following paragraphs as a kind of preview of the next section.

First, the fact that Merge combines two expressions entails that syntactic structures are confined to binary branching and excludes various analyses that have been assumed within HPSG and other frameworks. Second, the assumption that expressions produced by Merge have the same label as one of the expressions that they consist of (Chomsky 2008: 145) is essentially the assumption that all complex expressions are headed. For HPSG, as for many other approaches, there are headed expressions and non-headed expressions, e.g., coordination and the NPN construction discussed in Sections 4.2.2 and 4.2.3, respectively.

As emphasized above, a further important feature of Minimalism is the view that semantics and morphology are simple reflections of syntax. The basic archititecture assumed in Minimalism is shown in Figure 6. Both phonology and semantics are read off the structures produced by syntax. The idea that seman-

lexicon ⟶ numeration

PHON   SEM

lexicon

overt syntax

PHON   SEM

← Spell-Out

covert syntax

PHON   SEM

LF/CI   PF/AP
(meaning) (sound)

PHON   SEM
...

PHON   SEM

Transfer
Transfer
Transfer
Transfer
Transfer

Figure 6: Syntax-centristic architecture in Minimalism before the Phase model (left) and in the Phase model (right) according to Richards (2015: 812, 830))

tics is a simple reflection of syntax goes back to the early years of Transformational Grammar. One aspect of this idea was formalized as the Uniform Theta Assignment Hypothesis (UTAH) by Baker (1988: 46).

(15)   Uniform Theta Assignment Hypothesis
     Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.

Minimalism abandoned the notion of D-structure, but within Minimalism the Hypothesis can be reformulated as follows:

(16)   Uniform Theta Assignment Hypothesis (revised)
     Identical thematic relationships between items are represented by identical structural relationships between those items when introduced into the structure.

We will look at some of the implications of this in the next section.

The idea that morphology is a simple reflection of syntax is also important. As we will discuss in the next section, it leads to abstract underlying structures and complex derivations and to functional heads corresponding to various suffixes. Again, we will say more about this in the next section.

4 Different views of syntactic structure

The very different views of grammar that are assumed in Minimalism and HPSG naturally lead to very different views of syntactic structure. The syntactic structures of Minimalism are both very complex and very simple. This sounds paradoxical but it isn't. They are very complex in that they involve much more structure than those assumed in HPSG and other approaches. But they are very simple in that they have just a single ingredient – they consist entirely of local trees in which there is a head and a single non-head. From the standpoint of HPSG, they are both too complex and too simple. We will consider the complexity in Section 4.1 and then turn to the simplicity in Section 4.2.

4.1 The complexity of Minimalist structures

For HPSG, as the chapters in this volume illustrate, linguistic expressions have a single relatively simple constituent structure with a minimum of phonologically empty elements.[16] For Minimalism, they have a complex structure containing a variety of empty elements and with various constituents occupying more than one position in the course of the derivation. Thus the structures assumed within Minimalism are not at all minimalist. But this complexity is a more or less inevitable consequence of the Minimalist view of grammar outlined above.

4.1.1 Uniformity of structures due to semantic representation

There are a variety of sources of complexity, and some predate Minimalism.[17] This is true especially of the idea that semantics and morphology are simple reflections of syntax (on morphology see Section 4.1.3). For the syntax-semantics relation, UTAH, which we introduced on p. 665, is particularly important. It leads

---

[16]The relatively simple structures of HPSG are not an automatic consequence of its declarative nature. Postal's Metagraph Grammar framework (formerly known Arc Pair Grammar) is a declarative framework with structures that are similar in complexity to those of Minimalism (see Postal 2010).

[17]For interesting discussion of the historical development of the ideas that characterize Minimalism, see Culicover & Jackendoff (2005: Chapters 2 and 3).

Draft of 26th December 2018, 19:51

to a variety of abstract representations and movement processes. Consider, for example, the following:

(17)   a.   Who did Lee see?
       b.   Lee saw who

Who bears the same thematic relation to the verb *see* in (17a) as in (17b). Assuming UTAH, it follows that *who* in (17a) should be introduced in the object position which it occupies in (17b) and then be moved to its superficial position. Consider next the following:

(18)   a.   Lee was seen by Kim.
       b.   Kim saw Lee.

Here, *Lee* bears the same thematic relation to the verb *see* in (18a) as in (18b). Hence, it follows that *Lee* in (18a) should be introduced in the object position which it occupies in (18b) and then be moved to its superficial subject position. Finally, consider these examples:

(19)   a.   Lee seems to be ill.
       b.   It seems that Lee is ill.

Here, *Lee* bears the same thematic relation to *ill* in (19a) as in (19b). Thus, it follows that *Lee* in (19a) should be introduced in the same position as *Lee* in (19b). The standard Minimalist approach assumes that *Lee* in both examples originates in a position adjacent to *ill* and is moved a short distance in (19b) but a longer distance in (19a).

These analyses are more or less inevitable if one accepts UTAH. But how sound is UTAH? Work in HPSG shows that it is quite possible to capture both the syntactic and the semantic properties of these sentence types without the assumption that the crucial constituents occupy more than one position. Thus, there is no reason to accept UTAH.

4.1.2 Complex derived structures a la Generative Semantics

The idea that semantics is a simple reflection of syntax has led to other kinds of complexity. For example, it has led to revival of the idea once characteristic of Generative Semantics that lexical items may derive from complex expressions which in some sense represent their meanings.[18] Thus, Hale & Keyser (1993)

---

[18]For typical Generative Semantics proposals of this kind, see McCawley (1968) and Postal (1970). Like Minimalism, Generative Semantics was characterized by extremely complex syntactic structures and for similar reasons. See Newmeyer (1986: Chapter 4) for discussion.

argue that (20a) derives from a structure like that of (20b).

(20) a. Kim shelved the books.
    b. Kim put the books on the shelf.

One problem with this proposal is that *shelve X* means more than just *put X on the shelf.* Thus, (21a) is not equivalent to (21b).

(21) a. Kim put his elbow on the shelf.
    b. Kim shelved his elbow.

Moreover, as Culicover & Jackendoff (2005: 54–55) point out and as Hale & Keyser (1993: 105, Fn. 7) note themselves, denominal verbs can have many different interpretations.[19]

(22) a. Kim saddled the horse.
        (Kim put the saddle on the horse.)
    b. He microwaved the food.
        (He put the food in the microwave and in addition he heated it.)
    c. Lee chaired the meeting.
        (Lee was the chairperson of the meeting.)
    d. Sandy skinned the rabbit.
        (Sandy removed the skin from the rabbit.)
    e. Kim pictured the scene.
        (Kim constructed a mental picture of the scene.)
    f. They stoned the criminal.
        (They threw stones at the criminal.)
    g. He fathered three children.
        (He was the biological father of three children.)
    h. He mothers his students.
        (He treats his students the way a mother would.)

Denominal verbs need to be associated with the correct meanings, but there is no reason to think that syntax has a role in this.[20]

---

[19]The examples in (22c), (22g) and (22h) are taken from (Culicover & Jackendoff 2005: 54–55) or parallel to examples they discussed.

[20]See Culicover & Jackendoff (2005: 53–56) for further discussion.

Draft of 26th December 2018, 19:51

4.1.3 Complex structures and morphology

The idea that morphology is a simple reflection of syntax also leads to syntactic complexity. The fact that verbs in English and many other languages are marked for tense leads to the assumption that there is a T(ense) head at the heart of clause structure. Thus the sentence in (23) has the analysis in Figure 7.

(23) The cat chased the dog.

Figure 7: TP/VP analysis of simple English sentences

The verbal stem moves to the T head to pick up the *-ed* suffix.

Similarly the fact that nouns in English and other languages are marked for number leads to the assumption that there is a Num(ber) head at the heart of noun phrase structure. These elements are not solely motivated by morphology. The assumption that verbs move to T and nouns to Num in some languages but not others provides a way of accounting for cross-linguistic word order differences. However, as discussed in Crysmann (2018), Chapter 23 of this volume, HPSG assumes a realizational approach to morphology, in which affixes are just bits of phonology. Hence, analyses like these are out of the question.

4.1.4 Binary branching

Another source of complexity which also predates Minimalism is the assumption that all structures are binary branching. As Culicover & Jackendoff (2005: 112–116) note, this idea goes back to the 1980s. It entails that there can be no structures of the form in Figure 8a. Rather all structure must take the form in

Figure 8: Flat and binary branching

Figure 8b or Figure 8c. As Culicover & Jackendoff discuss, the arguments for the binary branching restriction have never been very persuasive. Moreover, it is incompatible with various analyses which have been widely accepted in HPSG and other frameworks. We will return to this topic in Section 4.2.

4.1.5 Unbounded dependencies

As noted in Section 3, the simplicity of the Minimalist grammatical system means the properties of constructions must largely derive from the lexical items that they contain. Hence, the properties of lexical items are absolutely central to Minimalism and often this means the properties of phonologically empty items, especially empty functional heads. Thus, such elements are central feature of Minimalist syntactic structures. These elements do much the same work as phrase types and the associated constraints in HPSG.

The contrast between the two frameworks can be illustrated with unbounded dependency constructions. Detailed HPSG analyses of various unbounded dependency constructions are set out in Sag (1997; 2010) and Ginzburg & Sag (2000), involving a complex system of phrase types. For Minimalism, unbounded dependency constructions are headed by a phonologically empty complementizer (C) and have either an overt filler constituent or an invisible filler (an empty operator) in their specifier position. Essentially, then, they have the structure in Figure 9. All the properties of the construction must stem from the properties of the C that heads it.

An important unbounded dependency construction is relative clauses. In English there are *wh*-relatives and non-*wh*-relatives and finite and non-finite relatives. *Wh*-relatives are illustrated by the following:

(24)   a.   someone [who you can rely on]

CP
XP   C′
C   TP

Figure 9: CP structures in Minimalism

     b.   someone [on whom you can rely]

(25)   a.   * someone [who to rely on]
     b.   someone [on whom to rely]

These show that whereas finite *wh*-relatives allow either an NP or a PP as the filler, non-finite *wh*-relatives only allow a PP. In the HPSG analysis of Sag (1997), the facts are a consequence of constraints on two phrase types. A constraint on the type *fin-wh-fill-rel-cl* allows the first daughter to be an NP or a PP while a constraint on *inf-wh-fill-rel-cl* requires the first daughter to be a PP. For Minimalism, the facts must be attributed to the properties of the complementizer. There must be a complementizer which takes a finite TP complement and allows either an NP or a PP as its specifier and another complementizer which takes a non-finite TP complement (with an unexpressed subject) and only allows a PP as its specifier.

     Non-*wh*-relatives require further phrase types within HPSG and further complementizers in Minimalism. However, rather than consider this, we will look at another unbounded dependency construction: *wh*-interrogatives. The basic data that needs to be accounted for is illustrated by the following:

(26)   a.   Who knows?
     b.   I wonder [who knows].
     c.   Who did Kim talk to?
     d.   I wonder [who Kim talked to].
     e.   I wonder [who to talk to].

Like *wh*-relatives, *wh*-interrogatives can be finite and non-finite. When they are finite their form depends on whether the *wh*-phrase is subject of the highest verb or something else. When it is subject of the highest verb, it is followed by what looks like a VP although it may be a clause with a gap in subject position. When

the *wh*-phrase is something else, the following clause shows auxiliary-initial order if it is a main clause and subject-initial order if it is not. Non-finite *wh*-interrogatives are a simple matter, especially as the filler does not have to be restricted in the way that it does in non-finite *wh*-relatives. Ginzburg & Sag (2000) present an analysis which has two types for finite *wh*-interrogatives, one for subject-*wh*-interrogatives such as those in (26a) and (26b), and another for non-subject-*wh*-interrogatives such as those in (26c) and (26d). The latter is subject to a constraint requiring it to have the same value for the features IC (INDEPENDENT-CLAUSE) and INV (INVERTED). Main clauses are [IC +] and auxiliary-initial clauses are [INV +]. Hence the constraint ensures that a non-subject-*wh*-interrogative shows auxiliary-initial order just in case it is a main clause.

How can the facts be handled within Minimalism? As noted above, Minimalism analyses auxiliary-initial order as a result of movement of the auxiliary to C. It is triggered by some feature of C. Thus C must have this feature just in case (a) it heads a main clause and (b) the *wh*-phrase in its specifier position is not the subject of the highest verb. There are no doubt various ways in which this might be achieved, but the key point is the properties of a phonologically empty complementizer are crucial.

Borsley (2006b; 2017) discusses Minimalist analyses of relative clauses and *wh*-interrogatives and suggests that at least eight complementizers are necessary. One is optionally realized as *that*, and another is obligatorily realized as *for*. The other six are always phonologically empty. But it has been clear since Ross (1967) and Chomsky (1977) that relative clauses and *wh*-interrogatives are not the only unbounded dependency constructions. Here are some others:

(27) a. What a fool he is!   (*wh*-exclamative clause)
      b. The bagels, I like.   (topicalized clause)
      c. Kim is more intelligent [than Lee is].   (comparative-clause)
      d. Kim is hard [to talk to].   (*tough*-complement-clause)
      e. Lee is too important [to talk to].   (*too*-complement-clause)
      f. [The more people I met], [the happier I became].   (*the*-clauses)

Each of these constructions will require at least one empty complementizer. Thus, a comprehensive account of unbounded dependency constructions will require a large number of such elements. But with a large unstructured set of complementizers there can be no distinction between properties shared by some or all elements and properties restricted to a single element. There are a variety of shared properties. Many of the complementizers will take a finite complement, many others will take a non-finite complement, and some will take both. There

will also be complementizers which take the same set of specifiers. Most will not attract an auxiliary, but some will, not only the compementizer in an example like (26c) but also the complementizers in the following, where the auxiliary is in italics:

(28) a. Only in Colchester *could* such a thing happen.

  b. Kim is in Colchester, and so *is* Lee.

  c. Such *is* life.

  d. The more Bill smokes, the more *does* Susan hate him.

Thus, there are generalizations to be captured here. The obvious way to capture them is with the approach developed in the 1980s in HPSG work on the hierarchical lexicon (Flickinger, Pollard & Wasow 1985; Flickinger 1987), i.e. a detailed classification of complementizers which allows properties to be associated not just with individual complementizers but also with classes of complementizers. With this it should be possible for Minimalism not just to get the facts right but to capture the full set of generalizations. In many ways such an analysis would be mimicking the HPSG approach with its hierarchy of phrase types.[21] But in the present context the main point is the simplicity of the Minimalist grammatical system is another factor which leads to more complex syntactic structures than those of HPSG.

4.1.6 Syntactification of semantic categories

Further complexity arises from the syntactification of semantic properties (Rizzi 2014). This is especially apparent in a sub-school that calls itself "cartographic". Generative Grammar in the Chomskyan tradition comes with strong claims about the autonomy of syntax. There is a syntactic component and than there are the components of Phonological Form (PF) and Logical Form (LF), in more recent versions of the theory this is the articulatory-perceptual system (AP) and the conceptual-intentional system (CI). Figure 6 shows the early Minimalist architecture and the architecture assumed in the Phase-based models. Syntax was always regarded primary and PF and LF derived from syntactic representations. This is similar in Minimalism. The problem is that questions of intonation are connected to semantic and information structural properties (Halliday 1970: 36). A way around this is to stipulate syntactic features that can be interpreted by both PF and LF (Gussenhoven 1983). Another way of dealing with the data is to employ empty elements that are responsible for certain ordering of elements and

---

[21]For a fuller discussion of the issues see Borsley (2006b; 2017).

that can be interpreted in the semantics. The accounts of Rizzi and Cinque are very prominent in this school of thought. For example, Rizzi (1997) suggests an analysis of the left periphery of clauses that incorporate special functional projections for topic and focus. His analysis is shown in Figure 10. In comparison no
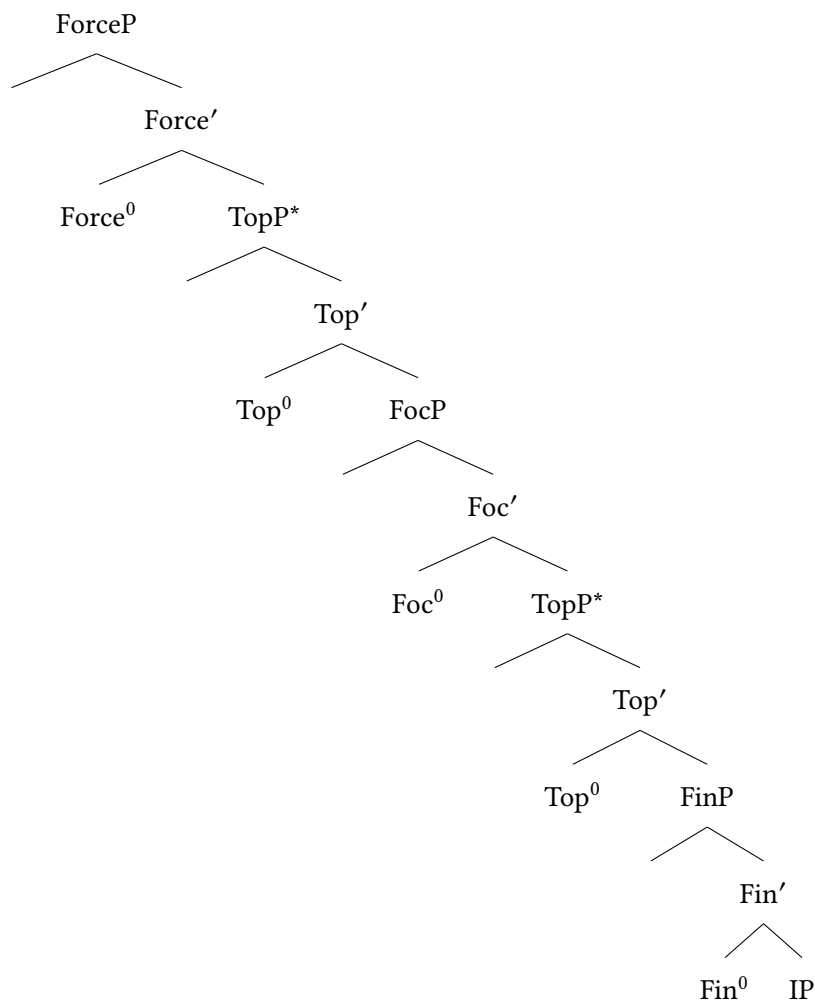

Figure 10: Syntactic structure of sentences following Rizzi (1997: 297)

such projections exist in HPSG theories. HPSG grammars are surface oriented and the syntactic labels correspond to classical part of speech categorizations. So in examples with frontings like (29) the whole object is a verbal projection and

not a Topic phrase, a Focus Phrase or a Force phrase.

(29)   Bagels, I like.

Of course the fronted elements may be topics or foci but this is a property that is represented independently of syntactic information in parts of feature descriptions having to do with information structure. For treatment of information structure in HPSG see Engdahl & Vallduví (1996), De Kuthy (2000) and also de Kuthy (2018), Chapter 25 of this volume. On determination of clause types see Ginzburg & Sag (2000) and Müller (2016b). For general discussion of the representation of information usually assigned to different linguistic "modules" and on "interfaces" between them in theories like LFG and HPSG see Kuhn (2007).

Cartographic approaches also assume a hierarchy of functional projections for the placement of adverbials. Some authors assume that all sentences in all languages have the same structure, which is supposed to explain orders of adverbials that seem to hold universally (e.g., Cinque 1999: 106 and Cinque & Rizzi 2010: 54–55). A functional projection selects for another functional projection to establish this hierarchy of functional projections and the respective adverbial phrases can be placed in the specifier of the corresponding functional projection. Cinque (1999: 106) assumes 32 functional projections in the verbal domain. Cinque & Rizzi (2010: 57, 65) assume at least four hundred functional heads, which are – according to them – all part of a genetically determined UG.

In comparison, HPSG analyses assume that verbs project: a verb that is combined with an argument is a verbal projection. If an adverb attaches, a verbal projection with the same valence but augmented semantics results. Figure 11 shows the Cartographic and the HPSG structures. While the adverbs (Adv_1 and Adv_2 in the figure) attach to verbal projections in the HPSG analysis (S and VP are abbreviations standing for verbal projections with different valence requirements), the Cartographic approach assumes empty heads that select a clausal projection and provide a specifier position in which the adverbs can be realized. For the shake of exposition we called these heads $FAdv_1$ and $FAdv_2$. For example, $FAdv_2$ can combine with the VP and licences an Adv_2 in its specifier position. As is clear from the figure, the Cartographic approach is more complex since it involves two additional categories ($FAdv_1$ and $FAdv_2$) and nine nodes for the adverbial combination rather than five.

An interesting difference is that verbal properties are projected in the HPSG analysis. By doing this it is clear whether a VP contains an infinitive or a participle.
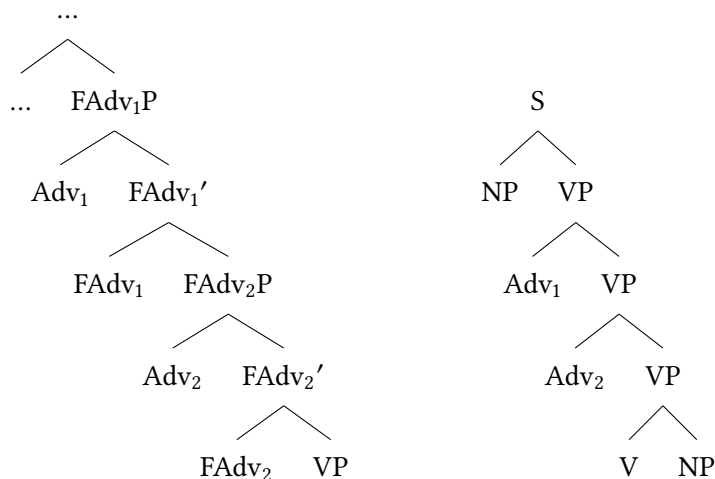
(30)   a.   Kim has met Sandy.

```
                    …
                  ╱   ╲
        …      FAdv₁P                          S
             ╱    ╲                          ╱  ╲
          Adv₁   FAdv₁′                    NP   VP
                ╱    ╲                         ╱  ╲
            FAdv₁   FAdv₂P                  Adv₁   VP
                   ╱    ╲                        ╱  ╲
               Adv₂   FAdv₂′                  Adv₂   VP
                     ╱    ╲                         ╱  ╲
                 FAdv₂   VP                        V    NP
```

Figure 11: Treatment of adverbial phrases in Cartographic approaches and in HPSG

b. Kim will meet Sandy.

This property is important for the selection by a superordinate head, the auxiliary in the examples in (30). In a Cartographic approach one either has to assume that adverbial projections have features correlated with verbal morphology or one has to assume that superordinate heads may check properties of linguistic items that are deeply embedded.

> Bob: I am not sure whether it is worth it to have the last paragraph …
> Stefan: For me this is one of the absurdities of the Cartographic approach. The category labels are distribution classes. For Cartographic approaches we have an Aspect Phrase but from the outside we are looking for a VP with certain properties. Of course Cartography works if you always have the full battery of functional projections. But then you have to make sure that it is of the right kind, that is the verbal form deep inside has to be right.
> Bob: I now think this is quite important. I assume they think that the full battery of functional projections is always present, but that doesn't alter the fact that you need not just any kind if verb but a verb of the right form to satisfy the requirements of an auxiliary.
> However, they have got a way of getting adverbs in the right order and I'm not really sure how we would do that.

add "a complex"

If one believed in Universal Grammar (which researchers working in HPSG usually do not) and in innately specified constraints on adverb order one would not assume that all languages contain the same structures, some of these struc-

tures being invisible. Rather one would assume linearization constraints (see Müller (2018: Section 2), Chapter 10 of this volume) to hold crosslinguistically. If adverbs of a certain type do not exist in a language, the linearization constraints would not do any harm. They just would never apply since there is nothing to apply to (Müller 2015b: 46).

4.1.7 Summary

Having discussed uniformity in theta role assignment, Generative Semantics-like approaches, branching, nonlocal dependencies and Cartographic approaches to the left periphery and adverb order within clauses, we conclude that a variety of features of Minimalism lead to structures that are much more complex than those of HPSG. HPSG shows that this complexity is unnecessary given a somewhat richer conception of grammar.

4.2 The simplicity of Minimalist structures

As we emphasized above, while minimalist structures are very complex, they are also simple in the sense that they have just a single ingredient, local trees consisting of a head and a single non-head. Most outsiders agree that this is too simple.

4.2.1 Binary branching, VPs, and verb-initial clauses

We look first at binary branching.[22] As we noted above, the assumption that all branching is binary is incompatible with various analyses which have been widely accepted in HPSG and other frameworks. For example, it means that the bracketed VP in (31), which contains two complements, cannot have the ternary branching structure in Figure 12, which is suggested in Pollard & Sag (1994: 36) and much other work.

---

[22]In addition to structures with two or more branches, HPSG uses unary branching structures both in syntax and in the lexicon (lexical rules basically are unary branching structures, see Meurers (2001) and Davis & Koenig (2018: Section 5), Chapter 4 of this volume. For example, unary branching syntactic rules are used for semantic type shifting (Partee 1987). For respective HPSG analyses see Flickinger (2008: 91–92) and Müller (2009; 2012). The lack of unary branching structures in Minimalism is no problem since empty heads can be used instead. The empty head projects the properties that would be otherwise assigned to the mother node of the unary projection. See for example Ramchand (2005: 370). So, while the effects of unary projections can be modeled, the resulting structures are more complex. For a general discussion of empty elements and unary projections and lexical rules see Müller (2016a: Sections 19.2 and 19.5).

(31) Kim [gave a book to Lee].

VP

V   NP   NP

gave   Lee   a book

Figure 12: Flat structure for the VP *gave Lee a book*

Instead it has been assumed since Larson (1988) that the VP in examples like (31) has something like the structure in Figure 13. It is assumed that the verb origi-
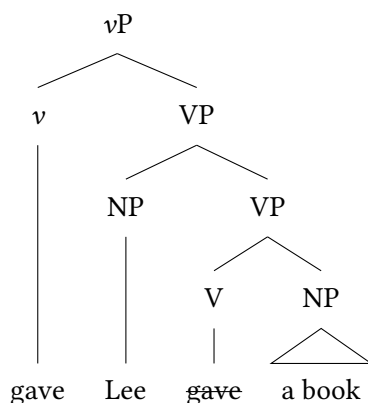
*v*P

v   VP

NP   VP

V   NP

gave   Lee   ~~gave~~   a book

Figure 13: Larson-type analysis of VPs involving little *v*

nates in the lower VP and is moved into the higher VP. The higher V position to which the verb moves is commonly labelled *v* ("little *v*") and the higher phrase *v*P. The main argument for such an analysis appears to involve anaphora, especially contrasts like the following:

(32) a. John showed Mary herself in the picture.
      b. * John showed herself Mary in the picture.

The first complement can be the antecedent of a reflexive which is the second complement, but the reverse is not possible. If constraints on anaphora refer to constituent structure as suggested by Chomsky (1981), the contrast suggests that

the second NP should be lower in the structure than the first NP. But, as suggested by Pollard & Sag (1992), it is assumed in HPSG that constraints on anaphora refer not to constituent structure but to a list containing all arguments in order of obliquenss, in recent versions of HPSG the ARG-ST list (see also Branco (2018), Chapter **??** of this volume). On this view, anaphora can provide no argument for the complex structure in Figure 13. Therefore, both flat structures and binary branching structures with different branching directions as in Figure 14 are a viable option in HPSG. Müller (2015a: Section 2.4; 2018a) argues for such binary
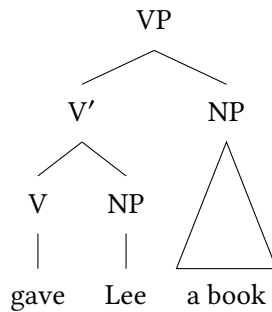

Figure 14: Possible analysis of VPs in HPSG with a branching direction differing from Larson-type structures

branching structures as a result of parametrizing the Head-Complement Schema for various variants of constituent order (head-initial and head-final languages with fixed constituent order and languages like German and Japanese with freer constituent order).

The fact that Merge combines two expressions also means that the auxiliary-initial clause in (33) cannot have a flat structure with both subjects and complement(s) as sisters of the verb, as in Figure 15.

(33)   Will Kim be here?

It is standardly assumed in Minimalism that it has a structure of the form in Figure 16 or more complicated structures, as explained in Section 4.1.6. *Will* is analysed as a T(ense) element which moves to the C(omplementizer) position. An analysis like this is essentially the only possibility within Minimalism.

It is not just English auxiliary-initial clauses that cannot have a ternary branching analysis within Minimalism but verb-initial clauses in any language. A notable example is Welsh, which has verb-initial order in all types of finite clause.

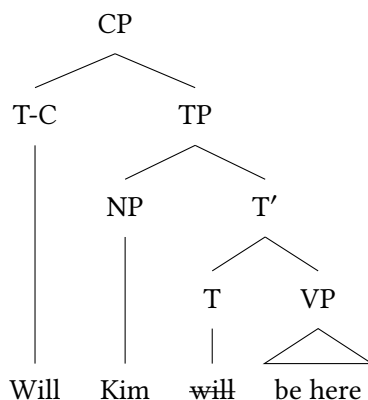Figure 15: Flat structure for *Will Kim be there?*

Figure 16: CP/TP structure for *Will Kim be there?*

Here are some relevant examples:[23]

(34)   a. Mi/Fe gerddith   Emrys i ’r dre.   (Welsh)
          PRT   walk.FUT.3SG Emrys to the town
          ‘Emrys will walk to the town.’

      b. Dywedodd   Megan [cerddith   Emrys i ’r dre].
          say.PAST.3SG Megan   walk.FUT.3SG Emrys to the town
          ‘Megan said Emrys will walk to the town.’

A variety of transformational work, including work in Minimalism, has argued
for an analysis like Figure 16 for Welsh finite clauses (see e.g., Jones & Thomas
1977, Sproat 1985, Sadler 1988, Rouveret 1994, and Roberts 2005). But Borsley

---

[23]Positive main clause verbs are optionally preceded by a particle (*mi* or *fe*). We have included
this in (34a) but not in (34b). When it appears it triggers so-called soft mutation. Hence (34a)
has *gerddith* rather than the basic form *cerddith*, which is seen in (34b).

Draft of 26th December 2018, 19:51

(2006a) argues that there is no theory-neutral evidence for a structure of this kind. Hence, at least for Welsh, it seems that a simpler flat structure like Figure 15 is preferable.[24] Note, that we do not argue against structures like the one in Figure 16 categorically. The analog to head-movement analyses is standard among HPSG grammarians of German and there is data from apparent multiple frontings that seems to make a head-movement analysis unavoidable. See Müller (2017) for a book-length discussion of German clause structure. Müller (2018: Section 4.1), Chapter 10 of this volume also discusses head-movement in HPSG.

4.2.2 Headedness and coordination

We turn now to the idea that all structures are headed. For HPSG, and many other approaches, there are headed structures and non-headed structure. Probably the most important example of the latter are coordinate structures such as those in (35) (see Sag 2003 for an HPSG analysis).

(35)   that [Kim and Lee] [wrote poems and painted pictures].

Much work in Minimalism assumes that coordinate structures are headed by the conjunction (Larson 1990: 596; Radford 1993: 89; Kayne 1994: Chapter 6; Johannessen 1998; Van Koppen 2005: 8; Bošković 2009: 474; Citko 2011: 27).[25] This suggests that both coordinate structures in (35) are conjunction phrases. This is highly problematic since the category of the phrases plays a role in accounting for their external distribution. So the VPs *wrote poems* and *painted pictures* have to be combined with a DP/NP to form a complete sentence. But according to the ConjP theory *Kim and Lee* is not a DP or NP it is a ConjP and hence incompatible with any requirements. Similarly, a T head in the analysis of (35) requires a VP argument but instead of a VP *wrote poems and painted pictures* there is only a ConjP.[26] It is fairly clear that conjunctions cannot be ordinary heads. Johannessen (1996: 669) suggests an analysis in which the specifier projects its features.

---

[24]Borsley (2016) argues for a similar flat structure for the Caucasian ergative SOV language Archi.

[25]Kayne (1994: 57) differs from other proposals in not assuming the category Conj for the conjunction. Instead, he uses X^0 as the category in his structured examples. Since X is an underspecified variable his theory is underdetermined: while a ConjP is not compatible with any requirement by a governing head, an XP could appear as an argument of any dominating head. Kayne needs to work out a theory that determines the properties of the projected XP in relation to the coordinated items. We discuss this below.

[26]If one considers the part of speech labels only, one would expect the two ConjPs to be interchangeable, but of course they are not:

     (i)   * [Wrote poems and painted pictures] [Kim and Lee].

Of course the two ConjPs are not exactly of the same category since there may be further

---

She depicts the analysis as in Figure 17. The problem is that it is unclear how this

```
        CoP[X]
        /\
      X    Co′
            /\
          Co    Y
```

Figure 17: Analysis of coordination with projection of features from the first conjunct according to Johannessen (1996: 669)

should be formalized: either the head category of the complete object is ConjP or it is X. Governing heads have to know where to look for the category. If they look at X, why is the part of speech information of Co projected? Why would governing heads not look at the category of other specifiers rather than their heads? Johannessen's analysis seems to be equivalent to an HPSG analysis assuming two different head features, which would be highly problematic for the reasons just given. Furthermore, coordinations are not equivalent to the first conjunct. There are cases where the coordination is a sum of the parts. For example, *Kim and Sandy* is a plural NP, as the agreement with the verb shows:

(36)   Kim and Sandy laugh.

Johannessen's analysis seems to predict that the coordination of *Kim* and *Sandy* behaves like *Kim*, which is not the case. So, if one wants to assume an analysis with the conjunction as a head, one would have to assume that the head is a functor taking into account the properties of its specifier and complement, and projecting nominal information if they are nominal, verbal if they are verbal, etc (Steedman 1991). This would make them a unique type of a head with a unique relation to their specifier and complement. A problem for this approach is coordinate structures in which the conjuncts belong to different categories, e.g., the following:

(37)   a. Hobbs is [a linguist and proud of it].
        b. Hobbs is [angry and in pain].

---

features that distinguish the two ConjPs. But how these features are distributed between the conjuncts and the mother is not worked out.

For a more detailed critique of the ConjP approach see Borsley (2005).

Draft of 26th December 2018, 19:51

Such examples have led to HPSG analyses in which coordinate structures have whatever properties are common to the two conjuncts (Sag 2003). Within Minimalism, one might try to mimic such analyses by proposing that conjunctions have whatever properties are common to their specifier and complement. But a problem arises with an example like (38), where the conjuncts are not phrases but words.

(38) Kim [criticized and insulted] his boss.

To accommodate such examples, conjunctions would have to acquire not only part of speech information from the conjuncts but also selectional information. They would be heads which combine with a specifier and a complement to form an expression which, like a typical head, combines with a specifier and a complement. This would be a very strange situation and in fact it would make wrong predictions since the object *his boss* would be the third-merged item. It would hence be "later-merged" in the sense of Chomsky (2008: 146) and therefore treated as a specifier rather than a complement.[27]

4.2.3 Binary branching and headless structures: The NPN construction

Another problem for Minimalist theories is the NPN Construction discussed by Matsuyama (2004) and Jackendoff (2008). Examples are provided in (39):

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

As Jackendoff argued it is not possible to identify one of the elements in the construction as the head. The construction has several peculiar properties and we share Jackendoff's view that these constructions are best treated by a phrasal configuration in which these highly idiosyncratic properties are handled. The construction is discussed in more detail in Müller (2018), Chapter 37 of this volume and Bargmann's analysis within HPSG is provided. Bargmann's analysis also captures multiple repetitions of the PN sequence as they occur in (39b). Until now there is one proposal for NPN in the Minimalist framework: G. Müller (2011). G. Müller develops a reduplication account. He states that reduplication applies to words only and claims that German differs from English in not allowing adjective noun sequences in NPN constructions. He is aware of the possibility of

---

[27]There have been attempts to argue that conjuncts are always phrases (Kayne 1994, Bruening 2018). But this position seems untenable (Abeillé 2006, Müller 2018c: Section 7).

these constructions in English (*miserable day after miserable day*) and states that his analysis is intended to account for the German data only. While this alone is a serious shortcoming of the analysis, the empirical claim does not hold water either as the following example from Müller (2018), Chapter 37 of this volume shows:

(40)   Die beiden tauchten nämlich geradewegs   wieder aus   dem heimischen
       the two   surfaced namely straightaway again   from the home
       Legoland auf, wo   sie   im   Wohnzimmer, schwarzen Stein um
       Legoland PART where they in.the living.room   black   brick after
       schwarzen Stein, vermeintliche Schusswaffen nachgebaut hatten.[28]
       black   brick alledged   firearms   recreated   had
       'The two surfaced straightaway from their home Legoland where they
       had recreated alledged firearms black brick after black brick.'

Apart from failing on the reduplication of adjective-noun combinations like *schwarzen Stein* 'black brick', the reduplication approach also fails on NPN patterns with several PN repetitions as in (39b): if the preposition is responsible for reduplicating content it is unclear how the first *after* is supposed to combine with *day* and *day after day*. We are convinced that it is possible to design analyses of the NPN construction involving several empty heads but it is clear that these solutions would come at a rather high price.

4.2.4 Movement for more local phenomena like scrambling, passive and raising

We want now to consider the dependencies that Minimalism analyzes in terms of Move/Internal Merge. In the next section we look at unbounded dependencies, but first we consider local dependencies in passives, unaccusatives, rasing sentences, and scrambling. The following illustrate the first three of these:

(41)   a. Kim has been hit.
       b. Kim has disappeared.
       c. Kim seems to be clever.

These differ from unbounded dependency constructions in that whereas the gaps in the latter are positions in which overt NPs can appear, this is not true of the supposed gap positions in (41):

---

[28]taz, 05.09.2018, p. 20, quoted from Müller (2018: 746).

(42) a. * It has been hit Kim.

  b. * It has disappeared Kim.

  c. * It seems Kim to be clever.

This is a complication if they involve the same mechanism, but is unsurprising if they involve different mechanisms, as in HPSG and most other frameworks.

4.2.4.1 Passive

In the classical analysis of the passive in MGG, it is assumed that the morphology of the participle suppresses the agent role and removes the ability to assign accusative case. In order to receive case the underlying object has to move to the subject position to Spec,TP where it gets nominative (Chomsky 1981: 124).

(43) a. The mother gave [the girl] [a cookie].

  b. [The girl] was given [a cookie] (by the mother).

The analysis assumed in recent Minimalist work differs in detail but is movement-based as its predecessors. While movement-based approaches seem to work well for SVO languages like English, they are problematic for SOV languages like German. To see why consider the examples in (44):

(44) a. weil   das   Mädchen dem   Jungen den   Ball schenkte
      because the.NOM girl   the.DAT boy   the.ACC ball gave
      'because the girl gave the ball to the boy'

  b. weil   dem   Jungen der   Ball geschenkt wurde
      because the.DAT boy   the.NOM ball given   was
      'because the ball was given to the boy'

  c. weil   der   Ball dem   Jungen geschenkt wurde
      because the.NOM ball the.DAT boy   given   was

In comparison to (44c), (44b) is the unmarked order (Höhle 1982). *der Ball* 'the ball' in (44b) occurs in the same position as *den Ball* in (44a), that is, no movement is necessary. Only the case differs. (44c) is, however, somewhat marked in comparison to (44b). So, if one assumed (44c) to be the normal order for passives and (44b) is derived from this by movement of *dem Jungen* 'the boy', (44b) should be more marked than (44c), contrary to the facts. To solve this problem, an analysis involving abstract movement has been proposed for cases such as (44b): the elements stay in their positions, but are connected to the subject position and receive their case information from there. Grewendorf (1993: 1311) assumes that

there is an empty expletive pronoun in the subject position of sentences such as (44b) as well as in the subject position of sentences with an impersonal passive such as (45):[29]

(45) weil   heute nicht gearbeitet wird
      because today not   worked   is
      'because there will be no work done today'

A silent expletive pronoun is something that one cannot see or hear and that does not carry any meaning. Such entities are not learnable from input and hence innate domain specific knowledge would be required and of course, approaches that do not have to assume very specific innate knowledge are preferable. For further discussion of language acquisition see Section 5.2 and Ginzburg (2018), Chapter 27 of this volume.

HPSG does not have this problem since passive is treated by lexical rules that map verbal stems onto participle forms with a reduced argument structure list. The first element (the subject in the active) is suppressed so that the second element (if there is any) becomes first. In SVO languages like English and Icelandic this element is realized before the verb: there is a valence feature for subjects/specifiers and items that are realized with the respective schema are serialized to the left of the verb. In SOV languages like German and Dutch the subject is treated like other arguments and hence it is not put in a designated position before the finite verb. No movement is involved in this valence-based analysis of the passive. The problem of MGG analyses is that they mix two phenomena: passive and subject requirement. Since these two phenomena are kept separate in HPSG, problems like that discussed above can be avoided. See Müller (2016a: Section 3.4, Chapter 20) for further discussion.

Stefan: Actually the chapter about order does not deal with SVO vs. SOV and SPR vs. COMPS. Maybe it should.

4.2.4.2 Scrambling

Discussing passive, we already touched on problems related to local reordering of arguments, so-called *scrambling*. In what follows, we want to discuss scrambling in more detail. Languages like German have a more liberal constituent order than English. A sentence with a ditransitive verb allows for six permutations of the arguments, two of which are given in (46):

---

[29]See Koster (1986: 11–12) for a parallel analysis for Dutch as well as Lohnstein (2014) for a movement-based account of the passive that also involves an empty expletive for the analysis of the impersonal passive.

Draft of 26th December 2018, 19:51

(46)   a.   [weil]   der   Mann der   Frau   das   Buch gibt
            because the.NOM man   the.DAT woman the.ACC book gives

       'because the man gives the book to the woman'

       b.   [weil]   das   Buch der   Mann der   Frau   gibt
            because the.ACC book the.NOM man   the.DAT woman gives

It was long argued that scrambling should be handled as movement as well (Frey 1993). An argument that has often been used to support the movement-based

Figure 18: Analysis of local reordering as movement to Spec TP and "base-generation" analysis assumed in HPSG

analysis is the fact that scope ambiguities exist in sentences with reorderings which are not present in sentences in the base order. The explanation of such ambiguities comes from the assumption that the scope of quantifiers can be derived from their position in the superficial structure as well as their position in the underlying structure. If the position in both the surface and deep structure are the same, that is, when there has not been any movement, then there is only one reading possible. If movement has taken place, however, then there are two possible readings (Frey 1993:):

(47)   a.   Es ist nicht der Fall, daß er mindestens einem Verleger   fast
            it is not   the case that he at.least   one   publisher almost

jedes Gedicht anbot.
every poem offered

'It is not the case that he offered at least one publisher almost every poem.'

b. Es ist nicht der Fall, daß er fast jedes Gedicht$_i$ mindestens einem
it is not the case that he almost every poem at.least one

Verleger $_{-i}$ anbot.
publisher offered

'It is not the case that he offered almost every poem to at least one publisher.'

It turns out that approaches assuming traces run into problems as they predict certain readings for sentences with multiple traces which do not exist (see Kiss 2001: 146 and Fanselow 2001: Section 2.6). For instance in an example such as (48), it should be possible to interpret *mindestens einem Verleger* 'at least one publisher' at the position of $_{-i}$, which would lead to a reading where *fast jedes Gedicht* 'almost every poem' has scope over *mindestens einem Verleger* 'at least one publisher'. However, this reading does not exist.

(48) Ich glaube, dass mindestens einem Verleger$_i$ fast jedes Gedicht$_j$ nur
I believe that at.least one publisher almost every poem only

dieser Dichter $_{-i}$ $_{-j}$ angeboten hat.
this poet offered has

'I think that only this poet offered almost every poem to at least one publisher.'

The alternative to movement-based approaches are so-called "base-generation" approaches in which the respective orders are derived directly. Fanselow (2001), working within the Minimalist Program, suggests such an analysis in which arguments can be combined with their heads in any order. This is the HPSG analysis that was suggested by Gunji (1986) for Japanese and is standardly used in HPSG grammars of German (Hinrichs & Nakazawa 1994; Kiss 1995; Meurers 1999; Müller 2003; 2017). See also Müller (2018), Chapter 10 of this volume.

Sauerland & Elbourne (2002: 308) discuss analogous examples from Japanese, which they credit to Kazuko Yatsushiro. They develop an analysis where the first step is to move the accusative object in front of the subject. Then, the dative object is placed in front of that and then, in a third movement, the accusative is moved once more. The last movement can take place to construct either a structure that is later passed to LF or as a movement to construct the Phonological

Form. In the latter case, this movement will not have any semantic effects. While this analysis can predict the correct available readings, it does require a number of additional movement operations with intermediate steps.

4.2.5 Nonloacl dependencies

Having dealt with phenomena treated via Move/Internal Merge in Minimalism but involving more local phenomena, we now turn to genuine nonlocal dependencies and compare the Move/Internal Merge approach to the HPSG approach to nonlocal dependencies.

4.2.5.1 Gaps without filler

The Move/Internal Merge approach seems quite plausible for typical examples of an unbounded dependency, but issues arise with less typical examples. Within this approach one expects to see a clause-initial filler-constituent and a gap somewhere in the following clause. This is what we commonly find, but there are unbounded dependency constructions in which there is a gap but no visible higher constituent matching it. Consider e.g., the following:

(49) a. the book [Kim bought _]
      b. Lee is too important [for you to talk to _].
      c. Lee is important enough [for you to talk to _].
      d. Kim is easy [for anyone to talk to _].

Within Minimalist assumptions, it is more or less necessary to assume that such examples contain an invisible filler (a so-called empty operator). Unless there is some independent evidence for such invisible fillers, they are little more than an ad hoc device to maintain the Move/Internal Merge approach. Within the HPSG SLASH-based approach to unbounded dependencies, there is no assumption that there should always be a filler at the top of an unbounded dependency (Pollard & Sag 1994: Chapter 4, see also Borsley & Crysmann (2018), Chapter 14 of this volume). Hence, the examples in (49) are completely unproblematic.

4.2.5.2 Filler without gaps: Resumptive pronouns

There are also unbounded dependency constructions which seem to have not a gap but a resumptive pronoun (RP). Among many languages that are relevant here is Welsh, which has RPs in both *wh*-interrogatives and relative clauses, as the following illustrate:

(50)   a. Pa   ddyn werthodd   Ieuan y ceffyl iddo   *fo*?
               which man sell.PAST.3SG Ieuan the horse to.3SGM he

               'Which man did Ieuan sell the horse to?'

        b. y   dyn werthodd   Ieuan y ceffyl iddo   *fo*
               the man sell.PAST.3SG Ieuan the horse to.3SGM he

               'the man that Ieuan sold the horse to'

Willis (2011) and Borsley (2010; 2013) present evidence that Welsh RPs involve the same mechanism as gaps. Within Minimalism, this means that they must involve Move/Internal Merge. But one expects to see a gap where Move/Internal Merge has applied. One Minimalist response suggests that instead of being deleted, the copy left behind by Move/Internal Merge is somehow turned into a pronoun (see McCloskey 2006). A problem for this approach is that it makes it surprising that RPs universally look like ordinary pronouns (McCloskey 2002). Another approach exploits the complexity of Minimalist structures and proposes that there is a gap in the structure somewhere near the RP. Thus, for example, Willis (2011) proposes that examples like those in (50) with an RP in prepositional object position have a coindexed operator in the specifier position of PP, which undergoes movement. Similar approaches are outlined in Aoun et al. (2001) and Boeckx (2003). For detailed objections to both approaches, see Borsley (2013: Section 3). Within the SLASH-based approach of HPSG, there is no reason to think that there will always be a gap at the bottom of a dependency, and it is not difficult to accommodate RPs. See Vaillette (2001), Taghvaipour (2010), Borsley (2013) and Crysmann (2012; 2016) for slightly different approaches.[30] See also Borsley & Crysmann (2018), Chapter 14 of this volume for a more detailed discussion of nonlocal dependencies.

4.3 Conclusion

Thus, there are variety of phenomena which suggest that the Minimalist view of constituent structure is too simple. The restriction to binary branching, the

---

[30]Also relevant here are examples with more than one gap such as the following:

(i)   a. Who does Kim like _ and Lee hate _?

       b. Which book did you criticize _ without reading _?

There have been various attempts to accommodate such examples within the Move/Internal Merge approach, but it is not clear that any of them is satisfactory. In contrast such examples are expect within the SLASH-based approach (Levine & Sag 2003). See also (Pollard & Sag 1994: Section 4.6).

Draft of 26th December 2018, 19:51

assumption that all structures are headed, and Move/Internal Merge all seem problematic. It looks, then, as if the Minimalist view is both too complex and too simple.

5 Psycholinguistic issues

Although they differ in a variety of ways, HPSG and Minimalism agree that grammatical theory is concerned with linguistic knowledge. They focus first and foremost on the question: what form does linguistic knowledge take? But there are other questions that arise here, notably the following:

- How is linguistic knowledge put to use?

- How is linguistic knowledge acquired?

Both questions are central concerns for psycholinguistics. Thus, in considering the answers that HPSG and Minimalism can give we are considering their relevance to psycholinguistics. Chomskyan approaches, including Minimalism, have focused mainly on the second question and have paid little attention to the first. HPSG has had more to say about the first and has shown less interest in the second. However, there is a large body of work on acquisition in Construction Grammar and since HPSG is a constructionist theory (Müller 2018, Chapter 37 of this volume) all the insights carry over to HPSG. Clearly an adequate grammatical theory should be able to give satisfactory answers to both questions. In this section we will look briefly at the relation of the two theories to processing and then consider more fully their relation to acquisition.

5.1 Processing

We noted in Section 3 that whereas HPSG is a declarative or constraint-based approach to grammar, Minimalism has a procedural view of grammar. This contrast means that HPSG is much more suitable then Minimalism for incorporation into an account of the processes that are involved in linguistic performance.

The most obvious fact about linguistic performance is that it involves both production and comprehension. As noted in Section 3, this suggests that the knowledge that is used in production and comprehension should have a declarative character as in HPSG and not a procedural character as in Minimalism.

A second important feature of linguistic performance is that it involves different kinds of information utilized in any order that is necessary. Sag & Wasow (2011: 367–368) illustrate with the following examples:

(51) a. The sheep that was sleeping in the pen stood up.

b. The sheep in the pen had been sleeping and were about to wake up.

In (51a), morphological information determines the number of sheep before non-linguistic information determines that *pen* means 'fenced enclosure' and not 'writing implement'. In (51b), on the other hand, non-linguistic information determines that *pen* means 'fenced enclosure' before morphological information determines the number of sheep. This is unproblematic for an approach like HPSG in which linguistic and non-linguistic knowledge takes the form of constraints which are not ordered in any way.[31] It is quite unclear how the facts can be accommodated within Minimalism given that linguistic knowledge with its procedural form is quite different from non-linguistic knowledge.

Other features of HPSG also make it attractive from a processing point of view. Firstly, there is the fact emphasized earlier that linguistic expressions have a single relatively simple constituent structure with a minimum of phonologically empty elements. Secondly there is the fact that all constraints are purely local and never affect anything larger than a local tree consisting of an expression and its daughters. Both these properties make processing easier than it would otherwise be. Minimalism has neither property and hence again seems less satisfactory than HPSG in this area.

Furthermore, it must be said that all approaches based on Chomsky's Phase-based approaches (Chomsky 1999; 2008) are incompatible with psycholinguistic results.

> Bob: Are phase-based approaches any worse than earlier versions of Minimaliism? I would have thought not.
>
> Stefan: True. But how was it back then. I do not remember. Was everything bottom up? When did they send stuff to the interfaces?

Most of Minimalist work assumes that structures are build in a bottom-up fashion. Chomsky assumes that there are so-called Phases and when a Phase is complete the complement of the Phase head is "shipped" to the interfaces (pronounciation and interpretation, see Figure 6). For the example in (52) this would mean that *a liar* is constructed first to be combined with *is,* which is then combined with *Peter* and *Peter is a liar* is combined with *that.* Since *that* is the Phase head *Peter is a liar* is shipped to the interfaces. The *that* clause can then be integrated into more complex constituents like *that Sandy claimed that Peter is a liar* but *Peter is a liar* is a complete unit the internal structure of which is no longer visible.

---

[31]See also Lücking (2018), Chapter 32 of this volume on the interaction of gesture and speech.

(52)   Kim believes that Sandy claimed that Peter is a liar.

The problem with this approach is that it uses a rhetoric talking about interfaces and so on that makes the impression as if it was related to processing and to psycho-linguistics, but it is fundamentally against all knowledge we have about language processing: language processing is incremental and all channels are used in parallel (Marslen-Wilson 1975; Tanenhaus et al. 1995; 1996). Information about phonology, morpho-syntax, semantics, information structure and even world knowledge (as in the examples (51) above) are used as soon as they are available. Hence, parsing (52) is an incremental process: the hearer hears *Kim* first and as soon as the first sounds of *believe* reach her the available information is integrated and hypothesis regarding further parts of the utterance are build.[32]

It is sometimes claimed in personal communication that current Minimalist theories are better suited to explain production (generation) than perception (parsing). But these models are as implausible for generation as they are for parsing. The reason is that it is assumed that there is a syntax component that generates structures that are then shipped to the interfaces. This is not what happens in generation though. Usually speakers know what they want to say (at least partly), that is, they start with semantics. If one assumed that information coming from the conceptual-intentional system triggers the creation of syntactic structure, this would not help, since one would then have to build complex utterances by creating linguistic objects starting with the most deeply embedded and keeping them in memory until they are finally spelled out. As Labelle (2007) pointed out this is unrealistic since our availible short-term memory is not sufficient for such tasks.

> references apart from personal communication

So talking about larger syntactic units that are shipped to interpretation is fundamentally wrong from both the parsing and production perspective. Of course one could claim that actual processing is performance and Minimalist theories are dealing with competence but this would beg the question why the vocabulary of performance theories is used in a competence theory.[33]

There are some researchers who are aware of the problematic aspects of the

---

[32]Note that the architecture in Figure 6 poses additional problems. An *numeration* is a selection of lexical items that is used in a derivation. Since a multitude of empty elements are assumed in Minimalist analyses it is unclear how such a numeration is constructed since it cannot be predicted at the lexical level which empty elements will be needed in the course of a derivation. Due to the empty elements, there are infinitely many possible numerations that might be appropriate for the analysis of a given input string. The problem would disappear if derivations could draw lexical items from the lexicon directly.

[33]Chomsky (2008), for example, talks of "memory [that] is required to identify these properties at the semantic interface C-I" (p. 145) and "computational resources" (p. 138).

buttom up view standardly assumed in Minimalism: Phillips (2003) and Chesi (2015). Phillips assumes that structures relevant for phenomena such as ellipsis, coordination and fronting are built up incrementally. These constituents are then reordered in later steps by transformations. For example, in the analysis of (53), the string *Wallace saw Gromit in* forms a constituent where *in* is dominated by a node with the label P(P). This node is then turned into a PP in a subsequent step (p. 43–44).

(53) Wallace saw Gromit in the kitchen.

While this approach is a transformation-based approach, the kind of transformation here is very idiosyncratic and incompatible with other variants of the theory. In particular, the modification of constituents contradicts the assumption of Structure Preservation when applying transformations as well as the *No Tampering Condition* of Chomsky (2008). Furthermore, the conditions under which an incomplete string such as *Wallace saw Gromit in* forms a constituent are not entirely clear. Furthermore, the *Parser is Grammar* approach that Phillips assumes is per definition not process-neutral as Sag & Wasow (2011: 371) point out. Hence, this model is not suited as a competence model that has any relevance for production, which starts out with meaning.

Chesi (2015) suggests a top-down derivational approach. For nonlocal dependencies he assumes a store that is basically equivalent to the percolation mechanism suggested by Gazdar (1981), the one that is used in GPSG and HPSG (Borsley & Crysmann 2018, Chapter 14 of this volume). While this top-down mechanism is better suited for parsing than the bottom-up procedures generally assumed, it fails in cases of noisy input. Humans are capable of working on partial input an this is perfectly compatible with constraint-based models of grammars (Pullum & Scholz 2001: Section 3.2) while derivational models be they top-down or bottom-up will be stuck depending on the situation. Like Phillips' proposal, Chesi's approach is also not suited for performance models of production.

A solution to the problem of compatibility with psycholinguistic results would be to drop the rhetoric about interfaces and any assumptions about directions of processing. This is possible for many of the works currently published in Minimalism. One would then get theories with Internal and External Merge that would basically be variants of Categorial Grammar (Berwick & Epstein 1995; Müller 2013: Section 2.3). As Müller (2013: 938–941) shows, Internal Merge corresponds to HPSG's Filler-Head-Schema and External Merge to the Specifier-Head Schema and the Head-Complement Schema. If statements about Phases and the order of derivations are avoided, certain Minimalist theories are translatable into

HPSG structures and hence are compatible with constraint-based views on grammar.[34]

5.2 Acquisition

Acquisition has long been a central concern for Chomskyans and it has long been argued that acquisition is made possible by the existence of a complex innate language faculty (Chomsky 1965: Section I.8). Since the early 1980s the dominant view has been that the language faculty consists of a set of principles responsible for the properties which languages share and a set of parameters responsible for the ways in which they may differ (Chomsky 1981: 6). On this view acquiring a grammatical system is a matter of parameter-setting (Chomsky 2000: 8). Proponents of HPSG have always been sceptical about these ideas (see e.g., the remarks about parameters in Pollard & Sag (1994: 31) and have favoured accounts with "an extremely minimal initial ontology of abstract linguistic elements and relations" (Green 2011: 378). Thus, the two frameworks appear to be very different in this area. It is not clear, however, that this is really the case.

The idea that acquiring a grammatical system is a matter of parameter-setting is only as plausible as the idea of a language faculty with a set of parameters. It seems fair to say that this idea has not been as successful as was hoped when it was first introduced in the early 1980s. Outsiders have always been sceptical, but they have been joined in recent times by researchers sympathetic to many Chomskyan ideas. Thus, Newmeyer (2005: 75) writes as follows:

> [...] empirical reality, as I see it, dictates that the hopeful vision of UG as providing a small number of principles each admitting of a small number of parameter settings is simply not workable. The variation that one finds among grammars is far too complex for such a vision to be realized.

At least some Minimalists have come to similar conclusions. Thus, Boeckx (2011: 206) suggests that:

---

[34]Of course some of the resulting analyses would be very different in spirit. For instance, it is usually not assumed that subjects of non-finite verbs are realized as traces and moved to positions higher up in the tree (see Section 4.2.4). The point here is that representational variants of Minimalism are possible much in the way of the representational variants of GB that were common in the 80ies (Koster 1978; 1987: 235; Kolb & Thiersch 1991; Haider 1993: Section 1.4; Frey 1993: 14; Lohnstein 1993: 87–88, 177–178; Fordham & Crocker 1994: 38). See also Veenstra (1998: 58) for a representational implementation of Minimalism.

> some of the most deeply-embedded tenets of the Principles-and-Parameters approach, and in particular the idea of Parameter, have outlived their usefulness. (Boeckx 2011: 206)

Much the same view is expressed in Hornstein (2009: 164–168).

A major reason for scepticism about parameters is that estimates of how many there are seem to have steadily increased. Fodor (2001: 734) considers that there might be just twenty parameters, so that acquiring a grammatical system is a matter of answering twenty yes/no-questions. Newmeyer (2005: 44) remarks that "I have never seen any estimate of the number of binary-valued parameters needed to capture all of the possibilities of core grammar that exceeded a few dozen". However, Roberts & Holmberg (2005) comment that "[n]early all estimates of the number of parameters in the literature judge the correct figure to be in the region of 50–100". Clearly, a hundred is a lot more than twenty. Newmeyer (2017: Section 6.3) speaks of "hundreds, if not thousands". This is worrying. As Newmeyer (2006: 6) observes, "it is an ABC of scientific investigation that if a theory is on the right track, then its overall complexity decreases with time as more and more problematic data fall within its scope. Just the opposite has happened with parametric theory. Year after year more new parameters are proposed, with no compensatory decrease in the number of previously proposed ones.".

The growing scepticism appears to tie in with the proposal by Hauser, Chomsky & Fitch (2002: 1573) that "FLN [the "Faculty of language–narrow sense"] comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces". On this view there seems to be no place for parameters within FLN. This conclusion is also suggested by Chomsky's remarks (2005) that "There is no longer a conceptual barrier to the hope that the UG [Universal Grammar] might be reduced to a much simpler form" (p. 8) and that "we need no longer assume that the means of generation of structured expressions are highly articulated and specific to language" (p. 9). It's hard to see how such remarks are compatible with the assumption that UG includes 50–100 parameters. But if parameters are not part of UG, it is not at all clear what their status might be.

It looks, then, as Chomskyans are gradually abandoning the idea of parameters. But if it is abandoned, grammar acquisition is not a matter of parameter-setting. Hence, it is not clear that Chomskyans can invoke any mechanisms that are not available to HPSG.

This might suggest that HPSG and Minimalism are essentially in the same boat where acquisition is concerned. However, this is not the case given the very different nature of grammatical systems in the two frameworks. The complex and

abstract structures that are the hallmark of Minimalism and earlier Chomskyan frameworks pose major problems for acquisition.

There is a related point about acquisition, which I make in my comparative syntax paper as follows:

One might think that the acquisition task is fairly simple if languages have essentially the same structures differing only in what is and what is not visible. But this seems doubtful. As Fodor (2001: 765) puts it, 'It is clear now that even if the structural scaffolding of sentences is everywhere fixed and the same, any particular sentence may be highly ambiguous with respect to how its words are attached to that scaffolding '. Essentially, the more complex the structure of sentences is and the more invisible material it may contain, the harder it is for the learner to determine where anything is. As Fodor (2001: 763) comments, on this view, 'natural language design is extremely cruel to children'.

Furthermore the machinery that is assumed in addition to the basic operations Internal and External Merge are by no means trivial. There are numerations (subsets of the lexicon) that are assumed to play a role in a derivation, there is Agree and it has to be somehow acquired what the restriction on possible probe/goal relations is. There is feature checking, some features are uninterpretable others are not, some are strong others are weak. Certain categories are Phase boundaries others are not. There are complex conditions on Labeling. It is this that has led to the assumption that acquisition must be assisted by a complex language faculty. In contrast, HPSG structures are quite closely related to the observable data and so pose less of a problem for acquisition and hence create less need for some innate apparatus. Thus, HPSG probably has an advantage over Minimalism in this area too.

5.3 Restrictiveness

There is one further issue that we should discuss here. It appears to be quite widely assumed that one advantage that Minimalism has over alternatives like HPSG is that it is more "restrictive", in other words that it makes more claims about what is and is not possible in language. It looks then as if there might be an argument for Minimalism here. It is not clear, however, that this is really the case.

Minimalism would be a restrictive theory making interesting claims about language if it assumed a relatively small number of parameters. However, it seems that the idea that there is just a small number of parameters seems to have been abandoned, and at least some minimalists have abandoned the idea of parame-

ters altogether (see Section 5.2). If there is either a large number of parameters or no parameters at all, Minimalism is not restrictive in the way that it once was. However, it does still embody some restrictions on grammatical systems. The assumption that syntactic structures are confined to binary branching is an important restriction, as is the assumption that expressions produced by Merge have the same label as one of the expressions that they consist of. But we have argued that both assumptions are quite dubious. It also seems to be assumed that case and agreement are features of all grammatical systems. This would be another important restriction, but this also seems dubious given that many languages show no clear evidence for one or both of these features. It looks to us, then, as if the restrictiveness of Minimalism is largely a matter of imposing certain dubious restrictions on grammatical systems.

Note also that there are problems with restrictiveness of a more formal nature. Earlier versions of MGG assumed $\overline{\text{X}}$ theory and although this was not assumed initially it was quickly argued that the $\overline{\text{X}}$ scheme is universal and that this is a restriction on grammatical systems that aids language acquisition (Haegeman 1994: 106). However, Kornai & Pullum (1990: 41, 47) show that $\overline{\text{X}}$ theory is not restrictive at all as soon as empty elements are allowed in grammars: all languages that can be analyzed with a context-free grammar can be analyzed with an $\overline{\text{X}}$ grammar with empty heads. Chomsky (1995b: Section 4.3) abandoned $\overline{\text{X}}$ theory and replaced it by notions like first-merged and later-merged (Chomsky 1995b: 245; 2008) but the principled problem remains. Since as many empty heads as needed can be assumed in any position the predictions as far as restrictiveness is concerned are limited. See also Hornstein (2009: 165) and Starke (2014: 140) on heads and features and restrictiveness.

An example that is usually discussed when it comes to restrictiveness is question formation (Musso et al. 2003). Researchers in MGG state that certain ways of expressing things do never occur although they may be imaginable. So the question is asked why questions are never formed by reversing the order of words in a string. So rather than (54b) the question that would correspond to (54a) would be (54c):

(54)   a.   Kim saw Sandy near the swimming pool.

      b.   Did Kim see Sandy near the swimming pool?

      c.   Pool swimming the near Sandy saw Kim?

Interestingly, such reorderings can be derived in systems that allow for so-called *remnant movement*, as Hubert Haider (p. c. 2018) pointed out. The element 5 can move to the left of 4. The unit containing 4 and 5 can move to the left of 3 and

[[5 [4 _]] [3 _]] can move to the left of 2 and so on.

(55) a. [1 [2 [3 4]]]

  b. [3 4] → [4 [3 _]] → [2 [4 [3 _]]] → [[4 [3 _]] [2 _]] →
     [1 [[4 [3 _]] [2 _]]] → [[[4 [3 _]] [2 _]] [1 _]]

Of course there are reasons for the absence of certain imaginable constructions in the languages of the world. The reason for the absence of question formation like (54c) is simply short-term memory. Operations like those are ruled out due to performance constraints and hence should not be modeled in competence grammars. So it is unproblematic that remnant movement systems allow the derivation of strings with reverse order and it is unproblematic that one might develop HPSG analyses that reverse strings. Similarly certain other restrictions have been argued not to be part of the grammar proper. For instance Subjacency (Baltin 1981: 262; 2017; Rizzi 1982: 57; Chomsky 1986: 38–40) does not hold in the form stated in MGG (Müller 2004; 2016a: Section 13.1.5) and it is argued that several of the island constraints should not be modeled by hard constraints in competence grammars. See Chaves (2018), Chapter 16 of this volume for further discussion.

It is true that the basic formalism does not pose any strong restrictions on what could be said in an HPSG theory. As Pollard (1996) points out, this is the way it should be. The formalism should not be the constraining factor. It should be powerful enough to allow everything to be expressed in insightful ways and in fact, the basic formalism of HPSG has Turing-power, the highest power in the Chomsky hierarchy (Pollard 1999). This means that the general formalism is above the complexity that is usually assumed for natural languages: mildly context-sensitive. What is important though is that theories of individual languages are much more restrictive, getting the generative power down (Müller 2016a: Chapter 17).

These remarks should not be understood as a suggestion that languages vary without limit, as Joos (1958: 96) suggested. No doubt there are universal tendencies and variation is limited, but the question is whether this is due to innate linguistic constraints or rather the question of what we do with language and how our general cognitive capabilities are structured. While Minimalism starts out with claims about universal features about languages and tries to confirm these claims in language after language, researchers working in HPSG aim to develop fragments of languages that are motivated by facts from these languages and generalize over several internally motivated grammars. This leaves the option open that languages have very little in common as far as syntax is concerned. For

example, Koenig & Michelson (2012) discuss the Northern Iroquoian language Oneida and argue that this language does not have syntactic valence. If they are correct, not even central concepts like valence and argument structure would be universal. The only remaining universal would be that we combine linguistic objects. This corresponds to Merge in Minimalism, without the restriction to binarity.

Acknowledgements

We know that this paper is very long. We nevertheless think that it is important to discuss the issues discussed in this paper and are very grateful to the editors for giving us this much space.

We thank Tom Wasow for discussion and Sebastian Nordhoff and Felix Kopecky▉ for help with Figure 1.

6 Todo

feature-driven movement

- there is altruistic movement

- If movement is cyclic and has to consume features and if it is unbounded, there have to be infinitely many lexical items per moved item. This follows from the argumentation of languages as infinite sets. Following the same logic the lexicon has to be infinite. This adds to the problem of determining the numeration.

References

Abeillé, Anne. 2006. In defense of lexical coordination. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in formal syntax and semantics*, vol. 6, 7–36. Paris: CNRS. http://www.cssp.cnrs.fr/eiss6/, accessed 2018-2-25.

An, Aixiu & Anne Abeillé. 2017. Agreement and interpretation of binominals in french. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 26–43. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-an-abeille.pdf.

Draft of 26th December 2018, 19:51

Aoun, Joseph E. & Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar*. Cambridge, MA: MIT Press.

Aoun, Joseph, Lina Choueiri & Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32(3). 371–404.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function change*. Chicago: The University of Chicago Press.

Baltin, Mark. 1981. Strict bounding. In Carl Lee Baker & John J. McCarthy (eds.), *The logical problem of language acquisition*, 257–295. Cambridge, MA: MIT Press.

Baltin, Mark. 2017. Extraposition. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, 2nd edn. (Blackwell Handbooks in Linguistics), 237–271. Oxford: Blackwell Publishers Ltd.   DOI:10.1002/9781118358733

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. 1990. 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: originally Gunter Narr Verlag now Stauffenburg Verlag.

Bender, Emily M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford University dissertation. http : / / faculty . washington.edu/ebender/dissertation/, accessed 2018-2-25.

Bender, Emily M. 2002. Review of *Step by step: Essays on Minimalist Syntax in honor of Howard Lasnik* by Roger Martin, David Michaels, and Juan Uriagereka. *Journal of Linguistics* 38(2). 432–439.

Berwick, Robert C. & Samuel David Epstein. 1995. On the convergence of 'Minimalist' Syntax and Categorial Grammar. In Anton Nijholt, Giuseppe Scollo & Rene Steetskamp (eds.), *Algebraic methods in language processing*, 143–148. Enschede: University of Twente. http://eprints.eemcs.utwente.nl/9555/01/twlt10.pdf, accessed 2018-9-25.

Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the Stimulus revisited. *Cognitive Science* 35(7). 1207–1242.

Bildhauer, Felix & Philippa Cook. 2010. German multiple fronting and expected topic-hood. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 68–79. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.

Boeckx, Cedric. 2003. *Islands and chains*. Amsterdam: John Benjamins Publishing Co.

Boeckx, Cedric. 2011. Approaching parameters from below. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 205–221. Oxford: Oxford University Press.

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.

Borsley, Robert D. 2006a. On the nature of Welsh VSO clauses. *Lingua* 116(4). 462–490.

Borsley, Robert D. 2006b. *Syntactic and lexical approaches to unbounded dependencies*. Essex Research Reports in Linguistics 49. University of Essex. 31–57. http://core.ac.uk/download/pdf/4187949.pdf#page=35, accessed 2018-2-25.

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

Borsley, Robert D. 2010. An HPSG approach to Welsh unbounded dependencies. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 47–67. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.

Borsley, Robert D. 2011. Constructions, functional heads, and comparative correlatives. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics*, vol. 8, 7–20. Paris: CNRS. http://www.cssp.cnrs.fr/eiss8/borsley-eiss8.pdf, accessed 2018-8-23.

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

Borsley, Robert D. 2016. HPSG and the nature of agreement in Archi. In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi: Complexities of agreement in cross-theoretical perspective*, 118–149. Oxford: Oxford University Press.

Borsley, Robert D. 2017. Explanations and 'engineering solutions'? Aspects of the relation between Minimalism and HPSG. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 82–102. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-borsley.pdf, accessed 2018-9-9.

Borsley, Robert D. 2018. Comparative syntax: An HPSG perspective. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Clausal architecture and its consequences: Synchronic and diachronic perspectives*, vol. 1. In Preparation.

Bošković, Željko. 2009. Unifying first and last conjunct agreement. *Natural Language and Linguistic Theory* 27(3). 455–496. DOI:10.1007/s11049-009-9072-6

Bresnan, Joan & Jane Grimshaw. 1978. The syntax of free relatives in English. *Linguistic Inquiry* 9. 331–392.

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

Bruening, Benjamin. 2018. The lexicalist hypothesis: Both wrong and superfluous. *Language* 94(1). 1–42. DOI:10.1353/lan.2018.0000

Chesi, Cristiano. 2015. On directionality of phrase structure building. *Journal of Psycholinguistic Research* 44(1). 65–89. DOI:10.1007/s10936-014-9330-6

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

Chomsky, Noam. 1964. Degrees of grammaticalness. In Jerry A. Fodor & Jerrold J. Katz (eds.), *The structure of language*, 384–389. Englewood Cliffs, NJ: Prentice-Hall.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum Press.

Chomsky, Noam. 1977. On wh-movement. In P. W. Culicover, T. Wasow & A. Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.

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

Chomsky, Noam. 1986. *Barriers* (Linguistic Inquiry Monographs 13). Cambridge, MA: MIT Press.

Chomsky, Noam. 1990. On formalization and formal linguistics. *Natural Language and Linguistic Theory* 8(1). 143–147.

Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), 1–52. Cambridge, MA/London: MIT Press.

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

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

Chomsky, Noam. 1999. *Derivation by phase.* MIT Occasional Papers in Linguistics 18. MIT. Reprint in: Michael Kenstowicz, ed. 2001. Ken Hale. A Life in Language. Cambridge, MA: MIT Press, 1–52.

Chomsky, Noam. 2000. *New horizons in the study of language and mind.* Cambridge, UK: Cambridge University Press.

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

Chomsky, Noam. 2002. *On nature and language.* Cambridge, UK: Cambridge University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.

Chomsky, Noam & Howard Lasnik. 1995. The theory of principles and parameters. In Noam Chomsky (ed.), *The Minimalist Program*, chap. 1, 13–127. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective.* New York, Oxford: Oxford University Press.

Cinque, Guglielmo & Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 51–65. Oxford: Oxford University Press.

Citko, Barbara. 2008. Missing labels. *Lingua* 118(7). 907–944.

Citko, Barbara. 2011. *Symmetry in syntax: Merge, move and labels.* Vol. 129. Cambridge, UK: Cambridge University Press. DOI:10.1017/CBO9780511794278

Cole, Peter, Wayne Harbert & Gabriella Hermon. 1982. Headless relative clauses in Quechua. *International Journal of American Linguistics* 48(2). 113–124. http://www.jstor.org/stable/1264676, accessed 2018-9-13.

Crysmann, Berthold. 2012. Resumption and island-hood in Hausa. In Philippe de Groote & Mark-Jan Nederhof (eds.), *Formal Grammar. 15th and 16th International Conference on Formal Grammar, FG 2010 Copenhagen, Denmark, August*

2010, FG 2011 Lubljana, Slovenia, August 2011, vol. 7395 (Lecture Notes in Computer Science). Berlin: Springer Verlag.

Crysmann, Berthold. 2016. An underspecification approach to Hausa resumption. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway-King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 194–214. Stanford, CA: CSLI Publications.

Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition. foundations of syntax*. Vol. 1 (Oxford Linguistics). Oxford: Oxford University Press.

Culicover, Peter W. & Ray Jackendoff. 1999. The view from the periphery: The English comparative correlative. *Linguistic Inquiry* 30(4). 543–571. DOI:10.1162/002438999554200

Culicover, Peter W. & Ray S. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

De Kuthy, Kordula. 2000. *Discontinuous NPs in German — A case study of the interaction of syntax, semantics and pragmatics*. Saarbrücken: Universität des Saarlandes dissertation.

Donati, Caterina. 2006. On *wh*-head-movement. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), Wh-*movement: Moving on* (Current Studies in Linguistics 42), 21–46. Cambridge, MA: MIT Press.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Claire Grover & Enric Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science, vol. 12: Studies in HPSG*, chap. 1, 1–32. Edinburgh: Centre for Cognitive Science, University of Edinburgh. ftp://ftp.cogsci.ed.ac.uk/pub/CCS-WPs/wp-12.ps.gz, accessed 2018-3-20.

Epstein, Samuel David & T. Daniel Seely. 2006. *Derivations in Minimalism* (Cambridge Studies in Linguistics 111). Cambridge, UK: Cambridge University Press.

Falk, Yehuda N. 2010. An unmediated analysis of relative clauses. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2010 conference*, 207–227. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/15/, accessed 2018-3-20.

Fanselow, Gisbert. 2001. Features, θ-roles, and free constituent order. *Linguistic Inquiry* 32(3). 405–437.

Flickinger, Daniel P. 1987. *Lexical rules in the hierarchical lexicon*. Stanford University dissertation.

Flickinger, Daniel P. 2008. Transparent heads. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*,

87–94. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/abstr-flickinger.shtml, accessed 2018-2-25.

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

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

Fong, Sandiway. 2014. Unification and efficient computation in the Minimalist Program. In L. Francis & L. Laurent (eds.), *Language and recursion*, 129–138. Berlin: Springer Verlag.

Fong, Sandiway & Jason Ginsburg. 2012. Computation with doubling constituents: Pronouns and antecedents in Phase Theory. In Anna Maria Di Sciullo (ed.), *Towards a Biolinguistic understanding of grammar: Essays on interfaces* (Linguistik Aktuell/Linguistics Today 194), 303–338. Amsterdam: John Benjamins Publishing Co.

Fordham, Andrew & Matthew Walter Crocker. 1994. Parsing with principles and probabilities. In Judith L. Klavans Philip Resnik (ed.), *The balancing act: combining symbolic and statistical approaches to language*. Las Cruces, New Mexico, USA: Association for Computational Linguistics.

Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus* (studia grammatica 35). Berlin: Akademie Verlag.

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

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

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

Green, Georgia M. 2011. Modeling grammar growth: Universal Grammar without innate principles or parameters. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 378–403. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd. DOI:10.1002/9781444395037.ch11

Grewendorf, Günther. 1993. German: A grammatical sketch. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax –*

Draft of 26th December 2018, 19:51

Ein internationales Handbuch zeitgenössischer Forschung, vol. 9.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1288–1319. Berlin: Walter de Gruyter Verlag. DOI:10.1515/9783110095869.1

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

Gunji, Takao. 1986. Subcategorization and word order. In William J. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, 1–21. Stanford, CA: CSLI Publications.

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19. 377–419.

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

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

Haider, Hubert. 2018. *On Minimalist theorizing and scientific ideology in grammar theory*. Ms. Universität Salzburg.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), 53–109. Cambridge, MA/London: MIT Press.

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

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

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

Höhle, Tilman N. 1982. Explikationen für „normale Betonung" und „normale Wortstellung". In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15), 75–153. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag. Republished as Höhle (2018).

Höhle, Tilman N. 2018. Explikationen für „normale Betonung" und „normale Wortstellung". In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 107–191. Berlin: Language Science Press. DOI:10.5281/zenodo.1145680

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar.* Cambridge University Press. DOI:10.1017/CBO9780511575129

Hornstein, Norbert, Jairo Nunes & Kleantes K. Grohmann. 2005. *Understanding Minimalism* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.

Huybregts, Riny & Henk van Riemsdijk. 1982. *Noam chomsky on the Generative Enterprise.* Dordrecht: Foris.

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

Johannessen, Janne Bondi. 1996. Partial agreement and coordination. *Linguistic Inquiry* 27(4). 661–676.

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

Jones, Morris & Alun R. Thomas. 1977. *The Welsh language: Studies in its syntax and semantics.* Cardiff: University of Wales Press.

Joos, Martin. 1958. Comment on *Phonemic overlapping* by Bernhard Bloch. In Martin Joos (ed.), *Readings in linguistics: The development of descriptive linguistics in America since 1925.* New York: American Council of Learned Societies. https://archive.org/details/readingsinlingui00joos.

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.

Kiss, Tibor. 2001. Configurational and relational scope determination in German. In Walt Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7), 141–175. Stanford, CA: CSLI Publications.

Koenig, Jean-Pierre & Karin Michelson. 2012. The (non)universality of syntactic selection and functional application. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, vol. 9, 185–205. Paris: CNRS.

Kolb, Hans-Peter & Craig L. Thiersch. 1991. Levels and empty categories in a Principles and Parameters based approach to parsing. In Hubert Haider & Klaus Netter (eds.), *Representation and derivation in the theory of grammar* (Studies

Draft of 26th December 2018, 19:51

in Natural Language and Linguistic Theory 22), 251–301. Dordrecht: Kluwer Academic Publishers.

Kornai, András & Geoffrey K. Pullum. 1990. The X-bar Theory of phrase structure. *Language* 66(1). 24–50.

Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris Publications.

Koster, Jan. 1986. The relation between pro-drop, scrambling, and verb movements. *Groningen Papers in Theoretical and Applied Linguistics* 1. 1–43.

Koster, Jan. 1987. *Domains and dynasties: The radical autonomy of syntax*. Dordrecht: Foris Publications.

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

Labelle, Marie. 2007. Biolinguistics, the Minimalist Program, and psycholinguistic reality. *Snippets* 14. 6–7. http://www.ledonline.it/snippets/allegati/snippets14002.pdf, accessed 2018-2-25.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–391.

Larson, Richard K. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* (21). 589–632.

Levine, Robert D. & Ivan A. Sag. 2003. Some empirical issues in the grammar of extraction. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 236–256. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-9-25.

Li, Yen-hui Audrey. 2008. Case, 20 years later. In Marjorie K.M. Chan & Hana Kang (eds.), *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, 41–68. Columbus, Ohio: The Ohio State University.

Lohnstein, Horst. 1993. *Projektion und Linking: Ein prinzipienbasierter Parser fürs Deutsche* (Linguistische Arbeiten 287). Tübingen: Max Niemeyer Verlag.

Lohnstein, Horst. 2014. Artenvielfalt in freier Wildbahn: Generative Grammatik. In Jörg Hagemann & Sven Staffeldt (eds.), *Syntaxtheorien: Analysen im Vergleich* (Stauffenburg Einführungen 28), 165–185. Tübingen: Stauffenburg Verlag.

Marslen-Wilson, William D. 1975. Sentence perception as an interactive parallel process. *Science* 189(4198). 226–228. DOI:10.1126/science.189.4198.226

Matsuyama, Tetsuya. 2004. The N *after* N Construction: A constructional idiom. *English Linguistics* 1(11). 55–84.

McCawley, James D. 1968. Lexical insertion in a Transformational Grammar without Deep Structure. In *Fourth regional meeting of the Chicago Linguistic Society*, 71–81. University of Chicago.

McCloskey, James. 2002. Resumption, successive cyclicity and the locality of operations. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the Minimalist Program*, 184–226. Oxford: Blackwell Publishers Ltd. DOI:10.1002/9780470755662

McCloskey, James. 2006. Resumption. In Martin Everaert & Henk van Riemsdijk (eds.), *The blackwell companion to syntax*, 94–117. Oxford: Blackwell Publishers Ltd.

Meinunger, André. 2000. *Syntactic aspects of topic and comment* (Linguistik Aktuell/Linguistics Today 38). Amsterdam: John Benjamins Publishing Co.

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

Meurers, Walt Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2). 161–217.

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

Michaelis, Laura A. & Knud Lambrecht. 1996. Toward a Construction-based theory of language function: The case of nominal extraposition. *Language* 72(2). 215–247.

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

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

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

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

Müller, Stefan. 2003. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. DOI:10.1017/S0022226703002032

Müller, Stefan. 2004. Complex NPs, subjacency, and extraposition. *Snippets* 8. 10–11.

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

Müller, Stefan. 2007b. The Grammix CD Rom: A software collection for developing typed feature structure grammars. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar Engineering across Frameworks 2007* (Studies in Computational Linguistics ONLINE). Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/, accessed 2018-2-25.

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.

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

Müller, Stefan. 2013. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950. DOI:10.1353/lan.2013.0061

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

Müller, Stefan. 2015a. HPSG – A synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), 937–973. Berlin: Walter de Gruyter. DOI:10.1515/9783110363708-004

Müller, Stefan. 2015b. 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. 2016a. *Grammatical theory: From Transformational Grammar to constraint-based approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2016b. Satztypen: Lexikalisch oder/und phrasal. In Rita Finkbeiner & Jörg Meibauer (eds.), *Satztypen und Konstruktionen im Deutschen* (Linguistik – Impulse und Tendenzen 65), 72–105. Berlin: de Gruyter.

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

Müller, Stefan. 2018a. *Germanic syntax*. Ms. Humboldt Universität zu Berlin, to be submitted to Language Science Press. Berlin. https://hpsg.hu‑berlin.de/~stefan/Pub/germanic.html, accessed 2018-3-20.

Müller, Stefan. 2018b. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 2nd edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.5281/zenodo.1193241

Müller, Stefan. 2018c. The end of lexicalism as we know it? *Language* 94(1). e54–e66. DOI:10.1353/lan.2018.0014

Müller, Stefan, Felix Bildhauer & Philippa Cook. 2012. Beschränkungen für die scheinbar mehrfache Vorfeldbesetzung im Deutschen. In Colette Cortès (ed.), *Satzeröffnung: Formen, Funktionen, Strategien* (Eurogermanistik 31), 113–128. Tübingen: Stauffenburg Verlag.

Musso, Mariacristina, Andrea Moro, Volkmar Glauche, Michel Rijntjes, Jürgen Reichenbach, Christian Büchel & Cornelius Weiller. 2003. Broca's area and the language instinct. *Nature Neuroscience* 6(7). 774–781. DOI:10.1038/nn1077

Newmeyer, Frederick J. 1986. *Linguistic theory in America*. 2nd edn. New York, NY: Academic Press.

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

Newmeyer, Frederick J. 2006. *A rejoinder to 'On the role of parameters in Universal Grammar: A reply to Newmeyer' by Ian Roberts and Anders Holmberg*. https://ling.auf.net/lingbuzz/000248, accessed 2018-9-25.

Newmeyer, Frederick J. 2017. Where, if anywhere, are parameters? A critical historical overview of parametric theory. In Claire Bowern, Laurence Horn & Raffaella Zanuttini (eds.), *On looking into words (and beyond): Structures, relations, analyses* (Empirically Oriented Theoretical Morphology and Syntax 3), 547–569. Language Science Press. DOI:10.5281/zenodo.495467

Ott, Dennis. 2011. A note on free relative clauses in the theory of Phases. *Linguistic Inquiry* 42(1). 183–192.

Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen A. G. Groenendijk, Dick de Jongh & Martin J. B. Stokhof (eds.), *Studies in Discourse Representation Theory and the theory of generalized quantifiers*

(Groningen-Amsterdam Studies in Semantics 8), 115–143. Dordrecht: Foris Publications.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34(1). 37–90.

Pollard, Carl. 1997. *Lectures on the foundations of hpsg*. Ms. Ohio State University. https://www1.essex.ac.uk/linguistics/external/clmt/papers/hpsg/pollard-foundations.ps, accessed 2018-9-29.

Pollard, Carl J. 1996. *The nature of constraint-based grammar*. http://lingo.stanford.edu/sag/L221a/pollard-96.txt, accessed 2018-2-25. Paper presented at the Pacific Asia Conference on Language, Information, and Computation, Kyung Hee University, Seoul, Korea.

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

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

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

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

Postal, Paul M. 1970. On the surface verb 'remind'. *Linguistic Inquiry* 1(1). 37–120.

Postal, Paul M. 2003. (Virtually) conceptually necessary. *Journal of Linguistics* 39(3). 599–620. DOI:10.1017/S0022226703002111

Postal, Paul M. 2004. *Skeptical linguistic essays*. Oxford: Oxford University Press.

Postal, Paul M. 2010. *Edge-based clausal syntax: A study of (mostly) English object structure*. Cambridge, MA: MIT Press.

Pullum, Geoffrey K. 1989. Formal linguistics meets the Boojum. *Natural Language and Linguistic Theory* 7(1). 137–143. DOI:10.1007/BF00141350

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

Radford, Andrew. 1993. Head-hunting: On the trail of the nominal Janus. In Greville G. Corbett, Norman M. Fraser & Scott McGlashan (eds.), *Heads in grammatical theory*, 73–113. Cambridge University Press. DOI:10.1017/CBO9780511659454.005

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

Ramchand, Gillian. 2005. Post-Davidsonianism. *Theoretical Linguistics* 31(3). 359–373.   DOI:10.1515/thli.2005.31.3.359

Richards, Marc. 2015. Minimalism. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42.3 (Handbooks of Linguistics and Communication Science), 803–839. Berlin: Mouton de Gruyter. DOI:10.1515/9783110363685

Rizzi, Luigi. 1982. Violations of the *wh* island constraint and the Subjacency Condition. In Luigi Rizzi (ed.), *Issues in Italian syntax* (Studies in Generative Grammar 11), 49–76. Dordrecht: Foris Publications.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 2014. Syntactic Cartography and the Syntacticisation of Scope-Discourse Semantics. In Anne Reboul (ed.), *Mind, Values, and Metaphysics: Philosophical Essays in Honor of Kevin Mulligan*, vol. 2, 517–533. Cham: Springer International Publishing.   DOI:10.1007/978-3-319-05146-8_30

Roberts, Ian. 2005. *Principles and Parameters in a VSO language: A case study in Welsh*. Oxford: Oxford University Press.

Roberts, Ian F. & Anders Holmberg. 2005. On the role of parameters in Universal Grammar: A reply to Newmeyer. In Hans Broekhuis, N. Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk* (Studies in Generative Grammar 86), 538–553. Berlin: Mouton de Gruyter.   DOI:10.1515/9783110892994.538

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT dissertation. http://files.eric.ed.gov/fulltext/ED016965.pdf, accessed 2018-2-25. Reproduced by the Indiana University Linguistics Club and later published as Ross (1986).

Ross, John Robert. 1986. *Infinite syntax!* (Language and Being 5). Norwood, New Jersey: Ablex Publishing Corporation.

Rouveret, Alain. 1994. *Syntaxe du gallois: principes généraux et typologie*. Paris: Editions du CNRS.

Sadler, Louisa. 1988. *Welsh syntax: a government binding approach*. London: Croom Helm.

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

Sag, Ivan A. 2003. Coordination and underspecification. In Jongbok Kim & Stephen Mark Wechsler (eds.), *The proceedings of the 9th International Confer-*

Draft of 26th December 2018, 19:51

ence on Head-Driven Phrase Structure Grammar, 267–291. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/3/, accessed 2018-9-25.

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

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

Sag, Ivan A. & Janet Dean Fodor. 1994. Extraction without traces. In Raul Aranovich, William Byrne, Susanne Preuss & Martha Senturia (eds.), *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, 365–384. Stanford, CA: CSLI Publications/SLA. http://lingo.stanford.edu/sag/papers/sag-fodor-wccfl.pdf, accessed 2018-2-25.

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

Sag, Ivan A. & Thomas Wasow. 2015. Flexible processing and the design of grammar. *Journal of Psycholinguistic Research* 44(1). 47–63.

Salkoff, Morris. 1988. Analysis by fusion. *Lingvisticæ Investigationes* 12(1). 49–84. DOI:10.1075/li.12.1.03sal

Sauerland, Uli & Paul Elbourne. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33(2). 283–319.

Sproat, Richard. 1985. Welsh syntax and VSO structure. *Natural Language and Linguistic Theory* 3(2). 173–216.

Stabler, Edward P. 2001. Minimalist Grammars and recognition. In Christian Rohrer, Antje Rossdeutscher & Hans Kamp (eds.), *Linguistic form and its computation* (Studies in Computational Linguistics 1), 327–352. Stanford, CA: CSLI Publications.

Starke, Michael. 2014. Towards elegant parameters: Language variation reduces to the size of lexically-stored trees. In M. Carme Picallo (ed.), *Linguistic variation in the Minimalist framework*, 140–152. Oxford: Oxford University Press.

Steedman, Mark J. 1991. Structure and intonation. *Language* 67(2). 260–296.

Sternefeld, Wolfgang. 2006. *Syntax: Eine morphologisch motivierte generative Beschreibung des Deutschen* (Stauffenburg Linguistik 31). Tübingen: Stauffenburg Verlag.

Sternefeld, Wolfgang & Frank Richter. 2012. Wo stehen wir in der Grammatiktheorie? — Bemerkungen anläßlich eines Buchs von Stefan Müller. *Zeitschrift für Sprachwissenschaft* 31(2). 263–291. DOI:10.1515/zfs-2012-0010

Taghvaipour, Mehran A. 2010. *Persian relative clauses in HPSG: An Analysis of Persian relative clause constructions in Head-Driven Phrase Structure Grammar*. Saarbrücken: LAMBERT Academic Publishing.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634. DOI:10.1126/science.7777863

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1996. Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In Toshio Inui & James L. McClelland (eds.), *Information integration in perception and communication* (Attention and Performance XVI), 457–478. Cambridge, MA: MIT Press.

Vaillette, Nathan. 2001. Irish gaps and resumptive pronouns in HPSG. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG-2001 Conference, Norwegian University of Science and Technology, Trondheim*, 284–299. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2001/, accessed 2015-12-23.

Van Eynde, Frank. 2015. *Predicative constructions: From the Fregean to a Montagovian treatment* (Studies in Constraint-Based Lexicalism). Stanford, CA: CSLI Publications.

Van Koppen, Marjo. 2005. *One probe–two goals: Aspects of agreement in Dutch dialects*. Vol. 105 (LOT Dissertation Series). Utrecht Institute of Linguistics: Netherlands Graduate School of Linguistics.

Veenstra, Mettina Jolanda Arnoldina. 1998. *Formalizing the Minimalist Program*. Rijksuniversiteit Groningen Ph.D. thesis.

Williams, Edwin. 1989. The anaphoric nature of θ-roles. *Linguistic Inquiry* 20(3). 425–456.

Willis, David. 2011. The limits of resumption in Welsh *wh*-dependencies. In Alain Rouveret (ed.), *Resumptive pronouns at the interfaces*, 189–222. Amsterdam: John Benjamins Publishing Co.

Wurmbrand, Susanne. 2003. *Long passive (corpus search results)*. Ms. University of Connecticut.

Draft of 26th December 2018, 19:51