Evaluating theories: Counting nodes and the question of constituency

Stefan Müller[a]

[a] Humboldt Universität Berlin, Institut für deutsche Sprache und Linguistik,
St.Mueller@hu-berlin.de.

Abstract. This paper is a reply to Timothy Osborne's paper *Tests for constituents: What they really reveal about the nature of syntactic structure* that appeared 2018 in *Language under Discussion*. This paper discusses how constituent tests work and why it is no problem if they are not applicable. It is argued that Osborne's claims regarding simplicity of Dependency Grammar (DG) in comparison to Phrase Structure Grammar (PSG) are unwarranted and that DG models that include semantics make use of auxiliary structure that is equivalent to the nodes assumed in PSG. A final section of the paper discusses the general validity of counting nodes for theory evaluation and the assumption of empty elements vs. specialized phrasal rules.

Keywords: Dependency Grammar, Phrase Structure Grammar, complexity, constituency

1 Introduction

Timothy Osborne has published a series of papers in which he claimed that Dependency Grammar is simpler than phrase structure grammar and hence has to be preferred for reasons of parsimony (Osborne & Groß 2016: 132, Osborne 2018: 2). One of these papers (Osborne 2018) is a target paper in *Language under Discussion*. Osborne explained why he thinks that grammarians working in the phrase structure tradition got their constituent tests wrong and argued that constituents at the level of complete phrases are justified but constituents at the subphrasal level are not, which, according to him, is obvious if one looks at the results of the tests.

    This paper is a reply to Osborne's discussion article including some more general comments on theory evaluation at the end. The paper will be structured as follows: I first discuss constituent tests and some of the claims Osborne has made (Section 2). I then add remarks on evaluating theories in Section 3. I begin with the Dependency/Phrase Structure Grammar comparison and the claim that grammars without sub-phrasal constituents are simpler (Section 3.1). It is shown

that Dependency Grammar approaches are of comparable complexity once semantics is taken into account. Section 3.2 compares the actual grammar rules of Phrase Structure Grammars and Dependency Grammars and motivates nodes as attachment sides for adjuncts. Subsection 3.3 makes a more general point on theory evaluation discussing recent Minimalist proposals and comparing them to earlier phrase structure approaches with fewer nodes and more rules. Section 4 concludes the paper.

2 Constituency tests

This section deals with the way constituent tests are formulated, with the coordination test in particular and with partial constituents in German.

2.1 Words and constituents and the implications of test applicability

Osborne applies tests like topicalization and proform substitution and finds that they sometimes do not apply to words. For example, the pre-nominal adjective *syntactic* and the verb *show* in (1a) cannot be topicalized as (1b) and (1c) show, respectively:

(1) a. Trees can show syntactic structure.
     b. * ...and *syntactic*, trees can show structure.
     c. * ...and *show* trees can syntactic structure.
     d. * ...and *can show syntactic structure*, trees.

Osborne discusses 15 constituency tests in total and applies a selection of 10, among them pseudo clefting and answer fragments. He concludes that words at the subphrasal level should not be regarded as constituents since most of the applied constituency tests are negative. However, what he misses is the exact formulation of the tests. For example, the topicalization test says: Sequences that can be moved are constituents (von Stechow & Sternefeld 1988: 108, Eroms 2000: 35, Dürscheid 2003: Section 3, Duden 2005: §1170, Flohr & Lobin 2009: 133, Remberger 2016, Schäfer 2015: Section 10.3).[1,2] This means that the constituent test is a sufficient and not a necessary condition: if X holds then the string under consideration is a constituent. The formulation in Müller (2016: 10) is even more cautious and talks about "strong indicator of constituent status".[3] Similar statements can be found in Machicao y Priemer (2018). So we have: if X holds

---

[1] "Verschiebeprobe. Was sich verschieben, umstellen läßt, ist eine Konstituente." (von Stechow & Sternefeld 1988: 108) "Nach der Verschiebeprobe ist das, was verschoben werden kann, eine Konstituente." (Flohr & Lobin 2009: 133) "Grundsätzlich kann verallgemeinert werden, dass alles, was sich vor dem finiten Verb im Vorfeld befindet, also auch *wir* in (1), eine Konstituente (ein Syntagma oder eine selbständige Phrase, [...]) bildet." (Remberger 2016).

[2] I could not check all the references listed by Osborne but statements by Haegeman (1994: 35) are similar. Wöllstein (2010: 16) and Dürscheid (2003: 54) are exceptional. Dürscheid states that the other direction of the implication is relevant too and Wöllstein formulates four out of five tests with the implication in the opposite way: *A constituent is frontable.* This means that if something is a constituent it has to be frontable, replaceable by pronouns, deletable and cordinatable. These tests would exclude many partial phrases and in particular words. Since Wöllstein formulated one of the constituent tests in a different way, I believe these statements were made with different intentions.

[3] See Müller (2016: Section 1.3.2) for a discussion of problems related to the tests. While fronting in German is a rather reliable test there is a class of exceptions: so-called apparent multiple frontings (Müller 2003). While these constituents in front of the finite verb can and should be analyzed as one constituent (Müller 2005, 2017), the respective patterns are problematic for constituent tests.

then the string under consideration is likely a constituent. In the case of words, the question of constituent status is trivially decided: constituents are parts of a larger structure and since all more complex structures consist of words (ignoring the possibility of having affixes as parts of syntax), words are constituents (Adger 2003: 66, Schäfer 2015: 312, Müller 2016: 7). I said "trivially" above. Actually, the issue of determining what a word is is not trivial at all, but Dependency Grammar has to make and justify the same decision. Words have certain properties like case, gender, and number. They form distribution classes and on the basis of their distribution, we assign labels to them: something like 3rd person singular count noun. Determining this category involves a test that is also one of the constituent tests: the substitution test. It is clear that some of the standard constituent tests do not work at the subphrasal level. For instance, the pronominalization test with personal pronouns usually does not work since pronouns refer to discourse referents and partial noun phrases do not refer. But the non-replaceability of partial NPs is not a problem, since the test is: if you can replace something with a pronoun, this indicates that this is a constituent. Nothing is said about cases in which this is impossible.

2.2 Coordination

Osborne discusses coordination as a test and notes that it sometimes does not apply, but as he notes himself there are reasons for this and even if there were no sensible explanations this would just mean that it is impossible to apply one of the tests. Let's have a look at coordination and what it tells us about subconstituents. The nominal structure in (2a) can refer to a specific set of skillful children and smart parents, possibly present in the current situation, or it can refer to a specific set of skillful children and some smart parents. Assuming a phrase including nouns and adjectives ($\overline{N}$ in $\overline{X}$ theory, Jackendoff 1977), it is easy to get the coordination facts right: (2b) is a coordination of two $\overline{N}$s and (2c) is the coordination of two full NPs, *smart parents* being an NP without a determiner.

(2) a. these skillful children and smart parents
     b. [$_{NP}$ these [$_{\overline{N}}$ [$_{\overline{N}}$ skillful children] and [$_{\overline{N}}$ smart parents]]]
     c. [$_{NP}$ [$_{NP}$ these skillful parents] and [$_{NP}$ smart parents]]

Of course, coordination is a very complex phenomenon consisting of various subphenomena: there is gapping, right-node raising and so on (see Osborne, 2018 for examples and references). Some argue that coordination is something entirely different from the rest of syntax and should be treated by special means (Goodall 1983, Hudson 1988, Osborne 2006). Some develop approaches that allow multidominance[4] (Crysmann 2008, Beavers & Sag 2004). It seems to be reasonable to attempt to stay as close as possible to the rules normally used rather than introducing new mechanisms for basic cases. If one has the constituents on the subphrasal level (which are needed for independent reasons, as discussed below), nothing special is needed for cases like (2). So, what is relevant for constituent tests is only a subset of coordination phenomena, namely symmetric coordination, that is, the coordination of two items of the same syntactic category.

---

[4] Multidominance would also account for examples like the following one by Osborne (p. c. 2018):

(i) the children [over here playing soccer] and [over there playing tag]

Multidominance approaches actually allow for material being used twice.

In contrast to symmetric coordination, parts of what we would like to call a constituent are discontinuous in cases of right node raising. The general problem is that we cannot easily restrict the test to symmetric coordination since we would need a pretheoretic way to tell the various types of coordination apart without referring to constituents. So, as argued by several researchers (cited in Osborne's paper), coordination has to be used with care and it has to be checked whether the assumption of a constituent is warranted by the results of other tests and by further assumptions made in the grammar under development.

Osborne (p. c. 2018) argued against assuming $\overline{\text{X}}$-type constituents on the basis of examples like (3):

(3) a. these [skillful children] and [intelligent parents] in school
     b. these skillful [children in school] and [parents at work]

In examples that have both pre-nominal and post-nominal modifiers, the combination of pre-nominal modifier and noun and the combination of noun and post-nominal modifier can be coordinated. He concludes that—provided one accepts constituency tests—both structures are needed in principle, which would result in spurious ambiguities in noun phrases without coordination. For example, (4) would have two structures: one with *children* combining with *skillful* first and another one with *children* combining with *in school* first.

(4) these skillful children in school

This is indeed unfortunate and in some cases there would not be a difference in meaning but there are cases in which there are two different readings corresponding to the respective bracketings. For instance, (5) can be used to refer to a person that used to be a professor in Tübingen but is living in Stuttgart now (*former* attaches to *professor*) or to a person who was a professor in Stuttgart (*former* attaches to *professor in Stuttgart*).

(5) a former professor in Stuttgart

Section 3.1 discusses another example from Tesnière (2015: 150), for which Tesnière argued that the pre-nominal modifier and the noun form a unit.

So this shows that it is reasonable to assume different structures in principle. This leaves us with spurious ambiguities in cases of intersective modifiers at both sides of the noun. One could argue for a processing strategy that prefers early attachment: while both structures are available in principle, the one with early attachment is preferred. I discussed this analysis in Müller (1999: Section 4.3.1).

Note that approaches with flat structures as in Dependency Grammar would have to find ways of accounting for the ambiguity in (5) in semantics without the help of syntax. They would then face the same problem as the syntactic approaches namely that they have to explain why one of the possible attachment orders in semantics is ignored for intersective modifiers as in (4). Of course, this depends on the actual semantic representation language chosen.

2.3 Non-maximal constituents

Most of the constituent tests apply to maximal phrases only but such languages as German allow for fronting of partial constituents. The following examples show that the ditransitive verb *erzählen* 'tell' can be fronted with either one of the two objects:

(1) a. [Den   Wählern erzählen] sollte man diese   Märchen   nicht.
         the.DAT   voters   tell   should one   this.ACC stories   not
         'One should not tell the voters these stories.'

     b. [Märchen   erzählen] sollte   man   den   Wählern nicht.
         stories.ACC   tell   should one   the.DAT   voters   not
         'One should not tell the voters these stories.'

Groß & Osborne (2009) develop a projective Dependency Grammar, that is, discontinuous constituents are not allowed. Therefore their Dependency Grammar has to license a linguistic object consisting of a ditransitive verb and a dative object (6a) and another one consisting of a ditransitive verb and its accusative object (6b). The assumption of such constituents is standard in Head-Driven Phrase Structure Grammars (HPSG) of German (Pollard 1996, Müller 1996, 2017, Meurers 1999).

3 Evaluating theories

Osborne compares phrase structure trees with dependency trees and claims that Dependency Grammars are simpler since the structures they license involve fewer nodes (Osborne & Groß 2016: 132, Osborne 2018: 2). For example, the left tree in Figure 1 contains eight nodes, while the right one has only four nodes.

3.1 Counting nodes and semantics

Osborne is right: if we count nodes, we see that Dependency Grammars need fewer nodes than Phrase Structure Grammars. But does this mean that they are leaner theories? I think this is not the case. Here is why: Osborne's theories do not discuss the integration with semantics.[5] In fact, they are not even theories. Osborne's papers discuss dependency trees. They do not tell us how
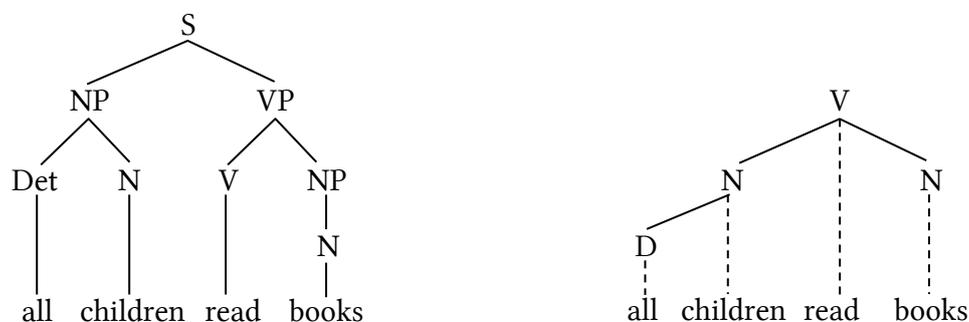

Figure 1. Phrase structure vs. Dependency Grammar analyses

---

[5] See also Dahl (1980) for a similar comment on a discussion note by Richard Hudson.

Figure 2. Tesnière's way of representing scope and the comparison with phrase structure-based analyses by Kahane & Osborne (2015: lix)

these trees are licensed.[6] But let's assume there are some rules of the type suggested by Gaifman (1965: 305) and Hays (1964: 513) licensing Osborne's trees. What we would have then is a syntactic theory. It would predict which word sequences belong to a certain language and which ones do not. But it would not tell us what a certain string means. In order to pair strings with meaning, one has to associate semantic representations to lexical items and to say something about the compositional accumulation of meaning that corresponds to syntactic structure. So, we have to explain that the meaning of *red cars* consists of a combination of the meaning of *red* and the meaning of *car*. Tesnière (2015: 150) noted that simply assuming a flat structure for nominal structures just containing the noun and all adjuncts may be insufficient and more fine-grained distinctions may be necessary. For such cases he suggested using something that Kahane & Osborne (2015: lix) called *polygraphs*. Rather than stating a dependency between a head and an adjunct, one can use polygraphs to state a dependency between a combination of two other items and an adjunct. An analysis using polygraphs is shown in Figure 2 on the left-hand side. As was pointed out by Kahane & Osborne (2015: lix) this polygraph-based analysis is basically equivalent to an $\overline{X}$ analysis, in which adjective and noun form a constituent (an $\overline{N}$) and this constituent is combined with a relative clause. The only difference is that the node gets a name in the $\overline{X}$ variant. (The name also marks which category is the head, something that is indicated in the left figure of Figure 2 by putting *car* above the other material.)

Osborne (p.c. 2016) stated that this was only one of Tesnière's stemmata, one out of 366. But this is irrelevant. The point is that there is a difference that one may want to capture. This is what Tesnière (2015: 150) wrote on his Stemma 149:

§24 Not all subordinates depend as closely on the noun as others. Some of them relate to the noun directly, whereas others are related to it more loosely (cf. Chapter 11, §11–13) and are placed further away. They can be so distant that they seem to depend not on the noun controlling the node but on the node formed by the noun and one of its subordinates.

§25 In order to represent these nuances, which merit a more detailed analysis than can be produced here, it is possible (but not necessary in the standard practice of stemmas) to adopt longer lines for more extended subordinates, and rather than connecting them directly to the node, to connect them to a point along the connection line between the noun and one of its close subordinates. By this process, the phrase *the red car that you saw yesterday* can be analyzed structurally in such a way that the connection line extending upward from the subordinate clause reaches the connection line connecting *red* to *car*. This means that *that you saw yesterday* is connected

---

[6] This question is not trivial. For instance, the tree in Figure 10a (on p. 62 below) could be licensed by six different phrase structure rules such as those in (9) or by two very abstract rules such as Merge and Move (Chomsky 1995) and conventions regarding the labeling of the mother node (Chomsky 2008, 2013).
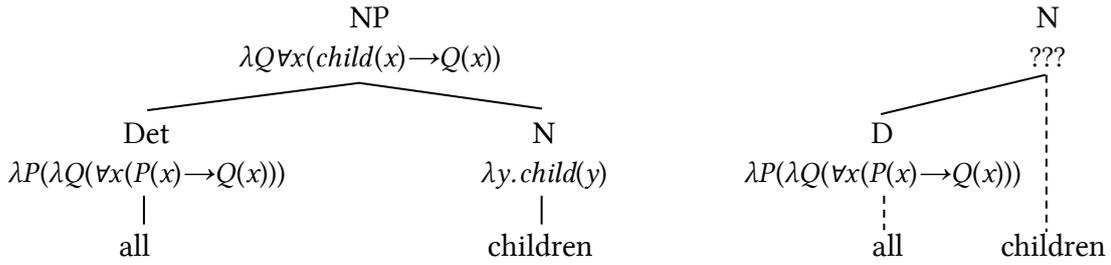
$$NP$$
$$\lambda Q\forall x(child(x)\rightarrow Q(x))$$

$$N$$
$$???$$

$$Det$$
$$\lambda P(\lambda Q(\forall x(P(x)\rightarrow Q(x))))$$

$$N$$
$$\lambda y.child(y)$$

$$D$$
$$\lambda P(\lambda Q(\forall x(P(x)\rightarrow Q(x))))$$

all   children   all   children

Figure 3. Nodes are used for meaning representation: determiner-noun combinations

not to *car* but to *red car*, since what you saw yesterday was not the car, but the red car (Stemma 149). (Tesnière 2015: 150)

Talking about semantics, consider the meaning of *all children*. While phrase structure grammars can represent the combined meaning of *all* and *children* at the NP node as shown in Figure 3, this is not possible in Dependency Grammar (DG). There is just the N and the Det. One could develop theories in which the meaning of the noun *children* is represented by *Qλx:child(x)* and the quantifier is filled into the meaning representation somehow (I use lambda terms here, see Heim & Kratzer (1998, Section 2.5) for an introduction). This would mean that the meaning associated with *children* in the phrase *all children* would be the meaning representation of the whole phrase (see Müller *20*18: Section 11.7.2.3 for details). While this is technically possible, it entails that the meaning of head nouns in NPs including adjuncts would have to include the meaning of the adjuncts as well.[7] So for instance *children* as in *all smart children* would have the meaning of the complete noun phrase (see Figure 4).

Note also that the left tree in Figure 1 above (p. 56) contains a unary branching from NP to N. Such unary branching structures can be used to introduce the semantics that is usually contributed by the determiner (Müller 1999: 284).[8] Again, one can imagine ways of doing this in the lexical item of the noun but this would be non-trivial and/or unintuitive.[9]
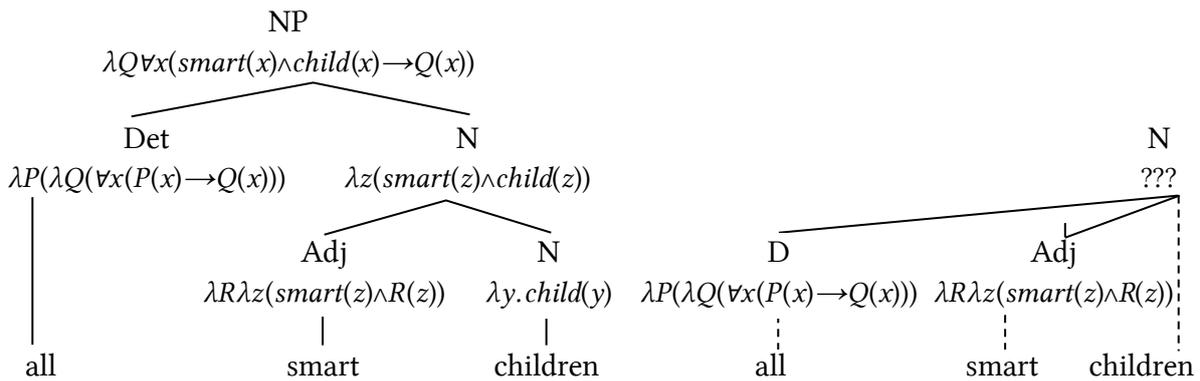
$$NP$$
$$\lambda Q\forall x(smart(x)\wedge child(x)\rightarrow Q(x))$$

$$Det$$
$$\lambda P(\lambda Q(\forall x(P(x)\rightarrow Q(x))))$$

$$N$$
$$\lambda z(smart(z)\wedge child(z))$$

$$N$$
$$???$$

$$Adj$$
$$\lambda R\lambda z(smart(z)\wedge R(z))$$

$$N$$
$$\lambda y.child(y)$$

$$D$$
$$\lambda P(\lambda Q(\forall x(P(x)\rightarrow Q(x))))$$

$$Adj$$
$$\lambda R\lambda z(smart(z)\wedge R(z))$$

all   smart   children   all   smart children

Figure 4. Nodes are used for meaning representation: determiner-adjective-noun combinations

---

[7] Hudson (2003: 391–392) is explicit about this: "In dependency analysis, the dependents modify the head word's meaning, so the latter carries the meaning of the whole phrase. For example, in *long books about linguistics*, the word *books* means 'long books about linguistics' thanks to the modifying effect of the dependents." For a concrete implementation of this idea see Figure 5.

[8] See also Müller (2018: 574, fn. 5) for a discussion of unary projections and semantics.

[9] See for example van Noord & Bouma (1994) for a proposal in HPSG to introduce adjuncts to verbs lexically. In such an approach, all lexical items for verbs are infinitely many times ambiguous. Such an analysis could of course be assumed for nouns as well.
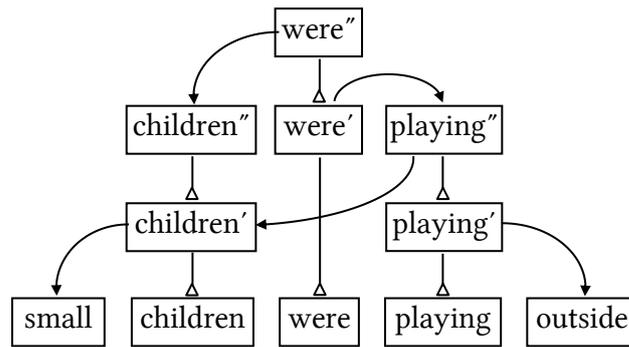
Figure 5. Analysis of *Small children were playing outside.* according to Hudson (2017: 105)
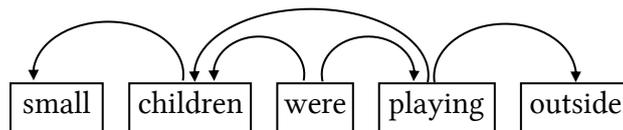

Figure 6. Abbreviated analysis of *Small children were playing outside.* according to Hudson (2017: 105)

There are probably other properties that differ between lexical nouns and the nodes of a complete phrase. A candidate for such a difference is information about nonlocal dependencies (e.g., slash information in HPSG, Pollard & Sag 1994: Chapter 4), but since Osborne's papers are not formalized it is difficult to tell (see Müller (2018: Section 11.5) on a discussion of all the things missing in Groß & Osborne's (2009) account of nonlocal dependencies).

One of the well-formalized versions of DG is Dick Hudson's Word Grammar (Hudson 1991, 2007) and as I show next, it is equivalent to phrase structure grammar in complexity. The representation in Figure 5 is a detailed description of what the abbreviated version in Figure 6 stands for. What is shown in the first diagram is that a combination of two nodes results in a new node. For instance, the combination of *playing* and *outside* yields *playing′*, the combination of *small* and *children* yields *children′*, and the combination of *children′* and *playing′* yields *playing″*. The combination of *were* and *playing″* results in *were′* and the combination of *children″* and *were′* yields *were″*. The only thing left to explain is why there is a node for *children* that is not the result of the combination of two nodes, namely *children″*. The line with the triangle at the bottom stands for default inheritance. That is, the upper node inherits all properties from the lower node by default. Defaults can be overridden, that is, information at the upper node may differ from information at the dominated node. This makes it possible to handle semantics compositionally: nodes that are the result of the combination of two nodes have a semantics that is the combination of the meaning of the two combined nodes. Turning to *children* again, *children′* has the property that it must be adjacent to *playing*, but since the structure is a raising structure in which *children* is raised to the subject of *were*, this property is overwritten in a new instance of *children*, namely *children″*.

The interesting point now is that we get almost a normal phrase structure tree if we replace the words in the diagram in Figure 5 by syntactic categories. The result of the replacement is shown in Figure 7. The only thing unusual in this graph (marked by dashed lines) is that N′ is combined with the *-ing* form of the verb V[*ing*]′ and the mother of N′, namely N″, is combined with the finite verb V[*fin*]′. As explained above, this is due to the analysis of raising in Word Grammar, which involves multiple dependencies between a raised item and its heads. There are
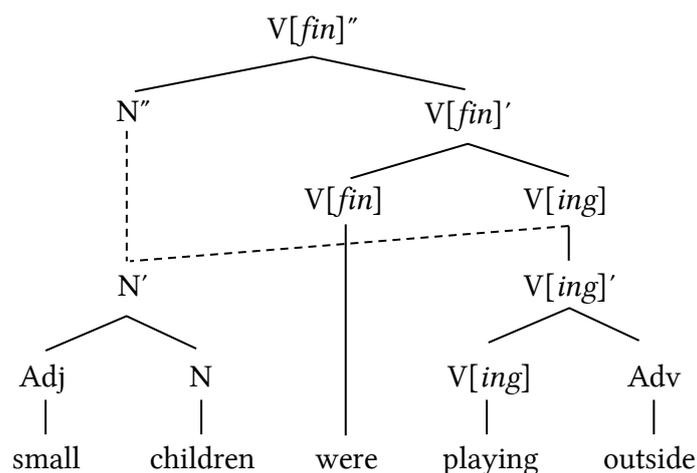
Figure 7. Analysis of *Small children are playing outside.* with category symbols

two N nodes (N′ and N″) in Figure 7 and two instances of *children* in Figure 5. Apart from this, the structure corresponds to what a Phrase Structure Grammar (PSG) would license.

Dick Hudson (p.c. 2018) pointed out to me that his diagram displays is-a relations, while the phrase structure trees visualize consists-of relations. So *playing* is-a *playing′*, while V[*ing*]′ contains-a V[*ing*]. This is an important difference in terms of viewing the domain to be modeled but in terms of complexity the two approaches are at the same level.

3.2 Dependency Grammars vs. Phrase Structure Grammars

Osborne claims again and again that Dependency Grammars are simpler than Phrase Structure Grammars but he does not provide a grammar. If one looks at actual grammar rules as suggested by Gaifman (1965: 305), Hays (1964: 513), Baumgärtner (1970: 61), Heringer (1996: Section 4.1) and others, one sees that the difference is not that large. For instance, Baumgärtner suggests the general rule format in (7):

(2) $\chi \longrightarrow \varphi_1...\varphi_i * \varphi_{i+2}...\varphi_n$, *where* $0 < i \leq n$

The asterisk in (7) corresponds to the central element of the rule, a word of category χ. In our example in Figure 1 ("All children read books", see p. 56 above), χ would be V, the position of the '*' would be taken by *read*, and φ_1 and φ_3 would be N. Together with the rule in (8b) for the determiner-noun combination, the rule in (8a) would license the dependency tree in Figure 1.

(3) a. V → N * N

  b. N → D *

Now compare these rules to phrase structure rules:

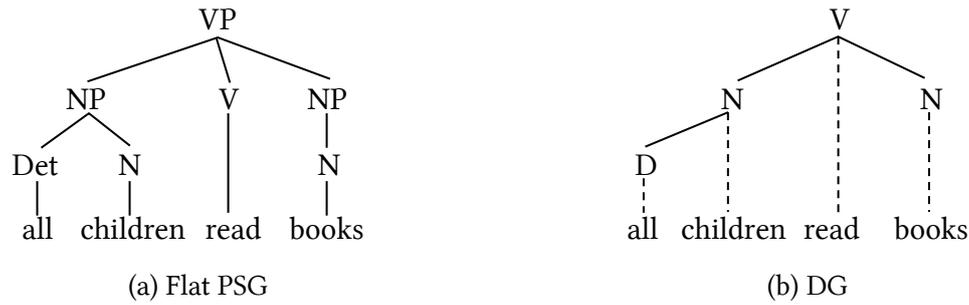(4) a. VP → N V N

  b. NP → D N

(a) Flat PSG (b) DG

Figure 8. Phrase structure vs. Dependency Grammar with rules for combining the same constituents

As is obvious, the difference is not big. The same holds for the tree visualizations of the structures licensed by the respective grammars, which are given in Figure 8.[10] It is clear how the two figures can be related: the dependency tree results from the phrase structure tree by pulling the lexical element up to the XP node (XP stands for NP, AP, PP, or VP or other maximal phrasal projections) and removing the XP node (Gaifman 1965, Covington 1990: 234, Oliva 2003, Hellwig 2006: 1093).

So, phrase structure grammarians could assume flat structures for sentences as they are assumed in DG. Why do they nevertheless stick to intermediate VP nodes? They could assume totally flat noun phrases having one node for NP dominating the determiner, the head noun, arguments of the head nouns and all adjuncts. Some phrase structure grammarians actually do assume such structures (Eisenberg 2004: 29, 144), but most do not. One reason for assuming intermediate nodes is that they serve as attachment sides for adjuncts. As (10) shows for the nominal domain, it would not be sensible to pose an upper limit on the number of possible adjuncts:

(5) a. A: All leather chairs are comfortable.
b. B: No, I know of one uncomfortable leather chair.
c. A: But all uncomfortable leather chairs are cheap.
d. B: No, I know of one expensive uncomfortable leather chair.

The same is true in the verbal domain as discussed in Frey & Gärtner (2002: 48). If one assumes a recursive rule like the one in (11), the NPs in (11) are accounted for:

(6) N′ → A N′

The computation of the semantic contribution is straightforward: the meaning of the adjective is paired with the meaning of the N′ that is modified. The result of this combination is represented at the resulting N′ node. Assuming these nodes also accounts for order: adjectives precede nouns but follow the determiner. It is clear that one can account for such examples as (10) with flat rules but this involves rules with arbitrarily many daughters.

---

[10] An additional rule for the projection of N to NP is needed in the phrase structure grammar:

NP → N

This rule introduces the determiner's semantics. See Section 3.1.

(a) PSG analysis with adjunct

(b) Totally flat DG analysis according to Osborne (2018: 24)
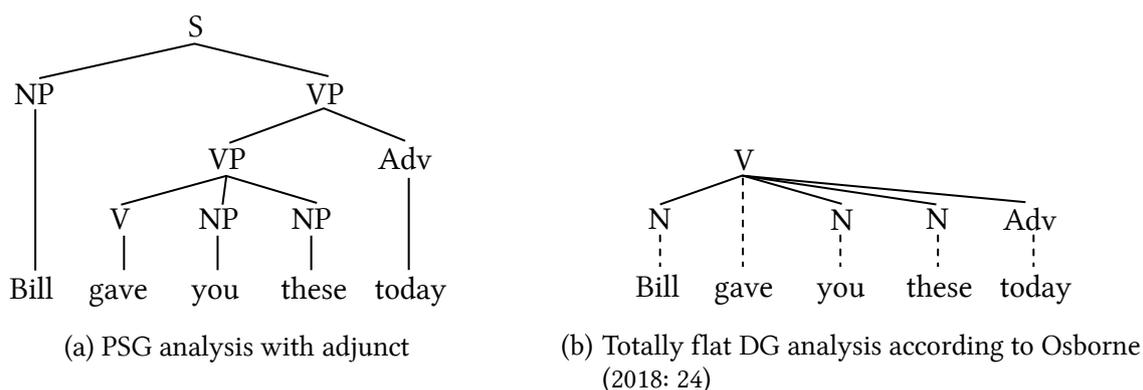
Figure 9. Phrase structure vs. Dependency Grammar with adjuncts

The situation with VPs (in SVO languages, such as English) is similar (see Figure 9): the VP is a certain well-defined area of the clause and adjuncts may attach to the left and to the right of it. The subject has to be placed in front of the VP not, say, somewhere between the objects. All this is accounted for in phrase structure grammars that acknowledge a VP node and recursive rules adding adverbial elements to the VP. Again, flat rules could be assumed for clauses but accounting for linearization of constituents and in particular the scope of adjuncts is not trivial. The computation of adjunct scope will be complex and involves machinery that is not needed in approaches with more structure. See Kasper (1994) for a proposal with flat structures in HPSG.

Linearization constraints are hardly ever provided in DG work. I could not find any statement regarding linearization within phrases in Osborne's work. Until this work is done, it is not justified to claim that DG is simpler since the theories are not comparable. For an extended discussion of these topics in general, and some of Osborne's analyses of German in particular, see the chapter on Dependency Grammar in Müller (2018).

3.3 Rules vs. lexical elements

During a discussion as part of a workshop on long-distance dependencies at the Humboldt-Universität zu Berlin in 2018, Gereon Müller claimed that lexical elements do not count when it comes to the evaluation of syntactic theories. According to him, the only thing that matters is the number of rules. If this was a valid criterion for evaluating theories, a theory assuming the structure in Figure 10a for a prepositional phrase would be better than one assuming the right structure, provided the left structure is licensed by very general rules as it is the case in Minimalism (Chomsky 1995). Radford's (1997) structure, represented in Figure 10a, involves several empty elements, two special part-of-speech labels (p for an additional prepositional category, and AgrO for something having to do with object agreement) and movements and is way more complex than the rather trivial combination of a preposition with its NP dependent in Figure 10b. Nobody outside the Minimalist community would accept the claim that the theory behind the structure on the left-hand side is simpler than others. In fact, such researchers as Wolfgang Sternefeld and Hubert Haider broke up with their school because of analyses like the one in Figure 10a and arguments for analyses similar to the one in Figure 10a (Sternefeld 2006: 549–550, Sternefeld & Richter 2012, Haider 2018).
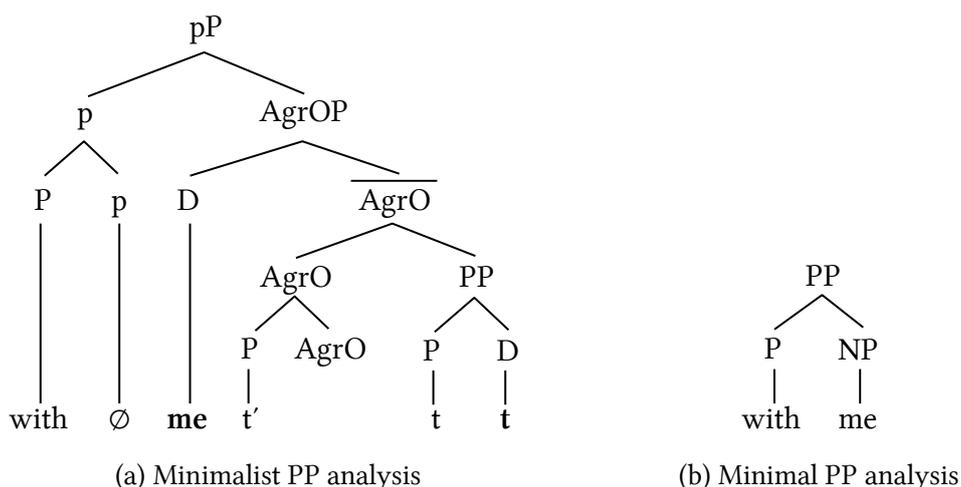
(a) Minimalist PP analysis   (b) Minimal PP analysis

Figure 10. Analysis of prepositional structures according to Radford (1997: 452) and conventional PP analysis

So, to evaluate theories with respect to their complexity, one has to count the distinctions that have to be made: the number of features, the number of their possible values, the constraints on identities of values, the number of rules, the number of lexical entries assumed. The number of lexical entries for normal words like *book* will not tell us much about the parsimony of theories since all theories have to assume a lexical item for *book* but not all theories assume empty elements.[11] Sometimes the empty elements are eliminated by assuming unary branching structures or additional phrasal rules (Müller 2016: Chapter 19). So there is a tradeoff that has to be taken into account when comparing theories.

When comparing theories, it is also important to look at the complexity of the lexicon. Such theories as Construction Grammar of the Goldberg type (Goldberg, 1995, 2006) assume a lot of phrasal schemata while lexical theories assume respective lexical items. For example, Goldberg & Jackendoff (2004) assume a phrase structure rule for resultative constructions while lexical approaches assume lexical rules (Wechsler & Noh 2001, Müller 2002, 2006, Müller & Wechsler 2014). While theories like Lexical Functional Grammar (LFG) and HPSG have elaborated theories of the lexicon, such subtheories are lacking from Osborne's DG and from Minimalism (Newmeyer 2005: 95, fn. 9). See also Borsley & Müller (2019) on the lexicon in Minimalism.

Before concluding the paper, I want to add a caveat: the question which theory is most appropriate depends on what we want to model and what the empirical domain is. If we want our theories to be rather directly related to human behavior and the way things are represented in the brain, developing the most compact and non-redundant theory may not result in the most appropriate theory. The reason is that a lot of information is stored redundantly in the brain. Patterns with high frequency are stored even though their structure follows general rules and they could be analyzed compositionally from their parts (Bybee 2006, Bannard & Matthews 2008).

---

[11] The situation is more complicated even for visible lexical items. Some theories introduce adjuncts lexically and hence assume that there are infinitely many possible lexical items for verbs. See van Noord & Bouma (1994) and Bouma, Malouf & Sag (2001).

4 Conclusion

I have shown that Osborne (2018) misunderstood the definitions of constituency tests and that the tests are unproblematic when they cannot apply. I further argued that his proposals lack an integration of semantics and if semantics is taken into account, some Dependency Grammar approaches get even closer to phrase structure-based ones than they already are. I discussed some other phenomena, such as constituent order and adjunct attachment interacting with intermediate phrasal nodes.

As for theory evaluation in general, I argued that rather than looking at the rules only as in Minimalism or looking at the nodes in the syntactic tree only as in Dependency Grammar, we have to look at the number of theoretical entities and mechanisms as a whole. We also have to look at the bigger picture and think about the integration of semantics. After all, doing syntax in isolation is fun but pointless: syntax is the bridge between form (sound, writing or signing) and meaning.

Acknowledgements

References

Adger, David. 2003. *Core syntax: A Minimalist approach* (Oxford Core Linguistics 1). Oxford: Oxford University Press.

Bannard, Colin & Danielle Matthews. 2008. Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science 19*(3). 241–248.

Baumgärtner, Klaus. 1970. Konstituenz und Dependenz: Zur Integration beider grammatischer Prinzipien. In Hugo Steger (ed.), *Vorschläge für eine strukturelle Grammatik des Deutschen* (Wege der Forschung 144), 52–77. Darmstadt: Wissenschaftliche Buchgesellschaft.

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 48–69. Stanford, CA: CSLI Publications.

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

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

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4).

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

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.

Covington, Michael A. 1990. Parsing discontinuous constituents in Dependency Grammar. *Computational Linguistics* 16(4). 234–236.

Crysmann, Berthold. 2008. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2003, Vienna, Austria*, 47–62. Stanford, CA: CSLI Publications.

Dahl, Östen. 1980. Some arguments for higher nodes in syntax: A reply to Hudson's 'constituency and dependency'. *Linguistics* 18(5–6). 485–488.

Duden. 2005. *Duden: Die Grammatik*. 7th edn. Vol. 4. Mannheim, Leipzig, Wien, Zürich: Dudenverlag.

Dürscheid, Christa. 2003. *Syntax: Grundlagen und Theorien*. 2nd edn. (Studienbücher zur Linguistik 3). Westdeutscher Verlag.

Eisenberg, Peter. 2004. *Grundriß der deutschen Grammatik*. 2nd edn. Vol. 2. Der Satz. Stuttgart, Weimar: Verlag J. B. Metzler.

Eroms, Hans-Werner. 2000. *Syntax der deutschen Sprache* (de Gruyter Studienbuch). Berlin: Walter de Gruyter Verlag.

Flohr, Horst & Henning Lobin. 2009. Die Struktur von Sätzen: Syntax. In Horst M. Müller (ed.), *Arbeitsbuch Linguistik*, 2nd edn. (UTB 2169), 125–147. Paderborn: Ferdinand Schöningh Verlag.

Frey, Werner & Hans-Martin Gärtner. 2002. On the treatment of scrambling and adjunction in Minimalist Grammars. In Gerhard Jäger, Paola Monachesi, Gerald Penn Shuly Wintner (eds.), *Proceedings of Formal Grammar 2002*, 41–52. Trento.

Gaifman, Haim. 1965. Dependency systems and phrase-structure systems. *Information and Control* 8(3). 304–397.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. (Cognitive Theory of Language and Culture). Chicago/London: University of Chicago Press.

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

Goldberg, Adele E. & Ray S. Jackendoff. 2004. The English resultative as a family of Constructions. *Language* 80(3). 532–568.

Goodall, Grant. 1983. A three-dimensional analysis of coordination. In *Papers from the nineteenth regional meeting of the Chicago Linguistics Society, Chicago 21–22 April 1983*. 146–154.

Groß, Thomas M. & Timothy Osborne. 2009. Toward a practical Dependency Grammar theory of discontinuities. *SKY Journal of Linguistics* 22. 43–90.

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

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

Hays, David G. 1964. Dependency Theory: A formalism and some observations. *Language* 40(4). 511–525.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. (Blackwell Textbooks in Linguistics 13). Oxford: Blackwell Publishers Ltd.

Hellwig, Peter. 2006. Parsing with Dependency Grammars. In Vilmos Ágel, Ludwig M. Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.), *Dependenz und Valenz / Dependency and valency: Ein internationales Handbuch der zeitgenössischen Forschung / An international handbook of contemporary research*, vol. 25.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1081–1109. Berlin: Walter de Gruyter.

Heringer, Hans-Jürgen. 1996. *Deutsche Syntax dependentiell* (Stauffenburg Linguistik). Tübingen: Stauffenburg Verlag.

Hudson, Richard. 1988. Coordination and grammatical relations. *Journal of Linguistics* 24(2). 303–342.

Hudson, Richard. 1991. *English Word Grammar*. Oxford: Basil Blackwell.

Hudson, Richard. 2003. Mismatches in default inheritance. In Elaine J. Francis & Laura A. Michaelis (eds.), *Mismatch: form-function incongruity and the architecture of grammar* (CSLI Lecture Notes 163), 355–402. Stanford, CA: CSLI Publications.

Hudson, Richard. 2007. *Language networks: The new Word Grammar*. Oxford: Oxford University Press.

Hudson, Richard. 2017. Pied piping in cognition. *Journal of Linguistics* 54(1). 85–138.

Jackendoff, Ray S. 1977. *X syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Kahane, Sylvain & Timothy Osborne. 2015. Translators' introduction. In Lucien Tesnière, *Elements of structural syntax*, xxix–lxxiii. Translated by Timothy Osborne and Sylvain Kahane. Amsterdam: John Benjamins Publishing Co.

Kasper, Robert T. 1994. Adjuncts in the Mittelfeld. In John Nerbonne, Klaus Netter Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 39–70. Stanford, CA: CSLI Publications.

Machicao y Priemer, Antonio. 2018. Permutationstest. In Stefan Schierholz & Pál Uzonyi (eds.), *Grammatik: Syntax* (Wörterbücher zur Sprach- und Kommunikationswissenschaft (Online) 1.2). Berlin: De Gruyter.

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.

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

Müller, Stefan. 1999. *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. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62.

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

Müller, Stefan. 2006. Phrasal or lexical Constructions? *Language* 82(4). 850–883.

Müller, Stefan. 2016. *Grammatical theory: From Transformational Grammar to constraint-based approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press.

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

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

Müller, Stefan & Stephen Mark Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76.

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

Oliva, Karel. 2003. Dependency, valency and Head-Driven Phrase-Structure Grammar. In Vilmos Ágel, Ludwig M. Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.), *Dependenz und Valenz / Dependency and valency: Ein internationales Handbuch der zeitgenössischen Forschung / An international handbook of contemporary research*, vol. 25.1 (Handbücher zur Sprach- und Kommunikationswissenschaft), 660–668. Berlin: Walter de Gruyter.

Osborne, Timothy. 2006. Shared material and grammar: Toward a Dependency Grammar theory of non-gapping coordination for English and German. *Zeitschrift für Sprachwissenschaft* 25(1). 39–93.

Osborne, Timothy. 2018. Tests for constituents: What they really reveal about the nature of syntactic structure. *Language Under Discussion* 5(1). 1–41.

Osborne, Timothy & Thomas Michael Groß. 2016. The *Do-So*-Diagnostic: Against finite VPs and for flat non-finite VPs. *Folia Linguistica* 50(1). 97–135.

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

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

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

Remberger, Eva. 2016. Vorfeldtest. In Stefan Schierholz & Pál Uzonyi (eds.), *Grammatik: Syntax* (Wörterbücher zur Sprach- und Kommunikationswissenschaft (Online) 1.2). Berlin: De Gruyter.

Schäfer, Roland. 2015. *Einführung in die grammatische Beschreibung des Deutschen* (Textbooks in Language Sciences 2). Berlin: Language Science Press.

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.

Tesnière, Lucien. 2015. *Elements of structural syntax*. Translated by Timothy Osborne and Sylvain Kahane. Amsterdam: John Benjamins Publishing Co.

van Noord, Gertjan & Gosse Bouma. 1994. The scope of adjuncts and the processing of lexical rules. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 250–256. Kyoto, Japan: Association for Computational Linguistics.

von Stechow, Arnim & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens: Ein Lehrbuch der Generativen Grammatik*. Opladen/Wiesbaden: Westdeutscher Verlag.

Wechsler, Stephen Mark & Bokyung Noh. 2001. On resultative predicates and clauses: Parallels between Korean and English. *Language Sciences* 23(4). 391–423.

Wöllstein, Angelika. 2010. *Topologisches Satzmodell* (Kurze Einführungen in die Germanistische Linguistik 8). Heidelberg: Universitätsverlag Winter.