Dependency or Phrase Structure Grammars

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

August 29, 2026

It is often claimed that Dependency Grammar (DG) is simpler than phrase structure grammars or even that it is "the simplest theory of grammar" (Gibson 2026). This paper shows that DGs appear simpler since many aspects of a formal theory are simply not provided. Many DG papers come without a rule component. I discuss the rules that Gibson provides in a syntax textbook and show that they are formally not sound and that they mimic phrase structure-based systems in their intention. Treating semantics requires more structure either in terms of nodes or in terms of additional features. Headless constructions like Jackendoff's NPN construction cannot be handled in DG without extensions or empty elements. If projectivity is given up, as suggested by Gibson (2025), other means are needed to avoid overgeneration, again adding to complexity. Some notion of constituency is needed to ensure the necessary adjacency requirements. I discuss examples from German that suggest that V2 should be treated by some kind of "movement" rather than by simple dependency relations. It is shown that psycholinguistic arguments (e.g. Dependency Locality) cannot be used to argue against phrase structure-based approaches.

Keywords: Dependency Grammar, Syntax, theory comparison, psycholinguistics, Construction Grammar

Word count: 10883 words

update

1 Introduction

Gibson (2026) argues in his paper *Dependency syntax as the simplest theory of grammar* against phrase structure grammars and in particular against transformational theories.

His main point is the alleged simplicity of Dependency Grammar. What I want to show in this paper is that his claim is not warranted: if there is a given set of phenomena and properties of language that have to be described, it is easy to isolate certain aspects and only display these. This looks rather simple in comparison to theories that take more properties into account. To take a concrete example, it is easy to draw dependency graphs for a given utterance, but a grammar has more tasks than to just provide structures for grammatical utterances. It also has to rule out ungrammatical word sequences. This is especially important since grammars are used for both parsing input and generating output. As I will show below, it is a non-trivial task to incorporate the constraints necessary for generation into (simple) Dependency Grammars.

To take a concrete example, have a look at the attested German sentence in (1):

(1) Stets einen wachen Blick mußten Prinz Claus II. und seine Begleiter auf das
always a woke glimps must Prinz Claus II. and his companions at the
lange Herrschaftssymbol an der Mütze haben, damit es die sich blitzschnell
long sovereignty.symbol at the cap have so.that it the REFL lightning-fast
drehenden Walzen des 1300 Tonnen schweren Kolosses ja nicht
spinning rollers of.the 1,300 metric-ton heavy colossus PART not
verschlingen.[1]
swallow.up
'Prince Claus II and his companions had to keep a close eye on the long symbol of sovereignty attached to his cap at all times, to ensure that it would not be swallowed up by the lightning-fast spinning rollers of the 1,300 metric-ton colossus.'

The sentence involves the phenomenon of multiple frontings (van de Velde 1978, Müller 2003, Bildhauer 2011), which are possible in German, although German is a V2 language (Holmberg 2015), that is, usually there is exactly one constituent before the finite verb. The adverb *stets* 'always' is an adjunct of *haben* 'have' and *einen wachen Blick* is a part of the accusative object of *haben* 'have'. The remaining part of the object (*auf das lange Herrschaftssymbol an der Mütze* 'on the long symbol of sovereignty attached to his cap') is realized to the right of the finite verb. Semantically the adverb and the part of the accusative object are not connected. Both phrases depend on the verb *haben* 'have'. If dependencies are analyzed without any further analytical devices, we get the dependency graph in Figure 1 for the relevant subpart of the sentence.[2]

It depends on further assumptions in the theory whether such graphs are possible at all. For example, some Dependency Grammarians assume projectivity (e.g. Hudson 2000: 23, Groß & Osborne 2009: 53). That is, crossing lines as in Figure 1 are forbidden. So, depending on basic assumptions that are made, there is either no worked out theory of apparently multiple frontings or the theory vastly overgenerates, since it allows all kinds of non-projective structures.

A further thing that is usually missing in papers about DG (and also in papers about Government & Binding and Minimalism) is semantics. As Müller (2019) pointed out, once semantics is integrated, DG analyses look suspiciously similar to phrase structure-based analyses.

[1] Mannheimer Morgen, 03.12.2005, Ressort: Mannheim; Beim Salto Mortale herrscht Spannung pur)

[2] I analyzed the subject *sie* 'they' as a dependent of the modal verb *mussten* 'must' as it is done by Heringer (1996: 73). Alternatively it could be analyzed as dependent of *haben* 'have', which would result in one further crossing line. See Eroms (2000: Section 5.3.4) for discussion.

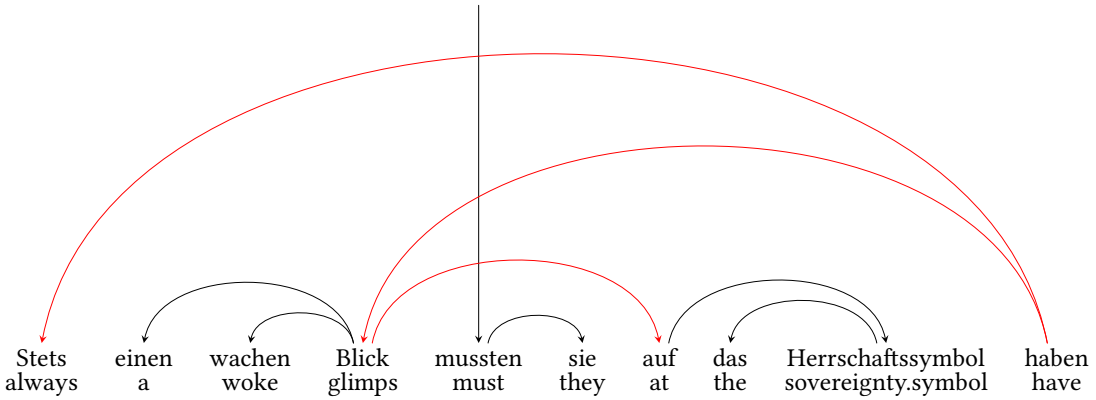


Figure 1: Analysis of an attested German example involving apparently multiple fronting together with NP split requiring crossing dependencies

Gibson (2026) raises various issues in his paper and I comment on many of his points in what follows. I also discuss some fundamental points related to his 2025 book. The paper is structured as follows: Section 2 discusses the fact that dependency trees do not constitute a theory. A grammar is needed to license the trees. If one looks at actual suggestions it is clear that they are not very different from phrase structure grammars. Section 3 shows that phrase structure grammar nodes are used to model certain kinds of recursive structures, which have to be modeled in different ways in DGs. Section 4 argues that some notion of unit is needed in DGs as well and that complaints about malfunctioning constituency tests are due to logical mistakes in the criticism. Section 5 discusses Gibson's own proposal, shows formal problems and points out problems related to the general Goldbergian approach to arguments structure constructions. Section 6 deals with claims about the alleged simplicity of DG. It shows that if semantics is integrated additional representational features or nodes are needed that make the alleged advantage of DG disappear. Section 7 discusses the problem that DG has with unheaded constructions. Section 8 points out problems with discontinuity. If adjacency between heads and dependents is enforced, grammars result that are rather similar to phrase structure grammars. Section 9 discusses the V2 property of German and shows that some sort of long distance dependency is needed even for simple V2 sentences. The problem of apparently multiple frontings is discussed and the example in (1) is taken up again. Section 10 deals with transformations. Gibson refers to studies showing that transformations cannot be psycholinguistically real and that language acquisition problems disappear if transformations are abandoned. I point out in this section that there have been representational models in Transformational Grammar that did not have the psycholinguistic problems. Section 11 deals with computational complexity of transformational accounts and the significance of complexity results in general. Section 12 discusses Gibson's claims regarding dependency locality and shows that the results are compatible with extraction accounts using the GPSG slash mechanism. Section 13 concludes the paper.

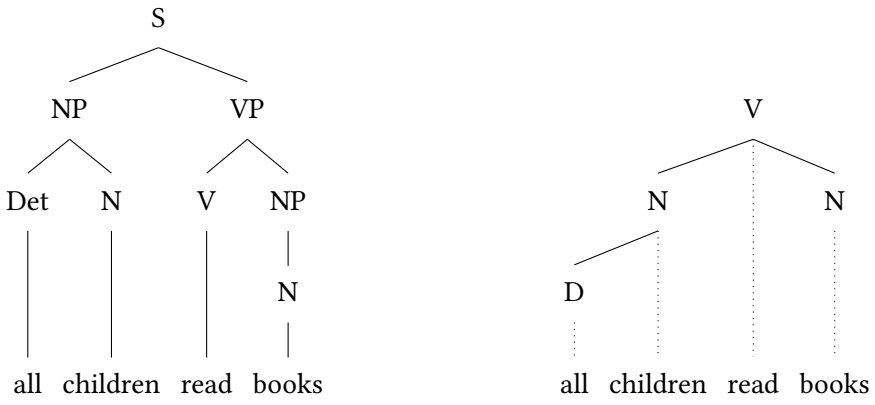


Figure 2: Phrase structure vs. Dependency Grammar analyses

2 Constituents vs. dependents and the missing grammar of Dependency Grammar

I already mentioned that drawing some trees or dependency graphs is not a linguistic theory although this is what is commonly done in theoretical work on Dependency Grammar and also in Mainstream Generative Grammar (MGG). There are some researchers, however, who do provide grammar rules (Gaifman 1965: 305, Hays 1964: 513, Baumgärtner 1970: 61, Heringer 1996: Section 4.1, Gibson 2025: Chapter 3). For instance Baumgärtner suggests the general rule format in (2):

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

The asterisk in (2) corresponds to the word of category χ. In the example in Figure 2, there are two uses of the rule. In one, χ would be V, the position of the '$*$' would be taken by *read*, and φ_1 and φ_3 would be N. Alternatively, χ can be N, the position of the '*' would be taken by *children* and φ_1 would be D. The two rules in (3) license the dependency tree in Figure 2.[3,4] (3a) combines a verb with two NPs and the rule in (3b) can be used for determiner-noun combinations:

(3) a. V → N * N
b. N → D *

Now compare these rules to phrase structure rules:[5]

(4) a. VP → NP V NP
b. NP → D N

[3]The dependency tree is displayed in a format that is common among German Dependency Grammarians (e.g., Eroms 2000). I use it in the following to make the comparison simpler.

[4]Gibson's own rules are discussed in Section 5 below.

[5]I assume a phrase structure grammar to be a quadruple consisting of non-terminal symbols, terminal symbols, rewrite rules and a start symbol. See Chomsky & Miller (1963: Section 4). Many modern grammar formalisms can be seen as variants of such simple phrase structure grammars.

They are completely parallel under the assumption that V is the head of VP and N is the head of NP. Phrase structure grammars usually do not assume completely flat representations for languages like English. They assume that the verb is combined with the objects to form a VP, which is then combined with the subject to form a complete sentence. The advantage of this analysis is that it is easy to account for adjuncts between the verb and the subject. The respective rules can be formulated recursively as in (5), which is not possible in DGs.

(5) a. VP → Adv VP
b. VP → VP Adv

Recursive rules of this kind are the topic of the next section.

3 Recursion

If adjuncts are to be covered in rules like (3) one would need regular expressions to allow for arbitrarily many adjuncts. Huddleston & Pullum (2002: 780) provide the following examples for multiple occurrences of adjuncts.

(6) a. For this reason, as soon as the meeting was over, he called his solicitor.
b. He probably deeply regretted having agreed to take part.
c. She left immediately in order to catch the early train.

The following data from Müller (2023c: 67) shows that there can be arbitrarily many adjuncts in NPs. One can add arbitrarily many adjectives and make the set of referents smaller.

(7) A: All grey squirrels are big.
B: No, I saw a small grey squirrel.
A: Yes, but small grey squirrels are ill.
B: No, I saw a healthy small grey squirrel.

There is an interesting way to generalize over the rules in (3) and (4): If dependents are represented in lists as was suggested by Bouma, Malouf & Sag (2001), one rule is sufficient:

(8) H → 1, H[DEPS 1 ⊕ 2], 2

The ⊕ stands for `append`, a relational constraint concatenating two lists. In (8), the DEPS list is split into 1 and 2, which can be realized before or after the head. This rule can be compared with (2): the H on the left-hand side is the χ, the H on the right-hand side is the '*', and the elements of 1 are the $\varphi_1 \ldots \varphi_i$ and those of 2 are $\varphi_{i+2} \ldots \varphi_n$.

Bouma et al. suggest a lexical introduction of adjuncts. They have a lexical representation of arguments in a list called argument structure list (ARG-ST) and then use the feature DEPS to represent all dependents of a head. The DEPS list is a list that contains all the elements of the ARG-ST list plus a list of arbitrary length containing adjuncts (Bouma et al. 2001: 42). So, both 1 and 2 can contain arbitrarily many elements.

The interesting thing to note here is that rules with regular expressions (as assumed by Gibson, see below) stand for infinitely many different rules.[6] The lexical items of Bouma et al. (2001: 42) admit arbitrarily many adjuncts in the DEPS list. If one would compute all possible lexical items, one would get an infinite lexicon. In comparison, phrase structure grammars have recursion in their rules.[7] For example, an adjective is combined with an $\overline{\text{N}}$ and the result is an $\overline{\text{N}}$. This can be repeated as often as it is required.

(9) $\overline{\text{N}} \rightarrow \text{Adj}\ \overline{\text{N}}$

If proponents of phrase structure grammar would want to work like Dependency Grammarians, they could do this and assume rules like (8), but apart from the issue with recursive structures, there are various reasons to set up grammars differently. Before I go into details, I want to point out one more thing: In order to develop a grammar like the one in (8), one has to identify the parts of a sentence and group words into larger units.

4 Constituents or heads with dependents

Both Dependency Grammarians and Phrase Structure Grammarians have to talk about units that are combined. The task of identifying these building blocks is the same for phrase structure grammars and for dependency grammars. A common mistake made by DGers when critizising constituent tests is that they ignore the exact formulation of the tests (Gibson 2026: 4, Osborne 2018). The tests say: If a sequence of words has property X, then it is (likely that the sequence is) a constituent. This means that if the sequence does not have the property X, the test does not fail. It is simply not applicable. So, for example, Gibson – referring to Osborne (2018) – notes that some PPs cannot be clefted although they are treated as constituents. The respective phrases do not "fail constituency tests" as claimed by Gibson. The test cannot be applied. But this is not a problem as long as there is other evidence for constituent status. See Müller (2023c: Section 1.3) for more discussion of constituency tests including problematic aspects. It is true that PSGs contain more nodes than most DGs, but there are reasons for assuming them, as will be discussed below. And some of the additional nodes are motivated by the constituency tests. If DG does not assume these tests, they have to come up with ad hoc solutions to the phenomena, not based on tests.

Note also that some of the Dependency Grammars are not projective. That is, they allow for crossing dependency arcs (e.g. Heringer 1996: 261, Eroms 2000: Section 9.6.2, Gibson 2025: 112, 137). Again, if this is wanted, the same can be achieved in a Phrase Structure Grammar setting. There were researchers allowing for discontinuous constituents (Zwicky 1986, Reape 1994, Kathol 2001, Müller 1999, 2004a). Within HPSG these models have been abandoned since they have problematic aspects (Müller 2024a: Section 6.4). And in any case, one has to make sure that material that belongs together is realized together (Behaghel 1909: 139) and hence one has the notion of constituent back in.

[6] See also the adjunct addition metarule of GPSG (Uszkoreit 1987: 146, Weisweber & Preuss 1992). This metarule added one adjunct at a time to flat phrase structure rules. By applying it recursively infinitely many phrase structure rules are licensed.

[7] Recursion can arise in Dependency Grammars as well. For example, nouns can govern other nouns and verbs can take sentential arguments.

Dependency information is available in phrase structure grammars: As Figure 3 shows, the information about dependencies is recoverable. It is clear that the N *children* of *all little children* would be the highest element in a Dependency Grammar. It is the head of the noun phrase, since its head information is shared in the substructure. The determiner is a dependent of *children* since it appears in its valence list, and the adjective also depends on the noun, since its MOD value is identical to (a projection of) the noun. The same works for the verbal projection: The

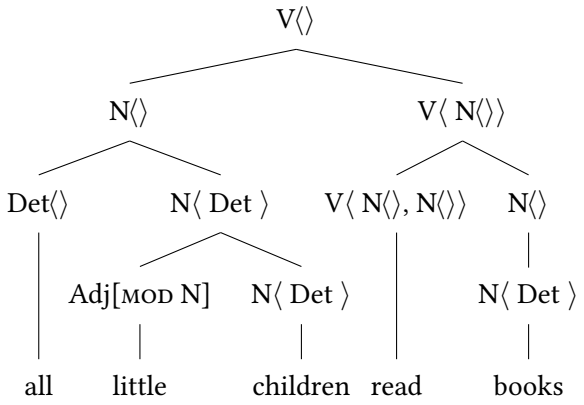


Figure 3: Dependency information is contained in (certain) phrase structure grammars

head of the complete sentence is *read*, since head information of *read* (in the figure just part of speech) is shared along the head path. The transitive verb selects two items, so it can be determined that *all little children* and *books* are dependents of *read*. The dependency information is contained in HPSG-like phrase structures and if it plays a role in psycholinguistic results, this is a good property of theories like HPSG. See Section 12 for a discussion of Dependency Locality.

5 Gibson's rules

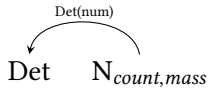
In comparison to his 2026 paper, Gibson's introduction to syntax (Gibson 2025: Chapter 3) is more explicit about the rules that are assumed. Gibson suggests for every type of head a grammar rule. So there are rules for intransitive verbs, for strictly transitive verbs, for ditransitive verbs, verbs with sentential objects and so on. There are rules combining nouns with determiners, rules for combining nouns with determiners and adjectives and for combining nouns with prepositional phrases. Gibson (2025: Section 3.14) justifies these rules by pointing to work by Goldberg (2006, 2019). He claims like proponents of phrasal Construction Grammar that the rules contribute meaning. This of course raises all the questions and causes almost all the problems of such proposals that were discussed earlier (Müller 2006, 2013, 2018, Müller & Wechsler 2014a,b).

Gibson (2025: Section 3.10) explains compactness and non-redundancy. So it comes with some surprise that he suggests many many rules although only one abstract rule of the kind given in (8) would be necessary since all the relevant information about valence is represented

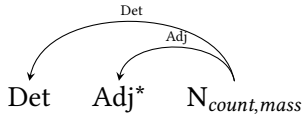
in his lexical items already.

In addition there are some interesting facts to note about the rules Gibson suggested. Gibson shows the two dependency graphs in the Figures 4 and 5 and states that they are derived with the rules in (10).

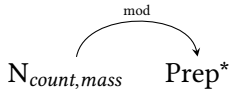
(10) a. Rule 3, p. 58:



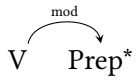
b. Rule for adjectives, p. 67:



c. Rule for prepositions modifying nouns, p. 68:



d. Rule for prepositions modifying verbs, p. 68:



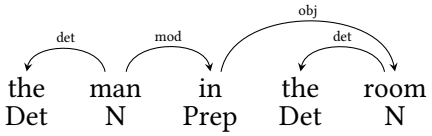


Figure 4: Dependency graphs containing adjuncts in the nominal domain according to Gibson (2025: 68)

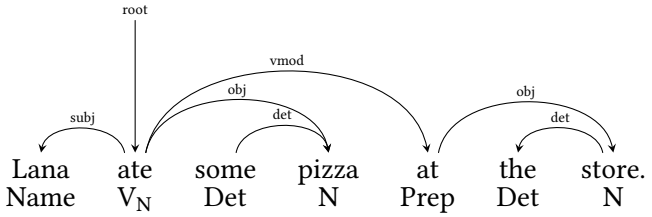


Figure 5: Dependency graphs containing adjuncts in the verbal domain according to Gibson (2025: 68)

In order to be able to analyze the sequence in Figure 4 with the rules in (10b) and (10c), Gibson has to assume that heads in rules may be used more than once. This is fundamentally

different from the rule systems, I talked about earlier. This can be seen maybe more clearly when Gibson's rules are reformulated into the format that was given above.

(11) a. N → Det Adj* *
b. N → * Prep*

The asterisk marks the position of the head, but there is no other slot for N. Either N is generated with (11a) or with (11b). It is impossible that both rules are used. Gibson did not provide details about the formalization of his rule system. Of course one can come up formalizations that allow this kind of rule interaction, but Gibson left this to the imagination of the reader. Przepiórkowski (2026: Section 5.1) came up with one interpretation. He pointed out that multiple applications of rules like the one in (10a) would permit the combination of a noun with arbitrarily many determiners. He then suggests, that each rule may be applied at most once in an application combining several rules. So, in the analysis in Figure 4 the two rules in (10a) and (10c) are applied in combination. This would fix many problems, but note that in grammars of German adjuncts can appear anywhere between arguments (see (14) below). This means that the rule for adjunct insertion may apply several times. What could be requested though is some sort of functional uniqueness as it is assumed in LFG. If this were to be assumed there could be at most on determiner, at most one object of a preposition and so on. Of course, adjuncts would have to be exempt from this uniqueness requirement (as in LFG).

Note also that the analysis in Figure 5 requires that the elements in the rule in (10d) combining a verb with a PP must allow for a discontinuous realization of V and PP. Otherwise it cannot be explained why *some pizza* can appear between the verb and the PP. But if such discontinuities are allowed, why does rule (10b) list all parts of the NP rather than just stating that adjectives can be combined with nouns? Rule (10a) could then license a discontinuous NP with a hole between Det and N and the adjective would be inserted by a rule that is parallel to (10c).

This brief discussion shows, that a lot more has to be said about Gibson's rule system, which also shows that his approach is less simple then what he claimed.

There is a further interesting aspect of rule (10c): if Gibson finds a way to allow multiple uses of head nodes, the Kleene star results into problems. First, Kleene star means zero or more occurrences of a symbol. This means that the rule in (10c) could just contain the N and no PP. So it could be applied to N arbitrarily often, making all nominal structures infinitely ambiguous. In phrase structures this would correspond to arbitrarily many applications of a unary branching rule. Gibson could go for the '+' instead of the Kleene star. This would require at least one PP if the rule is applied. But even this is problematic, since an NP with two PP adjuncts could then be analyzed as a noun with two PPs, that is one rule application, or as a double application of the rule adding one PP at the time. The problem worsens the more PPs are added to the noun. Note that this problem does not exist in phrase structure grammars since the result of the combination of an N with PPs can be of a different category, for example $\overline{\text{N}}$, or one can use a feature that marks the presence of PPs and forbid Ns with this feature as daughters.

Note that a rule system allowing an undetermined number of daughters in a sentence with adjuncts has to be combined with a semantics, which is possible (Kasper 1994) but not trivial. This is the topic of Section 6.

Discussing semantics, Gibson (2025: 170) states that "we can think of dependency grammar as an abstraction over categorial grammar, without the semantics of the phrases". This is not

quite true, since it is impossible to please Adele Goldberg and Mark Steedman at the same time. Steedman (2000) uses very general abstract rules for the combination of lexical material. Gibson argued for very specific phrasal schemata, one schema for every possible argument structure. For example, he has rules for monovalent intransitive and strictly transitive verbs (p. 56) and rules for ditransitive verbs (p. 79), and subject, verb, complement clause combinations (p. 81). The verbs use diacritics encoding their valence (V_N for transitive verbs, $V_{N, N}$ for a ditransitive verb, V_{Comp} for a verb taking a subject and a sentential complement) so that only verbs with the correct diacritic can enter one of these rules. See also Tomasello (2003: 98–99) for explicit statements against abstract rules.

A further problem for accounts like Gibson's that was also pointed out by Müller (2023c: Section 5.5.2) is partial verb phrase fronting. German allows for the fronting of verbs with any subset of their complements.

(12) a. [Erzählen] wird er seiner Tochter ein Märchen müssen.
tell will he.NOM his.DAT daughter a.ACC fairy.tale must
'He will have to tell his daughter a fairy tale.'

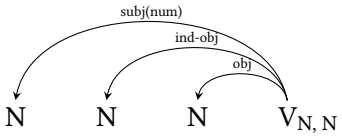
b. [Ein Märchen erzählen] wird er seiner Tochter müssen.
a.ACC fairy.tale tell will he.NOM his.DAT daughter must

c. [Seiner Tochter erzählen] wird er das Märchen müssen.
his.DAT daughter tell will he.NOM the.ACC fairy.tale must

d. [Seiner Tochter ein Märchen erzählen] wird er müssen.
his.DAT daughter a.ACC fairy.tale tell will he.NOM must

For ditransitive verbs in German, Gibson would have to assume a rule like (13), which is the head-final analogue to his rule for ditransitive verbs in English on p. 85.

(13) Dependency rule for German ditransitive verbs:



Note first, that adjuncts can appear anywhere between or before or after the arguments.

(14) *Gestern* hatte *in der Mittagspause* der Vorarbeiter *in der Werkzeugkammer* dem
yesterday had during the lunch.break the foreman in the tool.shop the
Lehrling *aus Boshaftigkeit langsam* zehn schmierige Gußeisenscheiben *unbemerkt* in
apprentice maliciously slowly ten greasy cast.iron.disks unnoticed in
die Hosentasche gesteckt.[8]
the pocket put
'Yesterday during the lunch break, the foreman maliciously put ten greasy cast iron disks slowly into the apprentice's pocket unnoticed.'

[8] Uszkoreit (1987: 145).

This means that Gibson either has to assume non-projective structures where the adjuncts are inserted into discontinuous parts of (13) or he has to allow multiple applications of an adjunct rule to the V symbol in (13). This of course would contradict the assumption that Przepiórkowski (2026: Section 5.1) made in a formalization of Gibson's rules that every rule may be used at most once in a combination of rules. Note further that the fact that there can be arbitrarily many adjuncts interspersed with arguments in German clauses is the reason for assuming binary branching structures in many phrase structure grammars of German (Berman 1996: Section 2.1.3, Müller 2024a: Section 3). This shows that there are many open ends in Gibson's approach that need fixing, which certainly will make his approach more complex, since the complexity that is saved in terms of nodes will show up elsewhere.

Returning to the examples in (12), they break all phrasal approaches, since ditransitive verbs have to be combined with one of their objects to the exclusion of the other one (and their subject) in (12b,c), unless one wants to assume non-projective structures (see Section 8 and especially Figure 12). The respective other object (and the subject) is realized elsewhere in the clause. This is against the very nature of phrasal approaches that claim that certain aspects of the meaning are only available when a verb enters a certain corresponding phrasal construction (Goldberg 1995: Chapter 6, Tomasello 2003: Section 5.1.3). For (12b) the required construction would be the phrasal construction for verbs taking an accusative object and for (12c) it would be the phrasal construction for verbs taking a dative object, but *erzählen* 'to tell' is a ditransitive verb in both cases. Nerbonne (1986) and Johnson (1986) solved this problem by incorporating Categorial Grammar-like treatments of valence information into GPSG. Shortly after this GPSG was given up and all researchers who worked in this framework switched to the Categorial Grammar-like framework HPSG (Pollard & Sag 1987). See Müller (2013: Section 2.3) for a comparison of Categorial Grammar, Minimalist Grammars and HPSG.

Another problem with Gibson's rules is the analysis of coordination. Gibson (2025: Section 3.17) assumes that the coordination is the head of the dependency. So *and* with part of speech Coord ist the head of *the man and the woman*. This is problematic since the governing verb does not select for a coordination but for an N (Borsley 2005, Müller 2023b: Section 4.10.2.5). Gibson writes "the Coord category matches any syntactic head; and all conjuncts (sometimes called "coordinands") agree with this type, including argument structures to the left and right that are being coordinated". This is not worked out. Maybe he has in mind an analysis like the Categorial Grammar analysis in which the conjunction is a functor taking two similar arguments to project something of the same category: (X\X)/X. But the text is not explicit about this and in any case, working this out would add to complexity.

6 Simplicity: Counting nodes and semantics

Gibson (2026) states at several places of his paper, that DG is simple. See also Osborne (2018: 6). It is not. Proponents of DG just use a subpart of what would be a complex theory if it were fully worked out.

Müller (2019: 59–60) pointed out that phrase structure is useful for determining the semantics of the unit of a head and its dependents. If the head is assigned a meaning that is usually assumed for lexical items and if there is no extra node for a phrase or constituent, then there

is no place to state what the meaning of the head and its dependents is. So, while in phrase structure grammars, the meaning of *children* is something like $\lambda y.child(y)$, in DGs, it must be the meaning of the whole nominal unit (Hudson 2003: 391–392).

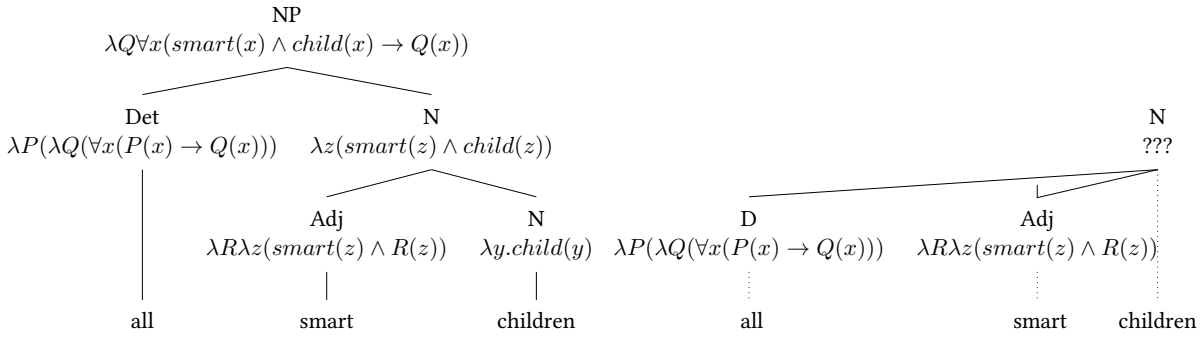


Figure 6: Nodes are used for meaning representation

A way to account for this is to use a special feature for lexical meaning and one for the meaning of a head with all dependents. But this is basically shifting complexity into nodes.

Hudson develops an explicit model including semantics. He states that the representation in Figure 7 is the worked out version of the usual dependency graphs. The abbreviated form is given in Figure 8.

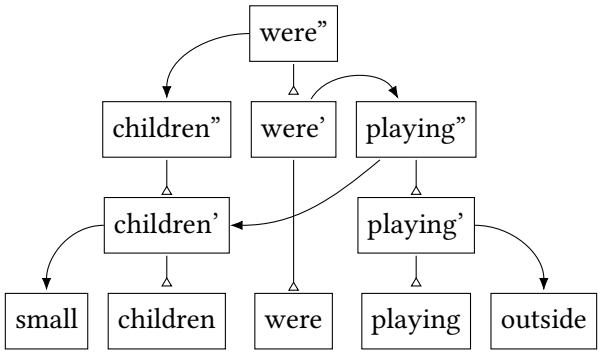


Figure 7: Analysis of *Small children were playing outside.* according to Hudson (2018: 105)

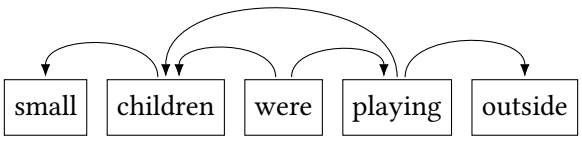


Figure 8: Abbreviated analysis of *Small children were playing outside.* according to Hudson (2018: 105)

Now, as Müller (2019) points out, the Figure 7 with category symbols filled in as in Figure 9

suspiciously looks like an $\overline{\text{X}}$ tree. There is no advantage for DG in terms of node numbers/node

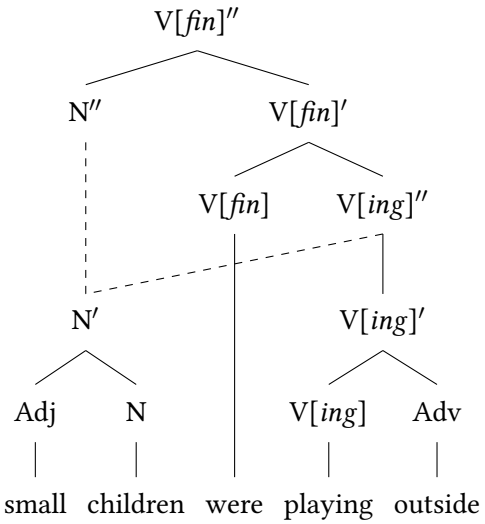


Figure 9: Analysis of *Small children are playing outside.* with category symbols

complexity as soon as one looks at fully worked out theories including semantics.

Before we go on to the next section, I want to point the reader to Tesnière (1959: 154, Stemma 149, 2015: 150, Stemma 149). Tesnière discussed the phrase in (15) and suggested two different ways of combining the adjunct with the head for semantic reasons.

(15) red cars that you saw yesterday

So he argued for a more articulate representation (polygraphs; Kahane & Osborne 2015: lix) and not just simple dependencies and flat representations. Figure 10 shows Tesnière's representation of scope information and the corresponding phrase structure grammar.

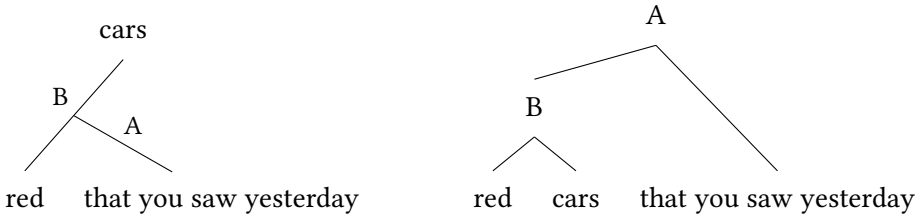


Figure 10: Tesnière's way of representing scope and the comparison with phrase structure-based analyses by Kahane & Osborne (2015: lix)

7 Dependency vs. constituency and the question of headedness

There is a clear case that decides the question of dependency vs. constituency: the question of unheaded structures. In many frameworks it is assumed that structures are headed or that there is a more prominent part that selects the other part or parts. This is true for Dependency Grammar (Osborne 2019: 48), Categorial Grammar (Steedman 2000), and Minimalism (Chomsky 1995). This means that they do not have accounts for unheaded structures. An example is the NPN construction (Matsuyama 2004, Jackendoff 2008):

(16) Student after student after student left the room.

One property of this construction is that there can be arbitrarily repetitions of the PN part and the repetition just means 'many' without the individual PN combinations adding anything to the meaning. The nouns appear without determiner and have to have the same form of the initial noun. Unless one assumes that all parts of the construction depend on an empty N head it is not clear how these constructions can be analyzed under DG assumptions.

The first analysis of this construction in DG seems to be by Hudson (2021), but Hudson works in Word Grammar and the expressivity of this framework includes more than the head-dependent configurations that are known from DG. Hudson uses arcs to create loops. His analysis is shown in Figure 11. The lower arc from the second 1 goes back to noun_{npn}, which accounts for the possibility to have arbitrarily many PN repetitions. As should be clear from the

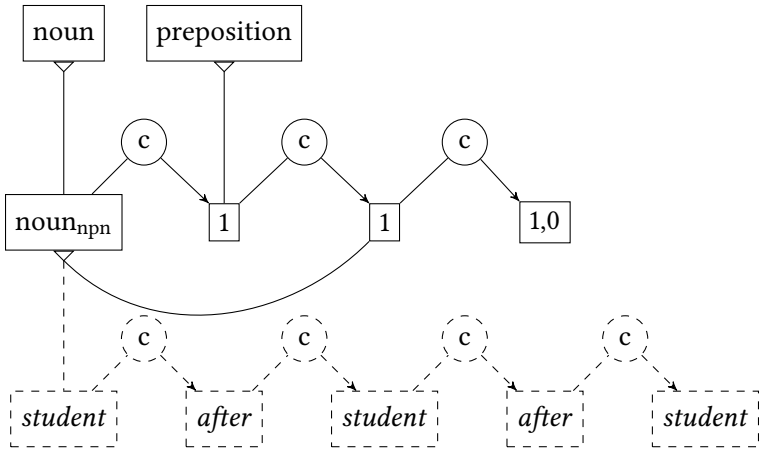


Figure 11: The NPN construction in Word Grammar according to Hudson (2021: 1477)

figure, this requires more expressivity than usual DG formalisms.

Covering the NPN construction for phrase structure-based approaches is relatively easy: There is a certain configuration with a certain meaning and a certain distribution.[9] The material

[9] A reviewer worries that "it might be a little too easy" and asks whether there are "any limits on when headlessness arises, or on the form of exocentric Phrase Structure Rules?" For phrase structure grammars as defined by Chomsky & Miller (1963: Section 4) there is no restriction: a symbol on the left-hand side of a rule can be rewritten as any sequence from the set of available symbols. The rule in (i) from before the introduction of IP (Chomsky 1981: 19, 25) is a famous example.

in (16) is categorized as an NP with the meaning that results from its parts (the nouns and the preposition in a certain configuration). See Bargmann (2015) for an analysis and Müller (2024b: Section 4.1) for discussion of earlier proposals in other frameworks.

A reviewer remarked that there seems to be some inconsistency in the argumentation in this paper. I first claimed that dependency information is valence information or – more general – functor information and that Dependency Grammars can be translated into (certain) phrase structure grammars (e.g., Figure 3 above) and then went on by claiming that Dependency Grammars cannot account for certain phenomena. There is no contradiction here. What I discussed above was head-dependent structures (subject-head, head-complement, head-adjunct). These are equivalent to DG rules. But for rules that just combine material without stating any dependency between the parts, there is not equivalent structure in DG, unless an empty head is assumed. To take a concrete example: the Valence Principle is defined for phrases of type *headed-phrase* in Head-Driven Phrase Structure Grammar (HPSG), but in addition to headed phrases, phrases of type *non-headed-phrase* are assumed (Ginzburg & Sag 2000: 363,364).

For further problems for theories that are entirely head-based like Minimalism, Categorial Grammar, and Dependency Grammar see Jacobs (2008) and Müller (2023c: Section 21.10). Gibson suggested a fusion of Dependency Grammar and Construction Grammar, but without empty heads and/or other extensions, it will not be possible to capture all phenomena that are discussed in the Construction Grammar literature.

(i) S → NP Infl VP

It is only additional theoretical considerations that require headedness. Chomsky (1986: 3) suggested a way to make (i) headed by introducing the category I.

(ii) a. IP → NP I′
b. I′ → I VP

The I element is an empty head in the analyses of some English sentences. Kornai & Pullum (1990: Section 3.2) showed that $\overline{\text{X}}$ theory as a set of constraints on phrase structure grammars does not help much, if one allows for empty elements. This can be demonstrated for the case at hand: Rather than declaring any element of the N-P-N sequence to be the head, one can assume a new empty head and let this empty head select the parts of the N-P-N construction:

(iii) NewP → New N P N

Some grammarians use empty element only as a last resort (Bender 2001, Sag, Wasow & Bender 2003: 464) and rather assume headless constructions. A further example is Bresnan (2001: Section 6.2.2), who assumes exocentric structures for a variety of languages.

(iv) S → C*

This phrase structure rule combines arbitrary constituents (C*) into an S.

The analysis of coordination in HPSG often assumes unheaded constructions (Abeillé & Chaves 2024). See (Borsley 2005, Müller 2023b: Section 4.10.2.5) on the ConjP analysis, which assumes the conjunction to be the head. See Müller (2022) for a general discussion of headless constructions.

8 The intricacies of syntax: Projectivity

The success of DG in computational linguistics and in corpus linguistics is probably due to the fact that dependency annotation is simpler as phrase structure annotations where ambiguities are reflected in phrase structure. But as I hope to make clear in this paper, dependency is not enough. One aspect of competence grammars is that they can be used both ways: they can be used for parsing and for generation. Corpus annotation is just the parsing part. If discontinuities like those in Figure 1 and Figure 12 are allowed, one has to make sure that generation does not result in word salad. As Müller (2023c: 410) pointed out, non-projective DGs (Heringer 1996:

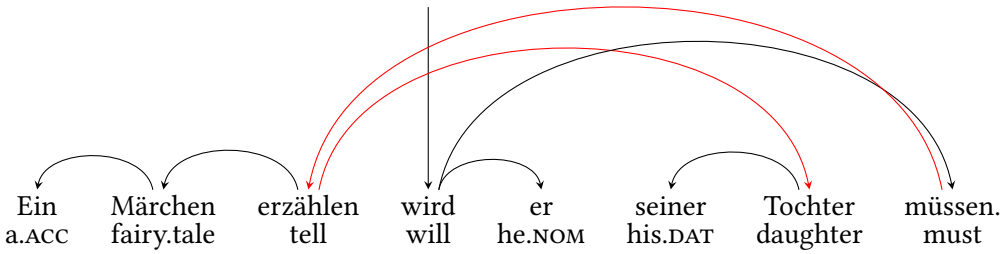


Figure 12: Analysis of (12b) requiring crossing dependencies

261, Eroms 2000: Section 9.6.2, Gibson 2025: 112, 137) would allow for the word salad in (17b):

(17) a. Deshalb klärt, dass Peter kommt, ob Klaus spielt.
therefore resolves that Peter comes whether Klaus plays
'That Peter comes does not resolve the question of whether Klaus comes.'

b. * Deshalb klärt dass ob Peter Klaus kommt spielt.
therefore resolves that whether Peter Klaus comes plays

The words from the sentences *dass Peter kommt* and *ob Klaus spielt* are all in the correct order as far as the subclauses are concerned, but they are interspersed. So, there has to be a statement somewhere in the grammar saying that both *dass Peter kommt* and *ob Klaus spielt* form continuous units.

9 V2 and "movement"

The following discusses simplistic approaches to German that are common in the DG literature. I first show that flat structures are not appropriate to account for verb second phenomena and then go into the details of a very complex phenomenon: apparently multiple frontings. The result is that DG descriptions of the German clause are way too simple and hence cannot be compared to phrase structure-based approaches unless they worked out analyses of the respective phenomena in detail.

German is regarded as a V2 language. The data in (18), which is taken from Müller (2023c: Section 11.7.1) suggests that fronted material differs from material in non-fronted position:

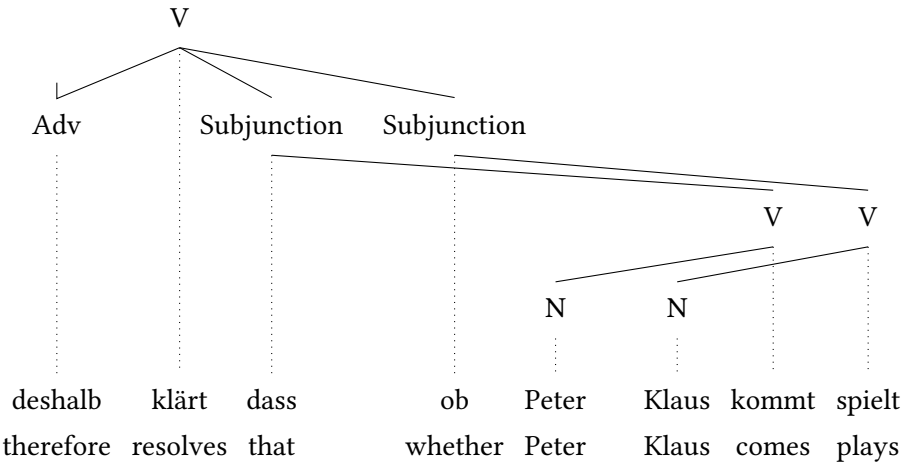


Figure 13: The dependency graph of the word salad *Deshalb klärt dass ob Peter Klaus kommt spielt.* 'Therefore resolves that whether Peter Klaus comes plays' which is admitted by non-projective Dependency Grammars that do not restrict discontinuity

(18) a. Oft liest er das Buch nicht.
often reads he the book not
'It is often that he does not read the book.' or 'It is not the case that he reads the book often.'

b. dass er das Buch nicht oft liest
that he the book not often reads
'It is not the case that he reads the book often.'

c. dass er das Buch oft nicht liest
that he the book often not reads
'It is often that he does not read the book.'

(18a) has two readings, while both (18b) and (18c) are not ambiguous in the same way. Since the relative order of *oft* and *nicht* is the same in (18a) and (18c), one would not expect that (18a) has additional readings if it were just a linearization variant of (18c). The additional reading can be explained by assuming that *oft* 'often' is extracted with different scopings in (18a) while it is not in the other two sentences. This shows that a flat representation of a head and all its dependents as it is common of much of the DG literature on German (Heringer 1996: 261, Eroms 2000: Section 9.6.2) is not sufficient. V2 should always be handled by a special mechanism, even the apparently simple cases. Heringer and Eroms allow for non-projectivity for V2. The sentences in (18) do not require non-projectivity, but the examples show that there is more going on than simple linearization in a flat structure.

Examples like (19) show that deeply embedded material can be fronted. If one assumes projectivity, some other mechanism has to be used for "movement".

(19) a. [Um zwei Millionen Mark]$_i$ soll er versucht haben, [eine
around two million Deutsche.Marks should he tried have an

Versicherung _$_i$ zu betrügen].[10]
insurance.company to deceive
'He apparently tried to cheat an insurance company out of two million Deutsche Marks.'

b. "'Wer$_i$, glaubt er, daß er _$_i$ ist?"' erregte sich ein Politiker vom Nil.[11]
who believes he that he is retort REFL a politician from.the Nile
' "Who does he think he is?", a politician from the Nile exclaimed.'

c. Wen$_i$ glaubst du, daß ich _$_i$ gesehen habe?[12]
who believe you that I seen have
'Who do you think I saw?'

d. [Gegen ihn]$_i$ falle es den Republikanern hingegen schwerer,
against him fall it the Republicans however more.difficult
[[Angriffe _$_i$] zu lancieren].[13]
attacks to launch
'It is, however, more difficult for the Republicans to launch attacks against him.'

If one allows for non-projectivity (Heringer 1996: 261, Eroms 2000: Section 9.6.2, Gibson 2025: Section 7.7), then ways have to be found to exclude non-projective structures as the one suggested in the previous section. Gibson (2025: 183) suggests that there are preferences for projective structures, but again, this is not worked out and the mathematics of setting up a respective system is not trivial.

The sentence in (20) suggests that both simple and complex sentences should be analyzed in parallel, since a simple sentence with a gap can be coordinated with a complex sentence with a gap and the filler can fill all gaps in the conjuncts.

(20) Wen$_i$ kennst _$_i$ du und glaubst du, dass _$_i$ jeder kennt?
who knows you and believe you that everybody knows
'Who do you know and do you believe that everybody knows?'

So the data in (18a) and the extraction cases discussed above suggest that German V2 is a unified phenomenon both for simple and complex sentences and it involves a nonlocal dependency that differs from normal local dependencies. This is acknowledged by researchers like Groß & Osborne (2009), who use a special mechanism for nonlocal dependencies. They indicate the nonlocal dependency by a little g in their representations. Figure 14 shows the analysis of (19c).[14] The authors call the mechanism rising but do not provide any formal details. One can imagine some GPSG-like slash passing mechanism. The DG analysis would then be equivalent to traceless GPSG and HPSG proposals (Müller 2023c: 189, Bouma et al. 2001). Gibson (2025: Section 3.19) indeed assumes a slash-based analysis for fronting in English.

[10] taz, 04.05.2001, p. 20.
[11] Spiegel, 8/1999, p. 18.
[12] Scherpenisse (1986: 84).
[13] taz, 08.02.2008, p. 9.
[14] The same idea can already be found in Heringer's work (1996: 261).

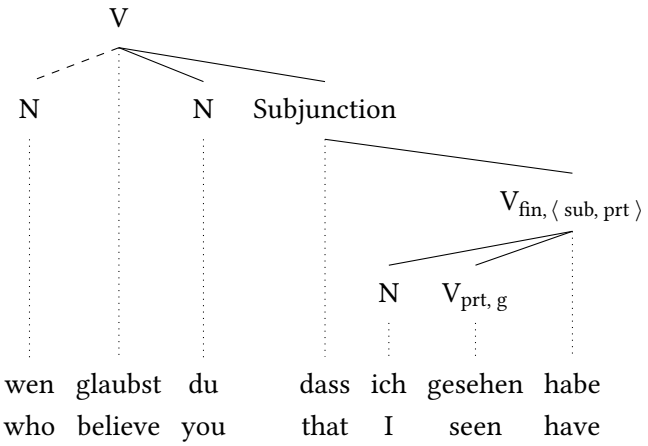


Figure 14: Projective analysis of *Wen glaubst du, dass ich gesehen habe?* 'Who do you think I saw?' involving rising

This brings us back to the example involving split NPs in the beginning of the paper (see Figure 1). Discussing the examples from Section 12 below, Gibson (2026: 7) writes: "nontransformational accounts postulate direct links between words that depend on one another for meaning". Such an approach would result in crossing dependencies as they are displayed in Figure 1 and hence in non-projectivity.

For sentences like (21), one would not get crossing edges but the V2 property of German would be violated.

(21) [Dem Saft] [eine kräftige Farbe] geben Blutorangen.[15]
the juice a strong color give blood.oranges
'Blood oranges give the juice a strong color.'

The dependency graph of this sentence is given in Figure 15.

The question is: What can appear before the finite verb? How many constituents? Where can they come from? A generalization about these frontings that has been confirmed in a broad study of 3,200 corpus examples (Bildhauer 2011) is that they have to depend on the same verb (Fanselow 1987: 99). Fanselow discussed the examples in (22):

(22) a. Ich glaube dem Linguisten nicht, einen Nobelpreis gewonnen zu haben.
I believe the linguist not a Nobel.prize won to have
'I don't believe the linguist's claim that he won a Nobel prize.'

b. * Dem Linguisten einen Nobelpreis glaube ich nicht gewonnen zu haben.
the linguist a Nobel.price believe I not won to have

(22b) is out since *dem Linguisten* 'the linguist' depends on *glauben* 'to believe' and *einen Nobelpreis* 'a nobel price' depends on *gewonnen* 'won'. *dem Linguisten* and *einen Nobelpreis* can be fronted alone but not together. The place in front of the finite verb is reserved for one constituent.

[15] Bildhauer & Cook (2010) found this example in the *Deutsches Referenzkorpus* (DeReKo), hosted at Institut für Deutsche Sprache, Mannheim: `http://www.ids-mannheim.de/kl/projekte/korpora`.

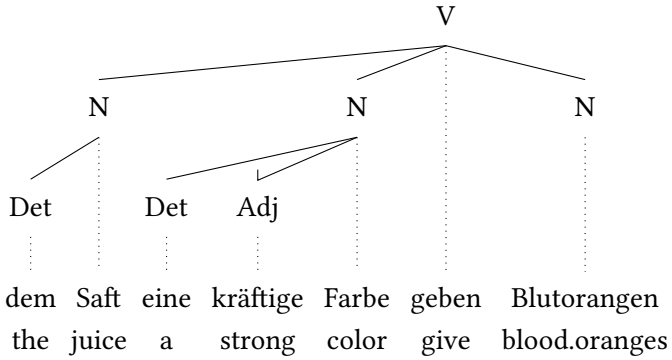


Figure 15: Dependency graph for *Dem Saft eine kräftige Farbe geben Blutorangen.* 'Blood oranges give the juice a strong color.'

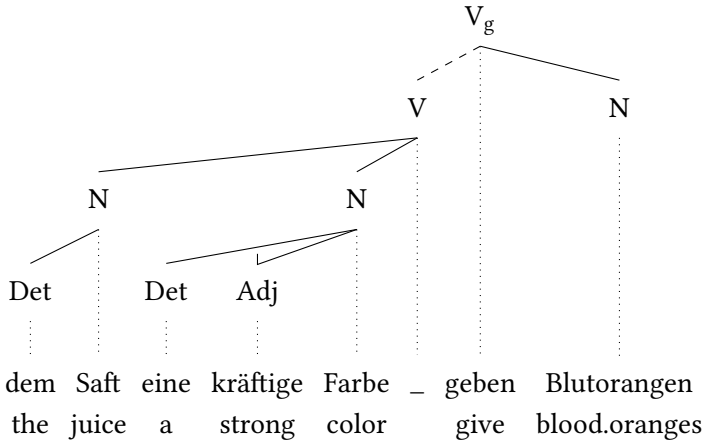


Figure 16: Dependency graph for *Dem Saft eine kräftige Farbe geben Blutorangen.* 'Blood oranges give the juice a strong color.' with an empty verbal head for the *Vorfeld*

How can this be modeled in DG without empty heads? See Müller (2005, 2023a) for an extensive discussion and rejection of possible analyses including multiple slash-based extractions.

Figure 16 shows a possibility for Groß & Osborne to analyze the example with an empty head. The fronted constituents would depend on an empty verbal head and this is related to a verb in the rest of the clause. (Note that Groß & Osborne (2009: 73) reject the assumption of empty elements.) Of course this begs the question what kind of empty head this is. If one assumes a verb-movement account together with formation of verbal complexes, the empty head needed for multiple frontings is just the element that is used for verb movement as well (Müller 2005: 316).

Returning to the sentence in Figure 1, this sentence shows a complex interaction of several phenomena. German allows for discontinuous NPs, which are usually analyzed via argument/adjunct attraction (De Kuthy 2002: 73–80). That is, part of an NP is raised to a higher head. In the case at hand, the PP dependent of *Blick* 'glance' is raised to *haben* 'have'. As a co-dependent of

haben 'have' it can be realized together with *stets* 'always' in fronted position.

Whatever the DG solution for the analysis of such examples will be, it is clear that it will involve much more than simply drawing lines between heads and dependents. The alledged simplicity of DG is just cherry picking.

10 Transformations

The following two sections deal with transformations. In principle, the issues concerning Transformational Grammar are independent of the question of whether Dependency Grammars or Phrase Structure Grammars are simpler. However, Gibson addresses Transformational Grammars in his contribution. Since some of the analyses in Mainstream Generative Grammar have been taken over to alternative frameworks it may be worthwhile to discuss Gibson's claims on Transformational Grammars.

Gibson (2026) argues that the assumption of transformations is dubious. Transformations as such are known not to be psycholinugistically plausible for a long time now (Fodor, Bever & Garrett 1974: 320–328). Psycholinguists like Frazier & Clifton (1996: 27) suggested that transformations could be precompiled before actual usage, but this then means that humans work with precompiled schemata and templates that basically would correspond to what Construction Grammar suggests, but their linguistic knowledge is represented as a Transformational Grammar. This is a very implausible assumption, if it is assumed that the Transformational Grammar corresponds to something that is represented in our brains combined with poverty of the stimulus and innateness claims (Berwick, Pietroski, Yankama & Chomsky 2011). It begs the question why we would need the transformational grammar in the first place and do not develop theories about the "precompiled stuff", which are proven to be more compatible with theories of language acquisition (Freudenthal et al. 2007, 2009). The reader may also consult Figure 17. It is clear that the acquisition of the fact that a certain preposition appears with an NP with a certain case is rather straight-forward in comparison to the acquisition of one of the structures suggested in Minimalist textbooks.

So, transformations mapping tree structures to other tree structures are psycholinguistically not plausible, but there have been representational variants of Transformational Grammar in the 1980ies and the 1990ies (Koster 1978, 1987: 235, Kolb & Thiersch 1991, Haider 1993: Section 1.4, Frey 1993: 14, Lohnstein 1993: 87–88, 177–178, Fordham & Crocker 1994: 38, Veenstra 1998: 58). Such representational variants are very similar to theories like LFG and HPSG in the versions that used a trace for extracted elements: there is a filler and somewhere else there is a corresponding trace. The representational variants of Transformational Grammar are basically grammars with just one structure representing the linguistic object. Movements are indicated by coindexed empty elements. Variants of HPSG and LFG that use traces are parallel to such representational variants of MGG. So, there are transformation-less versions of the mainstream theories that are basically surface-oriented phrase structure grammars. The criticism against general transformations of the era before GB does not apply to those theories. Note also, that some transformations, namely lexically-governed transformations correspond to lexical rules in non-transformational frameworks (Dowty 1978) and Gibson (2025: Section 3.9.1, Section 3.13, 95) is assuming lexical rules. Note also that Gibson (2025: 95) is using the SLASH mechanism of

GPSG. So this is the phrase structure equivalent to movement in Transformational Grammars.

With the Minimalist Programm (Chomsky 1995), Mainstream Generative Grammar took a more derivational turn again and hence, Gibson's criticism applies to these approaches but not to representational variants or variants like Stabler's Minimalist Grammars (Stabler 2011, Torr & Stabler 2016, Torr 2019). See Müller (2023c: 177-180) on Torr's version of Minimalist Grammars.

See Osborne, Putnam & Groß (2011) for a comparison of Minimalist approaches and Dependency Grammar.

11 (Transformational) Grammar and complexity results

Gibson (2026: 7) describes transformational analyses of the passive and argues that (23a) is derived from (23b) by several transformations.

(23) a. Lana will be kissed by Mitchell.
b. Mitchell will kiss Lana.

This is not quite true. It used to be the case that passive sentences were derived from patterns corresponding to active sentences (Chomsky 1957: 43), but in Government & Binding the passive actually had a lexical analysis (Gereon Müller, p. c. 2018). It was somehow claimed that the passive morphology blocked the subject and the participle lost the property to assign accusative. Due to the Case Filter, the object had to move to Spec,IP to get case there (Chomsky 1981: 124). Chomsky explicitly states that the D-structure of (24a) is (24b) rather than any active sentence:

(24) a. John was killed.
b. [$_{NP}$ e] was killed John.

Gibson concludes his discussion of transformations by saying "the compositional theories without such transformations are simpler than those with them [84]". [84] refers to Peters & Ritchie (1973), which is a study of the generative capacity of general transformations as they were common in the time before Peters & Ritchie's result. Transformations in later versions of the theory were different, they couldn't map arbitrary trees onto other arbitrary trees and they contained traces of material that was moved (Klenk 2003: Section 3.1.4). So, these complexity results are irrelevant for current versions of Transformational Grammar.

As for current Minimalist theories, there is often a lack of formalization (Stabler 2011: 397, 399, 400) or outright inconsistency, so that complexity results cannot be obtained. An example is Chomsky's solution to labeling (Chomsky 2008, 2013), which did not just have empirical problems but also was lacking the constraints for one possible combination of items for several years. The constraints provided in 2013 were simply inconsistent (Müller 2013: Section 2.1). There are formalized versions of Minimalist theories though and for these there are complexity results. Stabler's Minimalist Grammars (Stabler 2001, 2011) were shown to be mildly context sensitive (Michaelis 2001). This is exactly the computational complexity that is needed, since human languages are assumed to have this complexity (Joshi 1985).

So, Gibson cannot argue against transformational theories on the basis of computational complexity. What is more, I dare say that such complexity approximations of formal systems are irrelevant for (psycho)linguistics anyway. They are worst-case approximations of the formalism

available to the linguist, but the crucial factor is what elements of the formal language are used in an actual analysis. Such analyses are the things we should judge and the (psycho)linguistic predictions of which we should evaluate. Pollard (1997: 8–10) made this explicit: the formal apparatus of HPSG (features, values, relational constraints) has Turing power, but nobody cares about this, since it is not the formal power that is available in principle but what is actually used in analyses. Similarly, the LFG formalism has Turing power, but by adding a constraint (*Off-Line Parsability Constraint*; Johnson 1988: Section 3.5), the complexity can be reduced. So it is constraints that can be formulated within the general framework that are important.

To take a concrete example, let us compare the analyses of prepositional phrases which was discussed in Minimalist textbooks (Radford 1997: 452, Hornstein et al. 2005: 124). Radford's analysis is given in Figure 17 and the analysis in all non-MGG phrase structure grammars and Dependency Grammars is given next to it for comparison. The reader is referred to the

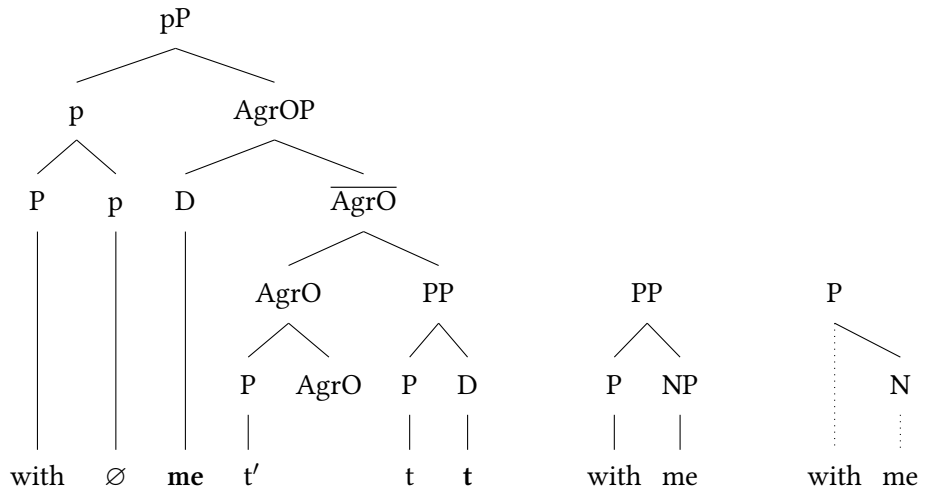


Figure 17: Minimalist PP analysis following Radford (1997: 452) compared with all other PSG and DG analyses

introductory texts cited above for the explanation of the Minimalist analyses and all the empty functional categories and the involved movement operations to derive a two word phrase. See also Sternefeld (2006: 549–550) for a critical discussion.

A reviewer pointed out to me that it might be unfair to criticize the analysis in Figure 17 without taking the specific goals of the analysis into account. This is true, but the reason for the complexity is that there was the assumption that all DPs check their case by moving into a specifier position of an agreement projection (Chomsky 1993: 7, Radford 1997: 451). Since the DP in a PP is not in a specifier position, the DP is moved to a specifier position of an empty head. Since the order of DP and preposition is then wrong, the preposition has to be moved into the position of another empty head. Of course, all this could be prevented by assuming that case requirements can be checked with complements directly.

In any case there is enough to criticize in current Minimalist theories for psycholinguists: for example, the idea that complete Phases (Phase is a special term in Minimalist reasoning;

Chomsky 2008) are build and then sent to the interfaces is incompatible with psycholinguistic findings regarding incremental processing (Marslen-Wilson 1975, Tanenhaus et al. 1996).

A reviewer remarks that various phenomena may require bottom up construction of linguistic objects (e. g. Bermádez-Otero 2011). Even if this were true, it would not mean that the whole linguistic system of all languages works this way. Constraint based systems are compatible with psycholinguistic findings since they allow for predictions of upcoming material. The phonological, morphological, syntactic, semantic, information-structural and pragmatic information that is available for a certain utterance at a certain point in time can be processed immediately (Wasow 2024). This would also work for material that is the result of a complex process. The processor would start to reverse the process at the time it occurs in the input. For example, Bruggeman et al. (2025) show for polysynthetic language Murrinhpatha that its complex one-word utterances are processed incrementally. Item-based Morphology (Bonami & Crysmann 2016) is compatible with incremental processing. See Davis & Koenig (2024) on the structure of the lexicon in HPSG.

12 Dependency Locality

The most convincing case for DG are Gibson's examples showing that for the analysis of relative clauses the length of dependencies between relative phrase and verb is important rather than the distance between filler and gap in movement-based approaches (Pickering & Barry 1991, da Cunha & Gibson 2025). This is illustrated in Figure 18.[16]

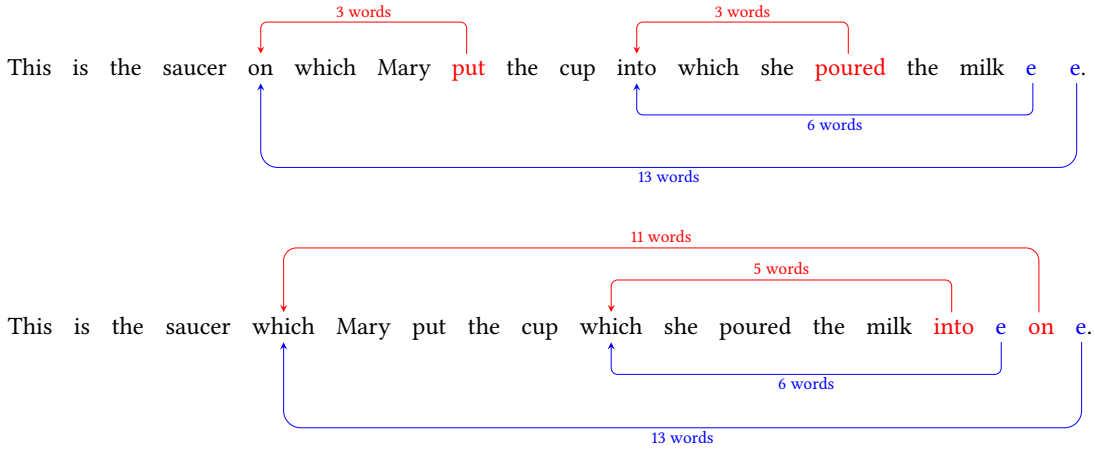


Figure 18: Distance between head and dependent vs. filler and gap

Transformational accounts (blue) predict that the processing costs for both examples in Figure 18 are similar, while dependency-based accounts (red) predict that the upper sentence is more easy to process since the distance between dependent element and head is much shorter: 3 vs. 11 words.

[16] Da Cunha & Gibson (2025: 14) and Gibson (2026: 7) state that the distance between the first *which* and the trace in the second figure is 12 words, but it is 13 words as in the first figure.

The interesting thing now is that there are traceless approaches to nonlocal dependencies as the one by Bouma et al. (2001) in HPSG and by Sag (2010: 508) in the HPSG variant Sign-Based Construction Grammar. The information about the missing constituent is only present at the head. So the Dependency Grammar approach makes exactly the same predictions as the non-transformational Phrase Structure Grammar approach.

The traceless approach to extraction is not accepted by everybody in the HPSG community because of problems with scope of extracted adjuncts (Levine & Hukari 2006), but even approaches with actual traces for extracted elements can account for Gibson's insights. The reason is that both *put* and *poured* contain PPs in their valence lists and the hearer/speaker of the utterance can identify the extracted element with the requirement in the verb.[17] Note also that the information that the PPs are extracted is also represented at the verbs, since the traces have a SLASH value that is identified with the SLASH value of the PP in the valence list of the verbs. So the valence lists of the verbs look the same independent of whether empty elements are assumed or not.

This is shown in Figure 19: the verb *put* selects for two NPs and a PP. Since the PP is realized as a trace, its category is PP/PP. Since the verbs requirement is identified with the object that is actually there, overt or non-overt (4), the information about the SLASH value of the PP trace is available in the valence list of the verb as well. So, when the relative clause is parsed

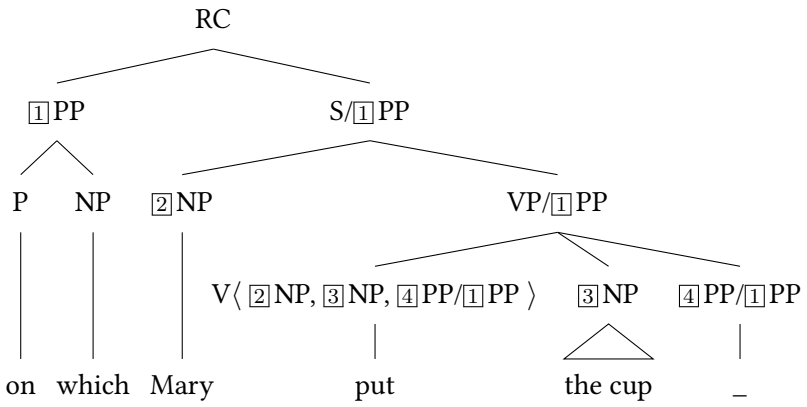


Figure 19: Information about extraction is available both at the head and at the trace.

incrementally, *on which* can be identified as the material that is needed by the verb *put*. Since the PP was first seen as a relative phrase, it is clear that it is a filler in a nonlocal dependency and it is expected that there will be a trace further to the right. To be precise: when a phrase like (25) is parsed, it is clear after the word *which* that the phrase *on which* is part of a relative clause and a sentence missing *on which* (1) is expected. 1 is percolated down to the trace. The trace identifies the nonlocal information (1) with the local information (4). This identifiaction is not shown in the figure, to make the "flow of information" more transparent. This is for explanatory

[17] See also Borsley's review of Gibson (2025). Borsley (2026: 3) makes a general remark that all constituent-based frameworks with selection of sisters incorporate the notion of dependency. See also the discussion of Figure 3 above.

purposes only. In reality there is no "flow", there is just identity. 4 is shared between the trace and the requirement in the valence list of *put* and hence the fact that the PP argument is extracted is available within the lexical item for *put*.

(25) the table on which Mary put the cup

This means that it is accessible for measuring the locality of dependencies. See Borsley & Crysmann (2024) for an overview of nonlocal dependencies in HPSG and Levine & Hukari (2006) for further details on adjunct extraction.

Da Cunha & Gibson (2025: 28) refer to earlier work discussing a trace-based analysis using the principle of First-Resort Gap Positing, but I showed another rather direct way that is possible for frameworks that register extraction at the head.

So, it seems to be the case that dependency-based and slash-based analyses of extraction are empirically indistinguishable (Sag & Fodor 1995: 376). Of course, analyses without empty elements are preferable over those with empty elements for reasons of parsimony and it is often possible to convert grammars with empty elements into grammars without any (Müller 2004b: Section 7.4), but there are situations in which they seem unavoidable or preferable (Sag et al. 2003: 464, Müller 2005, 2023a, Borsley 2013: Section 4.2).

13 Conclusion

I showed that dependency relations are important but they are not sufficient to account for grammatical facts. Constraints on adjacency and projectivity are needed. It follows that claims about the alleged simplicity of Dependency Grammars in comparison to other frameworks are not warranted. Gibson was right in pointing out the psycholinguistical implausibility of Transformational Grammars but representational models of Transformational Grammars do not need literal transformations but merely empty elements and nonlocal dependencies.

Since dependency information is contained in some phrase structure grammars, the psycholinguistic results can be used for such models as well. This means that the simpler DG representations can be used for psycholinguistic purposes, but for worked out theories of language more is needed.

Acknowledgements

I thank Anne Abeillé, Bob Borsley, and Rui Chaves for discussion and pointers to further literature. I thank Sten Vikner, Matthew Korte and the participants of the Syntax-Semantics Collogium at the Humboldt Universität zu Berlin for discussion. Thanks to Adam Przepiórkowski for comments on an earlier version of the paper and for discussion.

Competing interests

There are no competing interests.

References

Abeillé, Anne & Rui P. Chaves. 2024. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 2nd edn. (Empirically Oriented Theoretical Morphology and Syntax 9), 775–829. Berlin: Language Science Press. DOI: 10.5281/zenodo.13645041.

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.

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 146), 52–77. Darmstadt: Wissenschaftliche Buchgesellschaft.

Behaghel, Otto. 1909. Beziehung zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142. DOI: 10.1515/9783110242652.110.

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

Bermádez-Otero, Ricardo. 2011. Cyclicity. In *The blackwell companion to phonology*, chap. 85, 1–30. John Wiley & Sons. DOI: 10.1002/9781444335262.wbctp0085.

Berman, Judith. 1996. Eine LFG-Grammatik des Deutschen. In Judith Berman & Anette Frank (eds.), *Deutsche und französische Syntax im Formalismus der LFG* (Linguistische Arbeiten 344), 11–96. Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783110955354.

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

Bildhauer, Felix. 2011. Mehrfache Vorfeldbesetzung und Informationsstruktur: Eine Bestandsaufnahme. *Deutsche Sprache* 39(4). 362–379. DOI: 10.37307/j.1868-775X.2011.04.

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. DOI: 10.21248/hpsg.2010.4.

Bonami, Olivier & Berthold Crysmann. 2016. Morphology in constraint-based lexicalist approaches to grammar. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology* (Cambridge Handbooks in Language and Linguistics 30), 609–656. Cambridge: Cambridge University Press. DOI: 10.1017/9781139814720.022.

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

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. 2026. Edward A. F. Gibson, Syntax: A cognitive approach. Cambridge, MA: MIT Press, 2025. pp. vii + 366. *Journal of Linguistics*. 1–5. DOI: 10.1017/S0022226726101200.

Borsley, Robert D. & Berthold Crysmann. 2024. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 2nd edn. (Empirically Oriented Theoretical Morphology and Syntax 9), 571–634. Berlin: Language Science Press. DOI: 10.5281/zenodo.13644930.

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

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. 1st edn. (Blackwell Textbooks in Linguistics 16). Oxford: Blackwell Publishers Ltd.

Bruggeman, Laurence, Evan Kidd, Rachel Nordlinger & Anne Cutler. 2025. Incremental processing in a polysynthetic language (Murrinhpatha). *Cognition* 257. 1–7. DOI: 10.1016/j.cognition.2025.106075.

Chomsky, Noam. 1957. *Syntactic structures* (Janua Linguarum / Series Minor 4). Berlin: Mouton de Gruyter. DOI: 10.1515/9783112316009.

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

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

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: MIT Press.

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

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* (Current studies in linguistics series 45), 133–166. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/9780262062787.003.0007.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.

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

da Cunha, Yanis & Edward Gibson. 2025. Syntactic complexity phenomena are better explained without empty elements mediating long-distance dependencies. *Cognitive Science* 49(8). e70088. DOI: 10.1111/cogs.70088.

Davis, Anthony R. & Jean-Pierre Koenig. 2024. The nature and role of the lexicon in HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 2nd edn. (Empirically Oriented Theoretical Morphology and Syntax 9), 133–184. Berlin: Language Science Press. DOI: 10.5281/zenodo.13644994.

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

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

Eroms, Hans-Werner. 2000. *Syntax der deutschen Sprache* (de Gruyter Studienbuch). Berlin: Walter de Gruyter. DOI: 10.1515/9783110808124.

Fanselow, Gisbert. 1987. *Konfigurationalität* (Studien zur deutschen Grammatik 29). Tübingen: Gunter Narr Verlag.

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

Fordham, Andrew & Matthew Crocker. 1994. Parsing with principles and probabilities. In Judith L. Klavans & Philip Resnik (eds.), *The balancing act: Combining symbolic and statistical approaches to language*, 37–41. Las Cruces, NM: Association for Computational Linguistics.

Frazier, Lyn & Charles Clifton Jr. 1996. *Construal* (Language, Speech, and Communication 3). Cambridge, MA: MIT Press.

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

Freudenthal, Daniel, Julian M. Pine & Fernand Gobet. 2009. Simulating the referential properties of Dutch, German, and English root infinitives in MOSAIC. *Language Learning and Development* 5(1). 1–29. DOI: 10.1080/15475440802502437.

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

Gaifman, Haim. 1965. Dependency systems and phrase"=structure systems. *Information and Control* 8(3). 304–397. DOI: 10.1016/S0019-9958(65)90232-9.

Gibson, Edward. 2026. Dependency syntax as the simplest theory of grammar. *Trends in Cognitive Sciences*. DOI: 10.1016/j.tics.2026.03.002.

Gibson, Edward A. F. 2025. *Syntax: A cognitive approach*. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/15535.001.0001.

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

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

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language* (Oxford Linguistics). Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199268511.001.0001.

Goldberg, Adele E. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton, NJ: Princeton University Press.

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

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

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

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

Holmberg, Anders. 2015. Verb second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 1 (Handbooks of Linguistics and Communication Science 42), 342–383. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110377408.342.

Hornstein, Norbert, Jairo Nunes & Kleantes K. Grohmann. 2005. *Understanding Minimalism* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511840678.

Huddleston, Rodney & Geoffrey K. Pullum (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781316423530.

Hudson, Richard. 2000. Discontinuity. *Dependency Grammars, TAL* 41(1). 15–56.

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. 2018. Pied piping in cognition. *Journal of Linguistics* 54(1). 85–138. DOI: 10.1017/S0022226717000056.

Hudson, Richard. 2021. HPSG and Dependency Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax 9), 1447–1495. Berlin: Language Science Press. DOI: 10.5281/zenodo.5599880.

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

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

Johnson, Mark. 1986. A GPSG account of VP structure in German. *Linguistics* 24(5). 871–882. DOI: 10.1515/ling.1986.24.5.871.

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

Joshi, Aravind K. 1985. Tree Adjoining Grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In David R. Dowty, Lauri Karttunen & Arnold Zwicky (eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (Studies in Natural Language Processing 1), 206–250. Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511597855.007.

Kahane, Sylvain & Timothy Osborne. 2015. Translators' introduction. In *Elements of structural syntax*, xxix–lxxiii. Translated by Timothy Osborne and Sylvain Kahane. Amsterdam: John Benjamins Publishing Co. DOI: 10.1075/z.185.

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

Kathol, Andreas. 2001. Positional effects in a monostratal grammar of German. *Journal of Linguistics* 37(1). 35–66. DOI: 10.1017/S0022226701008805.

Klenk, Ursula. 2003. *Generative Syntax* (Narr Studienbücher). Tübingen: Gunter Narr Verlag.

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 in Natural Language and Linguistic Theory 22), 251–301. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-011-3446-0_8.

Kornai, András & Geoffrey K. Pullum. 1990. The X-bar Theory of phrase structure. *Language* 66(1). 24–50. DOI: 10.2307/415278.

Koster, Jan. 1978. *Locality principles in syntax* (Studies in Generative Grammar 5). Dordrecht: Foris Publications. DOI: 10.1515/9783110882339.

Koster, Jan. 1987. *Domains and dynasties: The radical autonomy of syntax* (Studies in Generative Grammar 30). Dordrecht: Foris Publications. DOI: 10.1515/9783110808520.

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

Lohnstein, Horst. 1993. *Projektion und Linking: Ein prinzipienbasierter Parser fürs Deutsche* (Linguistische Arbeiten 287). Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783111355733.

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* 21(1). 55–84. DOI: 10.9793/elsj1984.21.55.

Michaelis, Jens. 2001. *On formal properties of Minimalist Grammars*. Universität Potsdam. (Doctoral dissertation).

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

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

Müller, Stefan. 2004a. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation* 2(2). Special Issue on Linguistic Theory and Grammar Implementation, 209–257. DOI: 10.1023/B:ROLC.0000016785.49729.d7.

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

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

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

Müller, Stefan. 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. 2018. *A lexicalist account of argument structure: Template-based phrasal LFG approaches and a lexical HPSG alternative* (Conceptual Foundations of Language Science 2). Berlin: Language Science Press. DOI: 10.5281/zenodo.1441351.

Müller, Stefan. 2019. Evaluating theories: Counting nodes and the question of constituency. *Language Under Discussion* 5(1). 52–67. DOI: 10.31885/lud.5.1.226.

Müller, Stefan. 2022. Headless in Berlin: Headless (nominal) structures in Head-Driven Phrase Structure Grammar. In Ulrike Freywald, Horst J. Simon & Stefan Müller (eds.), *Headedness and/""or grammatical anarchy?* (Empirically Oriented Theoretical Morphology and Syntax 11), 73–121. Berlin: Language Science Press. DOI: 10.5281/zenodo.7142691.

Müller, Stefan. 2023a. *German clause structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Revise and resubmit Language Science Press.

Müller, Stefan. 2023b. *Germanic syntax: A constraint-based view* (Textbooks in Language Sciences 12). Berlin: Language Science Press. DOI: 10.5281/zenodo.7733033.
Müller, Stefan. 2023c. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 5th edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI: 10.5281/zenodo.7628029.
Müller, Stefan. 2024a. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 2nd edn. (Empirically Oriented Theoretical Morphology and Syntax 9), 391–441. Berlin: Language Science Press. DOI: 10.5281/zenodo.13644960.
Müller, Stefan. 2024b. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 2nd edn. (Empirically Oriented Theoretical Morphology and Syntax 9), 1497–1553. Berlin: Language Science Press. DOI: 10.5281/zenodo.5599882.
Müller, Stefan, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.). 2024. *Head-Driven Phrase Structure Grammar: The handbook*. 2nd edn. (Empirically Oriented Theoretical Morphology and Syntax 9). Berlin: Language Science Press. DOI: 10.5281/zenodo.13637708.
Müller, Stefan & Stephen Wechsler. 2014a. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI: 10.1515/tl-2014-0001.
Müller, Stefan & Stephen Wechsler. 2014b. Two sides of the same slim Boojum: Further arguments for a lexical approach to argument structure. *Theoretical Linguistics* 40(1–2). 187–224. DOI: 10.1515/tl-2014-0009.
Nerbonne, John. 1986. 'Phantoms' and German fronting: Poltergeist constituents? *Linguistics* 24(5). 857–870. DOI: 10.1515/ling.1986.24.5.857.
Nerbonne, John, Klaus Netter & Carl Pollard (eds.). 1994. *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46). Stanford, CA: CSLI Publications.
Osborne, Timothy. 2018. Tests for constituents: What they really reveal about the nature of syntactic structure. *Language Under Discussion* 5(1). 1–41. DOI: 10.31885/lud.5.1.223.
Osborne, Timothy. 2019. *A Dependency Grammar of English: An introduction and beyond*. Amsterdam: John Benjamins Publishing Company. DOI: doi:10.1075/z.224.
Osborne, Timothy, Michael Putnam & Thomas M. Groß. 2011. Bare Phrase Structure, labelless trees, and specifier-less syntax: Is Minimalism becoming a Dependency Grammar? *The Linguistic Review* 28(3). 315–364.
Peters, P. Stanley & Robert W. Ritchie. 1973. On the generative power of Transformational Grammar. *Information Sciences* 6. 49–83. DOI: 10.1016/0020-0255(73)90027-3.
Pickering, Martin & Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* 6(3). 229–259. DOI: 10.1080/01690969108406944.
Pollard, Carl. 1997. The nature of constraint-based grammar. *Linguistic Research* 15. 1–18. http://isli.khu.ac.kr/journal/content/data/15/1.pdf (10 February, 2021).
Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.
Przepiórkowski, Adam. 2026. *Model-theoretic Dependency Grammar*. https://ling.auf.net/lingbuzz/010221.

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

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545. DOI: 10.1353/lan.2010.0002.

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

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

Scherpenisse, Wim. 1986. *The connection between base structure and linearization restrictions in German and Dutch* (Europäische Hochschulschriften, Reihe XXI, Linguistik 47). Frankfurt/M.: Peter Lang.

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.

Stabler, Edward P. 2011. After Governement and Binding Theory. In Johan van Benthem & Alice ter Meulen (eds.), *Handbook of logic and language*, 2nd edn. (Elsevier Insights), 395–414. Amsterdam: Elsevier. DOI: 10.1016/C2010-0-65666-5.

Steedman, Mark. 2000. *The syntactic process* (Language, Speech, and Communication 24). Cambridge, MA: MIT Press.

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

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. DOI: 10.7551/mitpress/1479.003.0029.

Tesnière, Lucien. 1959. *Eléments de syntaxe structurale*. Paris: Librairie C. Klincksieck. Republished as *Elements of structural syntax*. Translated by Timothy Osborne and Sylvain Kahane. Amsterdam: John Benjamins Publishing Co., 2015. DOI: 10.1075/z.185.

Tesnière, Lucien. 2015. *Elements of structural syntax*. Translated by Timothy Osborne and Sylvain Kahane. Amsterdam: John Benjamins Publishing Co. DOI: 10.1075/z.185.

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

Torr, John. 2019. *Wide-coverage statistical parsing with Minimalist Grammars*. Edinburgh University. (Doctoral dissertation). http://hdl.handle.net/1842/36215 (1 February, 2020).

Torr, John & Edward P. Stabler. 2016. Coordination in Minimalist Grammars: Excorporation and across the board (head) movement. In David Chiang & Alexander Koller (eds.), *Proceedings of the 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12)*, 1–17. Düsseldorf. https://www.aclweb.org/anthology/W16-3301 (10 February, 2021).

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

van de Velde, Marc. 1978. Zur mehrfachen Vorfeldbesetzung im Deutschen. In Maria-Elisabeth Conte, Anna Giacalone Ramat & Paolo Ramat (eds.), *Wortstellung und Bedeutung: Akten des 12. Linguistischen Kolloquiums, Pavia 1977* (Linguistische Arbeiten 61), 131–141. Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783111713083.131.

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

Wasow, Thomas. 2024. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 2nd edn. (Empirically Oriented Theoretical Morphology and Syntax 9), 1153–1180. Berlin: Language Science Press. DOI: 10.5281/zenodo.13644940.

Weisweber, Wilhelm & Susanne Preuss. 1992. Direct parsing with metarules. In Antonio Zampolli (ed.), *Proceedings of the 15th International Conference on Computational Linguistics (COLING '92), August 23–28*, 1111–1115. Nantes, France: Association for Computational Linguistics.

Zwicky, Arnold M. 1986. Concatenation and liberation. In Anne M. Farley, Peter T. Farley & Karl-Eric McCullough (eds.), *Papers from the 22nd Regional Meeting of the Chicago Linguistic Society*, 65–74. Chicago, IL: Chicago Linguistic Society.