18 Binary branching, locality, and recursion

This chapter discusses three points: section 18.1 deals with the question whether all linguistic structures should be binary branching or not. Section 18.2 discusses the question what information should be available for selection, that is, whether governing heads can access the internal structure of selected elements or whether everything should be restricted to local selection. Finally, Section 18.3 discusses recursion and how/whether it is captured in the different grammar theories that are discussed in this book.

18.1 Binary branching

We have seen that the question of the kind of branching structures assumed has received differing treatments in various theories. Classical $\overline{\text{X}}$ theory assumes that a verb is combined with all its complements. In later variants of GB, all structures are strictly binary branching. Other frameworks do not treat the question of branching in a uniform way: there are proposals that assume binary branching structures and others that opt for flat structures.

Haegeman (1994: Section 2.5) uses learnability arguments (rate of acquisition, see Section 13.2 on this point). She discusses the example in (1) and claims that language learners have to choose one of eight structures if flat-branching structures can occur in natural language. If, on the other hand, there are only binary-branching structures, then the sentence in (1) cannot have the structures in Figure 18.1 to start with, and therefore a learner would not have to rule out the corresponding hypotheses.

(1) Mummy must leave now.

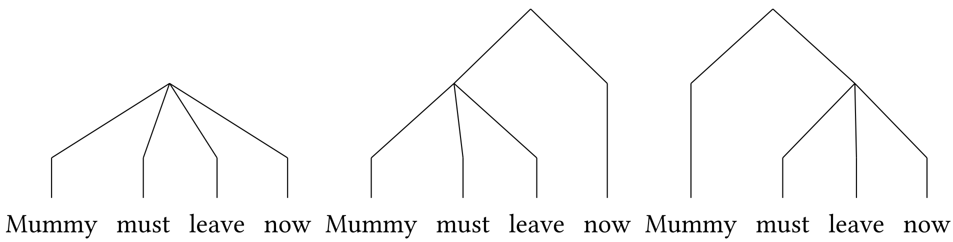

Figure 18.1: Structures with partial flat-branching

However, Haegeman (1994: 88) provides evidence for the fact that (1) has the structure in (2):

(2) [Mummy [must [leave now]]]

The relevant tests showing this included elliptical constructions, that is, the fact that it is possible to refer to the constituents in (2) with pronouns. This means that there is actually evidence for the structure of (1) that is assumed by linguists and we therefore do not have to assume that it is just hard-wired in our brains that only binary-branching structures are allowed. Haegeman (1994: 143) mentions a consequence of the binary branching hypothesis: if all structures are binary-branching, then it is not possible to straightforwardly account for sentences with ditransitive verbs in $\overline{\mathrm{X}}$ theory. In $\overline{\mathrm{X}}$ theory, it is assumed that a head is combined with all its complements at once (see Section 2.5). So in order to account for ditransitive verbs in $\overline{\mathrm{X}}$ theory, an empty element (little *v*) has to be assumed (see Section 4.1.4).

It should have become clear in the discussion of the arguments for the Poverty of the Stimulus in Section 13.8 that the assumption that only binary-branching structures are possible is part of our innate linguistic knowledge is nothing more than pure speculation. Haegeman offers no kind of evidence for this assumption. As shown in the discussions of the various theories we have seen, it is possible to capture the data with flat structures. For example, it is possible to assume that, in English, the verb is combined with its complements in a flat structure (Pollard & Sag 1994: 39). There are sometimes theory-internal reasons for deciding for one kind of branching or another, but these are not always applicable to other theories. For example, Binding Theory in GB theory is formulated with reference to dominance relations in trees (Chomsky 1981a: 188). If one assumes that syntactic structure plays a crucial role for the binding of pronouns (see page 92), then it is possible to make assumptions about syntactic structure based on the observable binding relations. Binding data have, however, received a very different treatment in various theories. In LFG, constraints on f-structure are used for Binding Theory (Dalrymple 1993), whereas Binding Theory in HPSG operates on argument structure lists (valence information that are ordered in a particular way, see Section 9.6.1).

The opposite of Haegeman's position is the argumentation for flat structures put forward by Croft (2001: Section 1.6.2). In his Radical Construction Grammar FAQ, Croft observes that a phrasal construction such as the one in (3a) can be translated into a Categorial Grammar lexical entry like (3b).

(3) a. [$_{\mathrm{VP}}$ V NP]
b. VP/NP

He claims that a disadvantage of Categorial Grammar is that it only allows for binary-branching structures and yet there exist constructions with more than two parts (p. 49). The exact reason why this is a problem is not explained, however. He even acknowledges himself that it is possible to represent constructions with more than two arguments in Categorial Grammar. For a ditransitive verb, the entry in Categorial Grammar of English would take the form of (4):

(4) ((s\np)/np)/np

If we consider the elementary trees for TAG in Figure 18.2, it becomes clear that it is equally possible to incorporate semantic information into a flat tree and a binary-branching tree. The binary-branching tree corresponds to a Categorial Grammar deriva-

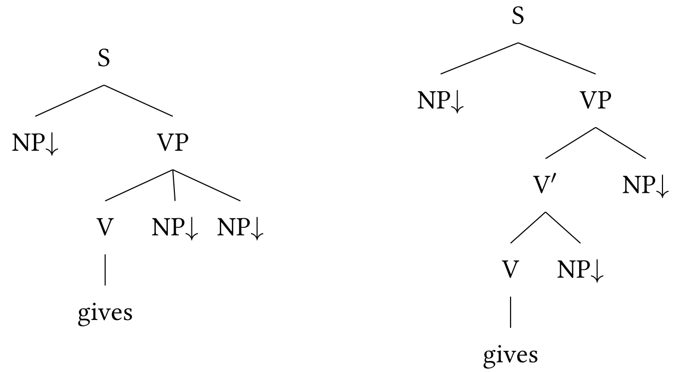

Figure 18.2: Flat and binary-branching elementary trees

tion. In both analyses in Figure 18.2, a meaning is assigned to a head that occurs with a certain number of arguments. Ultimately, the exact structure required depends on the kinds of restrictions on structures that one wishes to formulate. In this book, such restrictions were not discussed, but we have seen some theories model binding relations with reference to tree structures. Reflexive pronouns must be bound within a particular local domain inside the tree. In theories such as LFG and HPSG, these binding restrictions are formulated without any reference to trees. This means that evidence from binding data for one of the structures in Figure 18.2 (or for other tree structures) constitutes nothing more than theory-internal evidence.

Another reason to assume trees with more structure is the possibility to insert adjuncts on any node. In Chapter 9, an HPSG analysis for German that assumes binary-branching structures was proposed. With this analysis, it is possible to attach an adjunct to any node and thereby explain the free ordering of adjuncts in the middle field:

(5) a. [weil] der Mann der Frau das Buch *gestern* gab
because the man the woman the book yesterday gave
'because the man gave the woman the book yesterday'

b. [weil] der Mann der Frau *gestern* das Buch gab
because the man the woman yesterday the book gave

c. [weil] der Mann *gestern* der Frau das Buch gab
because the man yesterday the woman the book gave

d. [weil] *gestern* der Mann der Frau das Buch gab
because yesterday the man the woman the book gave

This analysis is not the only one possible, however. One could also assume an entirely flat structure where arguments and adjuncts are dominated by one node. Kasper (1994) suggests this kind of analysis in HPSG (see also Section 5.1.5 for GPSG analyses that make use of metarules for the introduction of adjuncts). Kasper requires complex relational constraints that create syntactic relations between elements in the tree and also compute the semantic contribution of the entire constituent using the meaning of both the verb and the adjuncts. The analysis with binary-branching structures is simpler than those with complex relational constraints and – in the absence of theory-external evidence for flat structures – should be preferred to the analysis with flat structures. At this point, one could object that adjuncts in English cannot occur in all positions between arguments and therefore the binary-branching Categorial Grammar analysis and the TAG analysis in Figure 18.2 are wrong. This is not correct, however, as it is the specification of adjuncts with regard to the adjunction site that is crucial in Categorial Grammar. An adverb has the category (s\np)\(s\np) or (s\np)/(s\np) and can therefore only be combined with constituents that correspond to the VP node in Figure 18.2. In the same way, an elementary tree for an adverb in TAG can only attach to the VP node (see Figure 12.3 on page 413). For the treatment of adjuncts in English, binary-branching structures therefore do not make any incorrect predictions.

18.2 Locality

The question of local accessibility of information has been treated in various ways by the theories discussed in this book. In the majority of theories, one tries to make information about the inner workings of phrases inaccessible for adjacent or higher heads, that is, *glaubt* 'believe' in (6) selects a sentential argument but it cannot "look inside" this sentential argument.

(6) a. Karl glaubt, dass morgen seine Schwester kommt.
Karl believes that tomorrow his sister comes
'Karl believes that his sister is coming tomorrow.'

b. Karl glaubt, dass seine Schwester morgen kommt.
Karl believes that his sister tomorrow comes

Thus for example, *glauben* cannot enforce that the subject of the verb has to begin with a consonant or that the complementizer has to be combined with a verbal projection starting with an adjunct. In Section 1.5, we saw that it is a good idea to classify constituents in terms of their distribution and independent of their internal structure. If we are talking about an NP box, then it is not important what this NP box actually contains. It is only of importance that a given head wants to be combined with an NP with a particular case marking. This is called *locality of selection*.

Various linguistic theories have tried to implement locality of selection. The simplest form of this implementation is shown by phrase structure grammars of the kind discussed in Chapter 2. The rule in (17) on page 61, repeated here as (7), states that a ditransitive verb can occur with three noun phrases, each with the relevant case:

(7) S → NP(Per1,Num1,nom)
NP(Per2,Num2,dat)
NP(Per3,Num3,acc)
V(Per1,Num1,ditransitive)

Since the symbols for NPs do not have any further internal structure, the verb cannot require that there has to be a relative clause in an NP, for example. The internal properties of the NP are not visible to the outside. We have already seen in the discussion in Chapter 2, however, that certain properties of phrases have to be outwardly visible. This was the information that was written on the boxes themselves. For noun phrases, at least information about person, number and case are required in order to correctly capture their relation to a head. The gender value is important in German as well, since adverbial phrases such as *einer nach dem anderen* 'one after the other' have to agree in gender with the noun they refer to (see example (12) on page 503). Apart from that, information about the length of the noun phrases is required, in order to determine their order in a clause. Heavy constituents are normally ordered after lighter ones, and are also often extraposed (cf. Behaghel's *Gesetz der wachsenden Glieder* 'Law of increasing constituents' (1909: 139; 1930: 86)).

Theories that strive to be as restrictive as possible with respect to locality therefore have to develop mechanisms that allow one to only access information that is required to explain the distribution of constituents. This is often achieved by projecting certain properties to the mother node of a phrase. In $\overline{\text{X}}$ theory, the part of speech a head belongs to is passed up to the maximal projection: if the head is an N, for example, then the maximal projection is an NP. In GPSG, HPSG and variants of CxG, there are Head Feature Principles responsible for the projection of features. Head Feature Principles ensure that an entire group of features, so-called head features, are present on the maximal projection of a head. Furthermore, every theory has to be capable of representing the fact that a constituent can lack one of its parts and this part is then realized via a long-distance dependency in another position in the clause. As previously discussed on page 298, there are languages in which complementizers inflect depending on whether their complement is missing a constituent or not. This means that this property must be somehow accessible. In GPSG, HPSG and variants of CxG, there are additional groups of features that are present at every node between a filler and a gap in a long-distance dependency. In LFG, there is f-structure instead. Using Functional Uncertainty, one can look for the position in the f-structure where a particular constituent is missing. In GB theory, movement proceeds cyclically, that is, an element is moved into the specifier of CP and can be moved from there into the next highest CP. It is assumed in GB theory that heads can look inside their arguments, at least they can see the elements in the specifier position. If complementizers can access the relevant specifier positions, then they can determine whether something is missing from an embedded phrase or not. In GB theory, there was also an analysis of case assignment in infinitive constructions in which the case-assigning verb governs into the embedded phrase and assigns case to the element in SpecIP. Figure 18.3 shows the relevant structure taken from Haegeman (1994: 170). Since the Case Principle is formulated in such a way that only finite I can assign

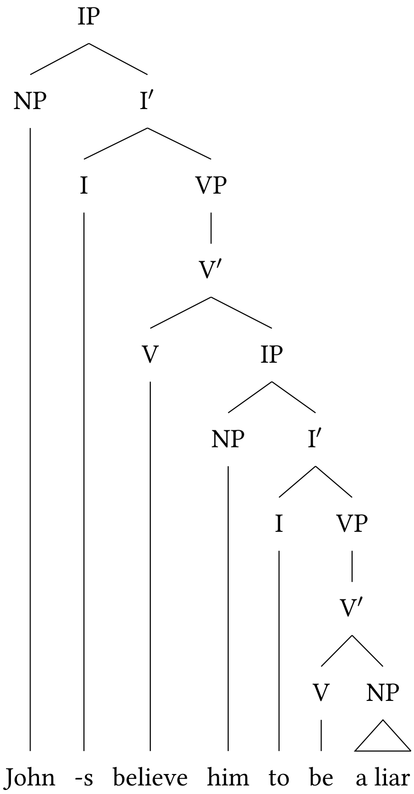

Figure 18.3: Analysis of the AcI construction with *Exceptional Case Marking*

case to the subject (cf. page 113), *him* does not receive case from I. Instead, it is assumed that the verb *believe* assigns case to the subject of the embedded infinitive.

Verbs that can assign case across phrase boundaries are referred to as ECM verbs, where ECM stands for *Exceptional Case Marking*. As the name suggests, this instance of case assignment into a phrase was viewed as an exception. In newer versions of the theory (e. g. Kratzer 1996: 120–123), all case assignment is to specifier positions. For example, the Voice head in Figure 18.4 on the facing page assigns accusative to the DP in the specifier of VP. Since the Voice head governs into the VP, case assignment to a run-of-the-mill object in this theory is an instance of exceptional case assignment as well. The same is true in Adger's version of Minimalism, which was discussed in Chapter 4: Adger (2010) argues that his theory is more restrictive than LFG or HPSG since it is only one feature that can be selected by a head, whereas in LFG and HPSG complex feature bundles are selected. However, the strength of this kind of locality constraint is weakened by the operation Agree, which allows for nonlocal feature checking. As in

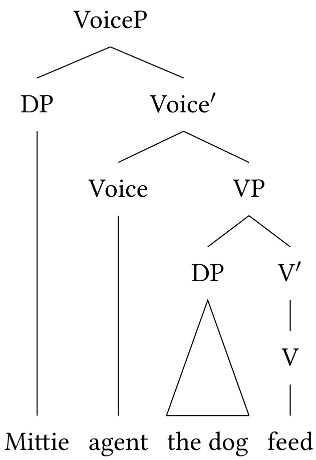

Figure 18.4: Analysis of structures with a transitive verb following Kratzer

Kratzer's proposal, case is assigned nonlocally by little *v* to the object inside the VP (see Section 4.1.5.2).

Adger discusses PP arguments of verbs like *depend* and notes that these verbs need specific PPs, that is, the form of the preposition in the PP has to be selectable. While this is trivial in Dependency Grammar, where the preposition is selected right away, the respective information is projected in theories like HPSG and is then selectable at the PP node. However, this requires that the governing verb can determine at least two properties of the selected element: its part of speech and the form of the preposition. This is not possible in Adger's system and he left this for further research. Of course it would be possible to assume an onP (a phrasal projection of *on* that has the category 'on'). Similar solutions have been proposed in Minimalist theories (see Section 4.6.1 on functional projections), but such a solution would obviously miss the generalization that all prepositional phrases have something in common, which would not be covered in a system with atomic categories that are word specific.

In theories such as LFG and HPSG, case assignment takes place locally in constructions such as those in (8):

(8) a. John believes him to be a liar.

b. Ich halte ihn für einen Lügner.
I hold him for a.ACC liar
'I take him to be a liar.'

c. Er scheint ein Lügner zu sein.
he seems a.NOM liar to be
'He seems to be a liar.'

d. Er fischt den Teich leer.
he fishes the.ACC pond empty
'He fishes (in) the pond (until it is) empty.'

Although *him*, *ihn* 'him', *er* 'he' and *den Teich* 'the pond' are not semantic arguments of the finite verbs, they are syntactic arguments (they are raised) and can therefore be assigned case locally. See Bresnan (1982a: 348–349 and Section 8.2) and Pollard & Sag (1994: Section 3.5) for an analysis of raising in LFG and HPSG respectively. See Meurers (1999c), Przepiórkowski (1999b), and Müller (2007b: Section 17.4) for case assignment in HPSG and for its interaction with raising.

There are various phenomena that are incompatible with strict locality and require the projection of at least some information. For example, there are question tags in English that must match the subject of the clause with which they are combined:

(9) a. She is very smart, isn't she / * he?
b. They are very smart, aren't they?

Bender & Flickinger (1999), Flickinger & Bender (2003) therefore propose making information about agreement or the referential index of the subject available on the sentence node.[1] In Sag (2007), all information about phonology, syntax and semantics of the subject is represented as the value of a feature XARG (EXTERNAL ARGUMENT). Here, *external argument* does not stand for what it does in GB theory, but should be understood in a more general sense. For example, it makes the possessive pronoun accessible on the node of the entire NP. Sag (2007) argues that this is needed to force coreference in English idioms:

(10) a. He$_i$ lost [his$_i$ / *her$_j$ marbles].
b. They$_i$ kept/lost [their$_i$ / *our$_j$ cool].

The use of the XARG feature looks like an exact parallel to accessing the specifier position as we saw in the discussion of GB. However, Sag proposes that complements of prepositions in Polish are also made accessible by XARG since there are data suggesting that higher heads can access elements inside PPs (Przepiórkowski 1999a: Section 5.4.1.2).

In Section 10.6.2 about Sign-based Construction Grammar, we already saw that a theory that only makes the reference to one argument available on the highest node of a projection cannot provide an analysis for idioms of the kind given in (11). This is because the subject is made available with verbal heads, however, it is the object that needs to be accessed in sentences such as (11). This means that one has to be able to formulate constraints affecting larger portions of syntactic structure.

(11) a. Ich glaube, mich / # dich tritt ein Pferd.[2]
I believes me you kicks a horse
'I am utterly surprised.'

[1] See also Sag & Pollard (1991: 89).
[2] Richter & Sailer (2009: 311).

b. Jonas glaubt, ihn tritt ein Pferd.[3]
Jonas believes him kicks a horse
'Jonas is utterly surprised.'

c. # Jonas glaubt, dich tritt ein Pferd.
Jonas believes you kicks a horse
'Jonas believes that a horse kicks you.'

Theories of grammar with extended locality domains do not have any problems with this kind of data.[4] An example for this kind of theory is TAG. In TAG, one can specify trees of exactly the right size (Abeillé 1988; Abeillé & Schabes 1989). All the material that is fixed in an idiom is simply determined in the elementary tree. Figure 18.5 shows the tree for *kick the bucket* as it is used in (12a).

(12) a. The cowboys kicked the bucket.
b. Cowboys often kick the bucket.
c. He kicked the proverbial bucket.

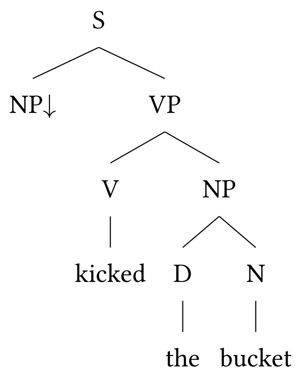

Figure 18.5: Elementary tree for *kick the bucket*

Since TAG trees can be split up by adjunction, it is possible to insert elements between the parts of an idiom as in (12b,c) and thus explain the flexibility of idioms with regard to adjunction and embedding.[5] Depending on whether the lexical rules for the passive and long-distance dependencies can be applied, the idiom can occur in the relevant variants.

[3] http://www.machandel-verlag.de/der-katzenschatz.html, 06.07.2015.

[4] Or more carefully put: they do not have any serious problems since the treatment of idioms in all their many aspects is by no means trivial (Sailer 2000).

[5] Interestingly, variants of Embodied CxG are strikingly similar to TAG. The Ditransitive Construction that was discussed on page 338 allows for additional material to occur between the subject and the verb.

The problems that arise for the semantics construction are also similar. Abeillé & Schabes (1989: 9) assume that the semantics of *John kicked the proverbial bucket* is computed from the parts *John′*, *kick-the-bucket′* and *proverbial′*, that is, the added modifiers always have scope over the entire idiom. This is not

In cases where the entire idiom or parts of the idiom are fixed, it is possible to rule out adjunction to the nodes of the idiom tree. Figure 18.6 shows a pertinent example from Abeillé & Schabes (1989: 7). The ban on adjunction is marked by a subscript NA.

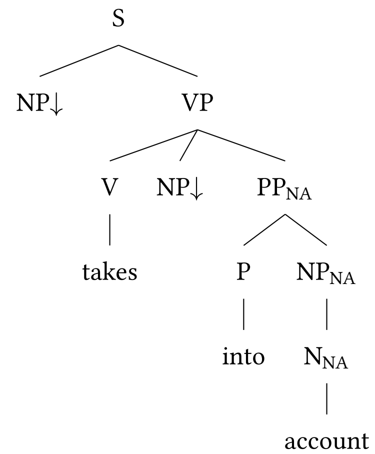

Figure 18.6: Elementary tree for *take into account*

The question that also arises for other theories is whether the efforts that have been made to enforce locality should be abandoned altogether. In our box model in Section 1.5, this would mean that all boxes were transparent. Since plastic boxes do not allow all of the light through, objects contained in multiple boxes cannot be seen as clearly as those in the top-most box (the path of Functional Uncertainty is longer). This is parallel to a suggestion made by Kay & Fillmore (1999) in CxG. Kay and Fillmore explicitly represent all the information about the internal structure of a phrase on the mother node and therefore have no locality restrictions at all in their theory. In principle, one can motivate this kind of theory in parallel to the argumentation in Chapter 17. The argument there made reference to the complexity of the grammatical formalism: the kind of complexity that the language of description has is unimportant, it is only important what one does with it. In the same way, one can say that regardless of what kind of information is in

adequate for all idioms (Fischer & Keil 1996):

(i) Er band ihr einen großen Bären auf.
he tied her a big bear on
'He pulled (a lot of) wool over her eyes.'

In the idiom in (i), *Bär* 'bear' actually means 'lie' and the adjective has to be interpreted accordingly. The relevant tree should therefore contain nodes that contribute semantic information and also say something about the composition of these features.

In the same way, when computing the semantics of noun phrases in TAG and Embodied Construction Grammar, one should bear in mind that the adjective that is combined with a discontinuous NP Construction (see page 335) or an NP tree can have narrow scope over the noun (*all alleged murderers*).

principle accessible, it is not accessed if this is not permitted. This was the approach taken by Pollard & Sag (1987: 143–145).

It is also possible to assume a world in which all the boxes contain transparent areas where it is possible to see parts of their contents. This is more or less the LFG world: the information about all levels of embedding contained in the f-structure is visible to both the inside and the outside. We have already discussed Nordlinger's (1998) LFG analysis of Wambaya on page 304. In Wambaya, words that form part of a noun phrase can be distributed throughout the clause. For example, an adjective that refers to a noun can occur in a separate position from it. Nordlinger models this by assuming that an adjective can make reference to an argument in the f-structure and then agrees with it in terms of case, number and gender. Bender (2008c) has shown that this analysis can be transferred to HPSG: instead of no longer representing an argument on the mother node after it has been combined with a head, simply marking the argument as realized allows us to keep it in the representation (Meurers 1999c; Przepiórkowski 1999b; Müller 2007b: Section 17.4). Detmar Meurers compares both of these HPSG approaches to different ways of working through a shopping list: in the standard approach taken by Pollard & Sag (1994), one tears away parts of the shopping list once the relevant item has been found. In the other case, the relevant item on the list is crossed out. At the end of the shopping trip, one ends up with a list of what has been bought as well as the items themselves.

I have proposed the crossing-out analysis for depictive predicates in German and English (Müller 2004a, 2008a). Depictive predicates say something about the state of a person or object during the event expressed by a verb:

(13) a. Er sah sie nackt.[6]
he saw her naked

b. He saw her naked.

In (13), the depictive adjective can either refer to the subject or the object. However, there is a strong preference for readings where the antecedent noun precedes the depictive predicate (Lötscher 1985: 208). Figure 18.7 on the next page shows analyses for the sentences in (14):

(14) a. dass er$_i$ die Äpfel$_j$ ungewaschen$_{i/j}$ isst
that he the apples unwashed eats
'that he eats the apples unwashed'

b. dass er$_i$ ungewaschen$_{i/*j}$ die Äpfel$_j$ isst
that he unwashed the apples eats
'that he eats the apples (while he is) unwashed'

Arguments that have been realized are still represented on the upper nodes, however, they are crossed-out and thereby marked as "realized". In German, this preference for the antecedent noun can be captured by assuming a restriction that states that the antecedent noun must not yet have been realized.

It is commonly assumed for English that adjuncts are combined with a VP.

[6] Haider (1985a: 94).

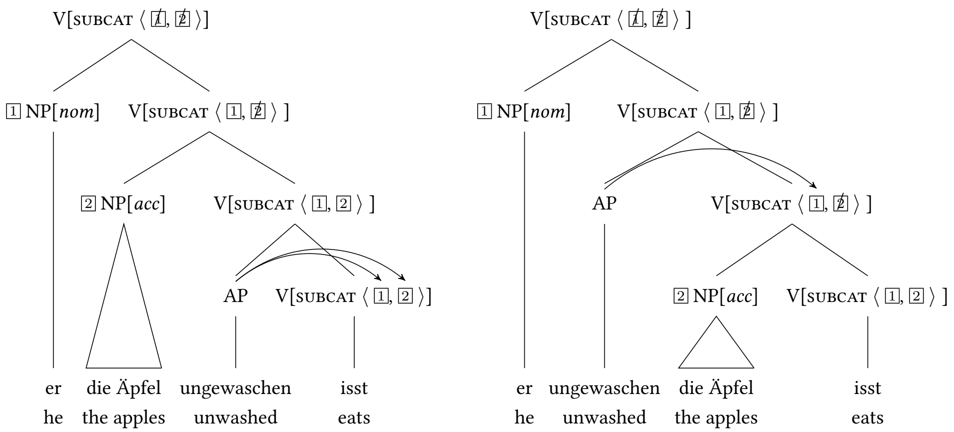

Figure 18.7: Analysis of *dass er die Äpfel ungewaschen isst* 'that he the apples unwashed eats' and *dass er ungewaschen die Äpfel isst* 'that he unwashed the apples eat'

(15) a. John [[$_{\text{VP}}$ ate the apples$_i$] unwashed$_i$].

b. You can't [[$_{\text{VP}}$ give them$_i$ injections] unconscious$_i$].[7]

In approaches where the arguments of the verb are accessible at the VP node, it is possible to establish a relation between the depictive predicate and an argument although the antecedent noun is inside the VP. English differs from German in that depictives can refer to both realized (*them* in (15b)) and unrealized (*you* in (15b)) arguments.

Higginbotham (1985: 560) and Winkler (1997) have proposed corresponding non-cancellation approaches in GB theory. There are also parallel suggestions in Minimalist theories: checked features are not deleted, but instead marked as already checked (Stabler 2011: 14). However, these features are still viewed as inaccessible.

Depending on how detailed the projected information is, it can be possible to see adjuncts and argument in embedded structures as well as their phonological, syntactic and semantic properties. In the CxG variant proposed by Kay and Fillmore, all information is available. In LFG, information about grammatical function, case and similar properties is accessible. However, the part of speech is not contained in the f-structure. If the part of speech does not stand in a one-to-one relation to grammatical function, it cannot be restricted using selection via f-structure. Nor is phonological information represented completely in the f-structure. If the analysis of idioms requires nonlocal access to phonological information or part of speech, then this has to be explicitly encoded in the f-structure (see Bresnan (1982c: 46–50) for more on idioms).

In the HPSG variant that I adopt, only information about arguments is projected. Since arguments are always represented by descriptions of type *synsem*, no information about their phonological realization is present. However, there are daughters in the structure

[7] Simpson (2005a: 17).

so that it is still possible to formulate restrictions for idioms as in TAG or Construction Grammar (see Richter & Sailer (2009) for an analysis of the 'horse' example in (11a)). This may seem somewhat like overkill: although we already have the tree structure, we are still projecting information about arguments that have already been realized (unfortunately these also contain information about their arguments and so on). At this point, one could be inclined to prefer TAG or LFG since these theories only make use of one extension of locality: TAG uses trees of arbitrary or rather exactly the necessary size and LFG makes reference to a complete f-structure. However, things are not quite that simple: if one wants to create a relation to an argument when adjoining a depictive predicate in TAG, then one requires a list of possible antecedents. Syntactic factors (e. g. reference to dative vs. accusative noun phrases, to argument vs. adjuncts, coordination of verbs vs. nouns) play a role in determining the referent noun, this cannot be reduced to semantic relations. Similarly, there are considerably different restrictions for different kinds of idioms and these cannot all be formulated in terms of restrictions on f-structure since f-structure does not contain information about parts of speech.

One should bear in mind that some phenomena require reference to larger portions of structure. The majority of phenomena can be treated in terms of head domains and extended head domains, however, there are idioms that go beyond the sentence level. Every theory has to account for this somehow.

18.3 Recursion

Every theory in this book can deal with self-embedding in language as it was discussed on page 6. The example (2) is repeated here as (16):

(16) that Max thinks [that Julia knows [that Otto claims [that Karl suspects [that Richard confirms [that Friederike is laughing]]]]]

Most theories capture this directly with recursive phrase structure rules or dominance schemata. TAG is special with regard to recursion since recursion is factored out of the trees. The corresponding effects are created by an adjunction operation that allows any amount of material to be inserted into trees. It is sometimes claimed that Construction Grammar cannot capture the existence of recursive structure in natural language (e. g. Leiss 2009: 269). This impression is understandable since many analyses are extremely surface-oriented. For example, one often talks of a [Sbj TrVerb Obj] construction. However, the grammars in question also become recursive as soon as they contain a sentence embedding or relative clause construction. A sentence embedding construction could have the form [Sbj that-Verb that-S], where a that-Verb is one that can take a sentential complement and that-S stands for the respective complement. A *that*-clause can then be inserted into the that-S slot. Since this *that*-clause can also be the result of the application of this construction, the grammar is able to produce recursive structures such as those in (17):

(17) Otto claims [$_{\text{that-S}}$ that Karl suspects [$_{\text{that-S}}$ that Richard sleeps]].

In (17), both *Karl suspects that Richard sleeps* and the entire clause are instances of the [Sbj that-Verb that-S] construction. The entire clause therefore contains an embedded subpart that is licensed by the same construction as the clause itself. (17) also contains a constituent of the category *that*-S that is embedded inside of *that*-S. For more on recursion and self-embedding in Construction Grammar, see Verhagen (2010).

Similarly, every Construction Grammar that allows a noun to combine with a genitive noun phrase also allows for recursive structures. The construction in question could have the form [Det N NP[gen]] or [N NP[gen]]. The [Det N NP[gen]] construction licenses structures such as (18):

(18)

[$_{\text{NP}}$	des Kragens	[$_{\text{NP}}$ des	Mantels	[$_{\text{NP}}$ der	Vorsitzenden]]]
	the collar	of.the coat		of.the chairwoman	

'the collar of the coat of the chairwoman'

Jurafsky (1996) and Bannard, Lieven & Tomasello (2009) use probabilistic context-free grammars (PCFG) for a Construction Grammar parser with a focus on psycholinguistic plausibility and modeling of acquisition. Context-free grammars have no problems with self-embedding structures like those in (18) and thus this kind of Construction Grammar itself does not encounter any problems with self-embedding.

Goldberg (1995: 192) assumes that the resultative construction for English has the following form:

(19) [SUBJ [V OBJ OBL]]

This corresponds to a complex structure as assumed for elementary trees in TAG. LTAG differs from Goldberg's approach in that every structure requires a lexical anchor, that is, for example (19), the verb would have to be fixed in LTAG. But in Goldberg's analysis, verbs can be inserted into independently existing constructions (see Section 21.1). In TAG publications, it is often emphasized that elementary trees do not contain any recursion. The entire grammar is recursive however, since additional elements can be added to the tree using adjunction and – as (17) and (18) show – insertion into substitution nodes can also create recursive structures.